

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2005-506049

(P2005-506049A)

(43) 公表日 平成17年3月3日(2005.3.3)

(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 38/00	A 6 1 K 39/395 D	4 B O 2 4
A 6 1 K 39/395	A 6 1 K 39/395 N	4 B O 2 9
A 6 1 K 45/00	A 6 1 K 45/00	4 B O 6 3
A 6 1 P 1/04	A 6 1 P 1/04	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2002-578444 (P2002-578444)	(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3月29日 (2002.3.29)	(74) 代理人	100089266 弁理士 大島 陽一
(85) 翻訳文提出日	平成15年9月29日 (2003.9.29)	(72) 発明者	ボーグン、マライア・アール アメリカ合衆国カリフォルニア州94577・サンレアンドロ・サンティアゴロード 14244
(86) 国際出願番号	PCT/US2002/009820		
(87) 国際公開番号	W02002/079441		
(87) 国際公開日	平成14年10月10日 (2002.10.10)		
(31) 優先権主張番号	60/280, 527		
(32) 優先日	平成13年3月30日 (2001.3.30)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/282, 112		
(32) 優先日	平成13年4月6日 (2001.4.6)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/282, 702		
(32) 優先日	平成13年4月9日 (2001.4.9)		
(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 (f) 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:1-25 (配列番号 1 乃至 25) 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

(b) SEQ ID NO:2-24 からなる群から選択した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持つポリペプチド

(c) SEQ ID NO:1 からなるアミノ酸配列に対して少なくとも 96% が同一であるような天然アミノ酸配列を有するポリペプチド

(d) SEQ ID NO:25 からなるアミノ酸配列に対して少なくとも 95% が同一であるような天然アミノ酸配列を有するポリペプチド 10

(e) SEQ ID NO:1-25 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

(f) SEQ ID NO:1-25 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免疫原性断片

【請求項 2】

SEQ ID NO:1-25 からなる群から選択したアミノ酸配列を有する、請求項 1 に記載の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする単離されたポリヌクレオチド。 20

【請求項 4】

請求項 2 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 5】

SEQ ID NO:26-50 からなる群から選択したポリヌクレオチド配列を含む、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。

【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。 30

【請求項 8】

請求項 6 に記載の組換えポリヌクレオチドを含む遺伝形質転換体。

【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、

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

(b) そのように発現した前記ポリペプチドを回収する過程。

【請求項 10】

前記ポリペプチドが、SEQ ID NO:1-25 からなる群から選択したアミノ酸配列を持つことを特徴とする、請求項 9 に記載の方法。 40

【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 (g) 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:26-50 からなる群から選択した或るポリヌクレオチド配列を有するポリヌクレオチド

(b) SEQ ID NO:27-50 からなる群から選択したポリヌクレオチド配列と少なくとも 90% が同一であるような天然のポリヌクレオチド配列を含むポリヌクレオチド

(c) SEQ ID NO:26 のポリヌクレオチド配列に対して少なくとも 96% が同一であるよう 50

な天然ポリヌクレオチド配列を有するポリヌクレオチド

(d) (a) のポリヌクレオチドに相補的なポリヌクレオチド

(e) (b) のポリヌクレオチドに相補的なポリヌクレオチド

(f) (c) のポリヌクレオチドに相補的なポリヌクレオチド

(g) (a) 乃至 (e) のRNA等価物

【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或る配列を有する少なくとも20の連続したヌクレオチド群を持つ或るプローブと前記サンプルとをハイブリダイズし、該プローブが、或るハイブリダイゼーション複合体が前記プローブと前記標的ポリヌクレオチドまたはその断片群との間に形成される条件下で、前記標的ポリヌクレオチドへ特異的にハイブリダイズする過程と、

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

【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、

(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、と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:1-25からなる群から選択されたアミノ酸配列を持つことを特徴とする、請求項17の組成物。

【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有効性を確認するために化合物をスクリーニングする方法であって、

(a) 請求項1のポリペプチドを含むサンプルを化合物に曝露する過程と、

(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とを含むことを特徴とするスクリーニング方法。

【請求項21】

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22】

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23】

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

(a) 請求項1のポリペプチドを含むサンプルを化合物に曝露する過程と、

(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、とを含むことを特徴とするスクリーニング方法。

【請求項24】

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。 10

【請求項25】

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26】

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、

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

(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ステップとを含むことを特徴とする方法。 20

【請求項27】

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

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

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

(c) 試験化合物の存在下での請求項1に記載のポリペプチドの活性を、試験化合物の不存在下での請求項1に記載のポリペプチドの活性と比較する過程とを含み、試験化合物の存在下での請求項1に記載のポリペプチドの活性の変化が、請求項1に記載のポリペプチドの活性を調節する化合物を標示することを特徴とする方法。 30

【請求項28】

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

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

(b) 前記標的ポリヌクレオチドの、改変された発現を検出する過程と、

(c) 可変量の前記化合物の存在下と前記化合物の不存在下で、前記標的ポリヌクレオチドの発現を比較する過程、とを含むことを特徴とする方法。 40

【請求項29】

試験化合物の毒性を算定する方法であって、

(a) 核酸群を有する或る生体サンプルを前記試験化合物で処理する過程と、

(b) 処理した前記生体サンプルの核酸群と、請求項12のポリヌクレオチドの少なくとも20の連続するヌクレオチド群を有する或る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が、前記プローブと前記生体サンプルの或る標的ポリヌクレオチドとの間で或る特異的なハイブリダイゼーション複合体が形成される諸条件下で行われ、前記標的ポリヌクレオチドが、請求項12のポリヌクレオチドのポリヌクレオチド配列またはその断片を有する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、 50

(d) 前記処理された生体サンプル中の前記ハイブリタイゼーション複合体の量を、或る処理されていない生体サンプル中の前記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、前記処理された生体サンプル中の前記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ことを特徴とする方法。

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

10

(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、前記生体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

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

【請求項32】

20

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】

30

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) 抗体応答を誘発する諸条件下で、SEQ ID NO:1-25からなる群から選択した或るアミノ酸配列またはその免疫原性断片を有する或るポリペプチドを用いて或る動物を免疫化する過程と、

(b) 前記動物から抗体を単離する過程と、

(c) 前記単離された抗体をポリペプチドでスクリーニングし、それによって、SEQ ID NO:1-25を有する群から選択したアミノ酸配列を有するポリペプチドに特異的に結合するようなポリクローナル抗体を同定する過程とを含むことを特徴とする方法。

【請求項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。

40

【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ーを含む組成物。

【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) 抗体応答を誘発する諸条件下で、SEQ ID NO:1-25からなる群から選択した或るアミノ酸配列またはその免疫原性断片を有する或るポリペプチドを用いて或る動物を免疫化する過程と、

(b) 前記動物から、抗体を産出する細胞を単離する過程と、

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

50

(d) 前記ハイブリドーマ細胞を培養する過程と、

(e) SEQ ID NO:1-25を有する群から選択したアミノ酸配列を有するポリペプチドに特異的に結合するようなモノクローナル抗体を前記培養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ーを含む組成物。

【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。 10

【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:1-25を有する群から選択したアミノ酸配列を持つポリペプチドをサンプル中に検出する方法であって、

(a) 前記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、或るサンプルと共に請求項11に記載の抗体をインキュベートする過程と、

(b) 特異結合を検出する過程とを含み、該特異結合が、SEQ ID NO:1-25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示唆することを特徴とする方法。 20

【請求項45】

SEQ ID NO:1-25 からなる群から選択したアミノ酸配列を有するポリペプチドを精製する方法であって、

(a) 前記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、或るサンプルと共に請求項11に記載の抗体をインキュベートする過程と、

(b) 前記サンプルから前記抗体を分離し、SEQ ID NO:1-25からなる群から選択した或るアミノ酸配列を有する精製ポリペプチドを得る過程、とを含むことを特徴とする方法。 30

【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ことを特徴とするマイクロアレイ。 30

【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ィールを作成する方法であって、

(a) サンプルのポリヌクレオチドを標識するステップと、

(b) 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、請求項46に記載のマイクロアレイの要素をサンプルの標識されたポリヌクレオチドと接触させるステップと

(c) サンプル中のポリヌクレオチドの発現を定量するステップとを含むサンプルの転写イメージを作製する方法。 40

【請求項48】

固体基板上の固有の物理的位置に付着された種々のヌクレオチド分子を含むアレイであって、少なくとも1つの前記ヌクレオチド分子が、標的ポリヌクレオチドの少なくとも30の連続したヌクレオチドと特異的にハイブリダイズ可能なオリゴヌクレオチドまたはポリヌクレオチドを有し、前記の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ことを特徴とするアレイ。 40

【請求項49】

請求項48に記載のアレイで、前記の最初のオリゴヌクレオチドまたはポリヌクレオチドの配列が前記の標的ポリヌクレオチドの少なくとも30の連続したヌクレオチドに完全に相補的であることを特徴とするアレイ。

【請求項50】

請求項48に記載のアレイで、前記の最初のオリゴヌクレオチドまたはポリヌクレオチドの配列が前記の標的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に完全に相補的であることを特徴とするアレイ。

【請求項51】

請求項48に記載のアレイで、前記の最初のオリゴヌクレオチドまたはポリヌクレオチドの配列が前記の標的ポリヌクレオチドに完全に相補的であることを特徴とするアレイ。

【請求項52】

請求項48に記載のアレイで、マイクロアレイであることを特徴とするアレイ。

【請求項53】

請求項48に記載のアレイで、前記のオリゴヌクレオチドまたはポリヌクレオチドの最初の配列を含むヌクレオチド分子にハイブリダイズした前記の標的ポリヌクレオチドを有することを特徴とするアレイ。 10

【請求項54】

請求項48に記載のアレイで、リンカーが少なくとも1つの前記のヌクレオチド分子と前記の固体基板を連結していることを特徴とするアレイ。

【請求項55】

請求項48に記載のアレイで、基板上の固有の物理的位置の各々が複数のヌクレオチド分子を含み、任意の単一の固有の物理的位置でのその複数のヌクレオチド分子は同一の配列を有し、基板上の固有の物理的位置の各々は、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は異なる配列を有するヌクレオチド分子を含むことを特徴とするアレイ 20

【請求項56】

SEQ ID NO:1のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項57】

SEQ ID NO:2のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項58】

SEQ ID NO:3のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項59】

SEQ ID NO:4のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項60】 30

SEQ ID NO:5のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項61】

SEQ ID NO:6のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項62】

SEQ ID NO:7のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項63】

SEQ ID NO:8のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項64】

SEQ ID NO:9のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項65】 40

SEQ ID NO:10のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項66】

SEQ ID NO:11のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項67】

SEQ ID NO:12のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項68】

SEQ ID NO:13のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項69】

SEQ ID NO:14のアミノ酸配列を有する請求項1に記載のポリペプチド。

【請求項70】 50

SEQ ID NO:15のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 7 1】

SEQ ID NO:16のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 7 2】

SEQ ID NO:17のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 7 3】

SEQ ID NO:18のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 7 4】

SEQ ID NO:19のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 7 5】

SEQ ID NO:20のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 7 6】

SEQ ID NO:21のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 7 7】

SEQ ID NO:22のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 7 8】

SEQ ID NO:23のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 7 9】

SEQ ID NO:24のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 8 0】

SEQ ID NO:25のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 8 1】

SEQ ID NO:26のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 8 2】

SEQ ID NO:27のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 8 3】

SEQ ID NO:28のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 8 4】

SEQ ID NO:29のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 8 5】

SEQ ID NO:30のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 8 6】

SEQ ID NO:31のポリヌクレオチド配列を有する請求項 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有する請求項 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有する請求項 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有する請求項 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有する請求項 12 に記載のポリヌクレオチド。

【請求項 9 1】

SEQ ID NO:36のポリヌクレオチド配列を有する請求項 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有する請求項 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有する請求項 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有する請求項 12 に記載のポリヌクレオチド。

【請求項 9 5】

10

20

30

40

50

SEQ ID NO:40のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項96】

SEQ ID NO:41のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項97】

SEQ ID NO:42のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項98】

SEQ ID NO:43のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項99】

SEQ ID NO:44のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項100】

SEQ ID NO:45のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項101】

SEQ ID NO:46のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項102】

SEQ ID NO:47のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項103】

SEQ ID NO:48のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項104】

SEQ ID NO:49のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【請求項105】

SEQ ID NO:50のポリヌクレオチド配列を有する請求項12に記載のポリヌクレオチド。

【発明の詳細な説明】

【技術分野】

【0001】

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、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】

【0002】

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、輸送または分泌されるタンパク質のアミノ末端に存在するシグナルペプチドによって媒介される。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と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現象によって活性化される。そのような現象の例として、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現象の例として、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とフォールディング、および受容体複合体若しくは細孔複合体を有する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が含まれる。

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)、細胞表面マーカー、および抗原認識分子がある(Alberts, B.他(1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, 557

10

20

30

40

50

-560,582, 582-592ページの概説を参照)。

【0003】

細胞表面マーカーの例としては、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)への共有結合性の接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ページの概説を参照)。

10

【0004】

基質タンパク質(MP)は、膜貫通タンパク質と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变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(概説に関しては、Ayad, S. 他(1994) The Extracellular Matrix Facts Book, Academic Press, San Diego, CA, 2-16ページ; Ruoslahti, E. (1997) *Kidney Int.* 51: 1413-1417; Sjaastad, M. D. および Nelson, W. J. (1997) *BioEssays* 19:47-55を参照)。

20

30

【0005】

ムチンは、非常にグリコシル化された糖タンパクで、粘液ゲルの主要な構成成分である。ムチンの生理学的機能は、細胞保護、機械的保護、分泌液の粘性維持、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)。ヘモムチンは新規なショウジョウバエ表面ムチンであり、抗菌エフェクター分子の誘導に関与し得る(Theopold, U. 他(1996) *J. Biol. Chem.* 271:12708-12715)。

40

【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知られている(Deutsch D. 他(1998) *Connect Tissue Res.* 39:177-184)。

50

【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、細胞外基質の分泌性糖タンパク質である。これらは、多様な組織に、また、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構成する。この5つのうちの1つであるTIGR/ミオシリン (myocilin) タンパク質は眼において発現され、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、ラット神経細胞のオルファクトメジンに関連し、ERに局在するタンパク質と96%の配列同一性を有するAMYと呼ばれる135アミノ酸のタンパク質が神経芽細胞腫細胞系cDNAライブラリで発見され、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)。ラット脳のcDNAライブラリから単離された神経細胞特異的オルファクトメジン関連糖タンパク質は、オルファクトメジンとの強い配列類似性を示す。この類似性は、神経細胞と神経分泌細胞において、これらの糖タンパク質が基質に関連した機能があることを示唆している (Danielson, P.E. 他 (1994) J. Neurosci. Res. 38:468-478)。

10

【0008】

Mac-2 結合タンパク質は90KDa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。成熟タンパク質の構造については、その長さが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システインリッチドメインによって定義される古代のタンパク質スーパーファミリーによく似たドメインを示す (Koths, 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、初代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が複数の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。この免疫刺激の原因は、発癌性形質転換、ウイルス感染または病原性侵襲であると提起されている (Ullrich, A., 他 (1994) J. Biol. Chem. 269:18401-18407)。

20

【0009】

セマフォリン類は少なくとも30の異なるメンバーからなる軸索誘導分子の大きなグループであり、脊椎動物、非脊椎動物、更にはある種のウイルスにおいても見つかっている。すべ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であるニューロピリン (neuropilin) は、*in vitro*で神経突起の生成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CUB、ジスコイジン (discoidin)、およびMAMドメインの3種のドメインからなる。ニューロピリンのCUBモチーフおよびMAMモチーフはタンパク質間相互作用において幾つかの役割を果たすことが示唆されており、更にセマドメインおよびC末端ドメインを介してセマフォリン類の結合に関与していると考えられる (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され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)。

30

40

【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。PSGは、癌胎児性抗原と共に、免疫グロブリンスーパーファミリー内の1サブファミリーを構成する (Plouzek C.A. およびChou J.Y. (1991) Endocrinology 129:950958)。PSGの様々な亜集団がヒト胎盤の

50

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* 14:277285)。

【0011】

自己分泌型運動促進因子 (AMF) は腫瘍細胞の遊走を調節する運動性サイトカインの1つであり、したがって、それと相まってシグナル伝達経路の同定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発見された (Ohta Y. 他 (2000) *Int. J. Oncol.* 17:259264)。AMFRは、分子量78K Da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、アミノ酸の誘導体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ページなど参照)。

【0013】

プロオピオメラノコルチン (POMC) は、脳下垂体前葉が合成するホルモンであるコルチコトロピン (ACTH) の前駆体ポリペプチドであり、副腎皮質の刺激で機能する。POMCはまた、リポトロピンホルモン (-LPH) の前駆体ポリペプチドでもある。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 は -LPHのペプチド成分であるが、-MSHは -LPH中に含有される。ACTH欠乏による副腎不全は、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しており、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、内分泌障害を生じる (Chretien, M. 他 (1979) *Can. J. Biochem.* 57:1111-1121, Krude, H. 他 (1998) *Nat. Genet.* 19:155-157, *Online Mendelian Inheritance in Man (OMIM)* 176830)。

【0014】

成長因子類および分化因子類は、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)。成長因子類および分化因子類には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3のクラスには、ボンベシン、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、およびブラジキニンなど、小さなペプチド因子類を含み、これらは増殖以外の細胞機能を調節するホルモンとして機能する。

【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、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)。

10

【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 (1998) Bra in Res. Mol. Brain Res. 62:175-186) 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。コードされるタンパク質群は推定上の分泌タンパク質であり、EFG様モチーフおよびロイシンリッチリピートを持つが、これらは共に、保存されたタンパク質間相互作用ドメインである。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、上記)。タンパク質のSlit ファミリーは神経組織内のglypican-1の機能性リガンドであることが示されており、それらの相互作用は中枢神経系の組織発生時の幾つかの段階において重要であり得ると示唆される (Liang, Y. 他 (1999) J. Biol. Chem. 274:17885-17892)。

20

【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は直接にシグナルを伝達し得る他、別の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作用することができる。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)。

30

【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などの薬剤は、血圧降下に用いられる (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)。

40

【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

50

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)。

【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、マトリックスメタロプロテイナーゼ類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. および Stetler-Stevenson, W. G. (1994) Eur. Respir. J. 7:2062-2072; Mignatti, P. および 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するアセチルCoA合成酵素である (Luong, A. 他 (2000) J. Biol. Chem. 275:26458-26466)。アセチルCoA合成酵素の活性の結果として、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、高血圧との関連が示されている (T oh, H. (1991) Protein Seq. Data Anal. 4:111-117 および Iwai, N. 他 (1994) Hypertension 23:375-380)。

10

【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つのファミリーは、FK506結合タンパク質類 (FKBP) とシクロフィリン類 (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) は、細胞の様々な領域に局在し、それらの領域で様々なタンパク質複合体と会合する (Coss, M. 他 (1995) J. Biol. Chem. 270:29336-29341; Schreiber, S.L. (1991) Science 251:283-287)。

20

30

【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集合または分解、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)。

40

大部分の通常の真核細胞は特定の回数、分裂した後、老化状態に入り、この状態において細胞は生存でき、代謝的に活性であるが、もはや複製はしない。細胞サイズの増加やpH依存性ベータガラクトシダーゼ活性のような多くの形質的变化、および特定遺伝子群の上方制御のような分子的变化が老化細胞 (senescent cells) において生じる (Shelton (1999) C

50

urrent Biology 9:939-945)。老化細胞が分裂促進因子に曝露されると、多数の遺伝子が上方制御されるが、細胞は増殖しない。老化細胞が *in vivo* で年齢とともに蓄積され、生体の老化 (aging) に寄与していることが示されている。さらに、老化は腫瘍形成を抑制し、老化に必要な多くの遺伝子もまた、p53 や網膜芽細胞腫感受性遺伝子のような癌抑制遺伝子として機能する。大部分の腫瘍はその複製限界を越えた細胞を含んでいる。つまり、これらは不死化されている。多くの癌遺伝子は腫瘍形成の第1段階として細胞を不死化する。

【0023】

酸化ストレス、放射、活性化腫瘍性タンパク質および細胞周期阻害剤のような多様な挑戦によって老化表現型が誘発され、老化が多くの増殖的シグナルや、抗増殖的シグナルによって影響されることが示されている (Shelton 前出)。老化は各々の細胞分裂とともに発生するテロメアの進行的短縮化と相関している。細胞内のテロメラーゼの触媒的成分の発現によってテロメアの短縮化が防がれ、また線維芽細胞や上皮細胞のような細胞が不死化されるが、別種の細胞 (例えば CD8+ T細胞) は不死化されない (Migliaccio 他 (2000) J. Immunol. 165:4978-4984)。このようにして、テロメアの短縮化および、細胞種によってはその他の機構によって老化は制御される。

10

【0024】

加齢における老化と腫瘍形成の役割を理解するための進行中の研究の一部として、老化細胞と前老化細胞の間で異なって発現される多数の遺伝子が同定されている。大部分の老化細胞は細胞周期のG1期において増殖が静止されている。多くの細胞周期遺伝子の発現は老化細胞および前老化細胞に類似している (Cristofalo (1992) Ann. N. Y. Acad. Sci. 663:187-194) が、増殖を抑制するサイクリン依存性キナーゼ p21 と p16、およびサイクリン D1 と E のような他の遺伝子の発現は老化細胞において上昇している。例えば細胞外基質タンパク質 (フィブロネクチン、プロコラーゲン、およびオステオネクチン)、およびプロテアーゼ (コラーゲナーゼ、ストロメライシンおよびカテプシン B) のような、細胞周期に直接関与しない他の遺伝子群もまた上方制御される (Chen (2000) Ann. NY Acad. Sci. 908:11125)。老化細胞において発現不足の遺伝子には、熱ショックタンパク質である c-fos と cdc-2 をコードする遺伝子が含まれる (Chen 前出)。

20

【0025】

プロリンリッチ -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参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. USA 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、血液凝固に参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 (bone-Gla protein)、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. Przysiecki (1987) Int. J. Biochem. 19:1-7; C. Vermeer (1990) Biochem. J. 266:625-636)。

30

40

【0026】

免疫グロブリン

抗原認識分子群は、すべての脊椎動物が、ウイルス、細菌、真菌および寄生虫による感染を防ぐために発達させてきた高度且つ複雑な免疫系において重要な役割を果たす。免疫系の主な特徴は、「自己」分子と外来分子すなわち抗原とを区別する能力である。この能力は、リンパ球、顆粒球および単球のような白血球によって発現される分泌タンパク質および膜貫通タンパク質によって主に媒介される。これらのタンパク質の多くは、保存された構造ドメインの 1 つ以上のリピートを持つメンバーたる免疫グロブリン (Ig) スーパーフ

50

ファミリーに属する。このIgドメインは、Igフォールド（折りたたみ）と呼ばれる或る配列でのジスルフィド結合によって接合されている、逆平行のシート群から構成されている。タンパク質がIgスーパーファミリーのメンバーであるための基準は、1つ以上のIgドメインを持ち、そのドメインが70～110アミノ酸残基の長さの領域であり、Ig可変領域様（V）ドメインまたはIg定常領域様（C）ドメインと相同であることである。Igスーパーファミリーのメンバーには抗体（Ab）、T細胞受容体（TCR）、クラスIおよびIIの主要組織適合性（MHC）タンパク質が、さらに、「分化クラスター」抗原すなわちCD抗原のCD2、CD3、CD4、CD8、ポリIg受容体、Fc受容体、神経細胞接着分子（NCAM）および血小板由来成長因子受容体（PDGFR）など、免疫細胞特異的表面マーカー群が含まれる。

【0027】

Igドメイン（VおよびC）は、保存されたアミノ酸残基の領域であり、これによってポリペプチドが免疫グロブリン（または抗体）フォールド（折畳み）と呼ばれる球状三次構造を形成する。免疫グロブリン（または抗体）フォールドはほぼ平行になった2層のシートからなる。55～75アミノ酸残基の長さをもつ、鎖内ジスルフィド結合ループを保存的なシステイン残基が形成し、2層のシートを接続する。各シートは、3または4本の逆平行ストランドを持ち、各ストランドは5～10アミノ酸残基の長さである。ストランド群内のアミノ酸残基の疎水性および親水性の相互作用がIgフォールドを安定させる（疎水性領域はストランドの内向きに面しているアミノ酸残基にあり、親水性領域は外向きに面している部分のアミノ酸残基にある）。VドメインはCドメインより長いポリペプチドで構成され、Igフォールド内に、付加的な1対のストランドを持つ。

10

20

【0028】

Igスーパーファミリー遺伝子の一貫した特徴は、あるIgドメインの各配列が1つのエキソンによってコードされることである。Igスーパーファミリーは、細胞間相互作用の仲介に参与する1つのIgドメインをコードする遺伝子から進化した可能性がある。そしてスーパーファミリーの新たなメンバーは、エキソンおよび遺伝子の複製によって生じた。現代のIgスーパーファミリータンパク質は、それぞれVおよび/またはCドメインの数が異なる。このスーパーファミリーの進化上の別の特徴は、DNAの再編成を起こす能力である。これは独自の機能で、ファミリーの抗原受容体メンバーが保持している。

【0029】

Igスーパーファミリーのメンバーの多くは膜内在性の形質膜タンパク質であり、細胞外Igドメインを持つ。それらの膜貫通ドメインおよびそれらの細胞質内尾部の疎水性アミノ酸残基は非常に多様で、Igファミリーのメンバー内または既知のシグナル伝達をする構造との相同性は少ないかまたは欠いている。スーパーファミリーに関するこの一般的な記述には例外がいくつかある。例えば、PDGFRの細胞質内尾部は、チロシンキナーゼ活性を持つ。また、Thy-1は、胸腺細胞およびT細胞で見られる糖タンパク質である。このタンパク質は細胞質内尾部を持たないが、その代わりに原形質膜に共有結合によるグリコシルホスファチジルイノシトール結合で固着している。

30

【0030】

Igスーパーファミリータンパク質の多くはまた、これらの分子の機能に必須の、Igドメイン間の相互作用という共通の特徴を持っている。多量体タンパク質のIgドメイン間の相互作用には、同種親和性のものと異種親和性（すなわち、同じIgドメイン間の作用、または異なるIgドメイン間の作用）のものがある。抗体は多量体タンパク質であり、Igドメイン間の同種親和性と異種親和性の双方の相互作用を持つ。重鎖の定常領域の対形成が抗体のFc領域を形成し、軽鎖および重鎖の可変領域の対形成が抗体の抗原結合部位を形成する。異種親和性の相互作用はまた、異なる分子のIgドメイン間でも起こる。これらの相互作用は免疫系における、または発生中および成熟した神経系における重要な細胞間相互作用のための細胞間の接着を提供する。（Abbas, A.K. 他（1991）Cellular and Molecular Immunology, W.B. Saunders Company, Philadelphia, PA, 142-145ページの概説を参照）。

40

【0031】

50

抗体

MHCタンパク質群は、外来抗原に結合し、これらをT細胞に提示する細胞表面マーカーである。MHC分子はクラスIまたはIIのどちらかに分類される。クラスIのMHC分子(MHC I)はほとんどすべての細胞表面に発現され、またキラーT細胞に対する抗原の提示に関与する。例えば、ウイルスに感染した細胞は細胞内ウイルスタンパク質を分解し、細胞表面のMHC I分子に結合したタンパク質断片を表現する。MHC I/抗原複合体は、感染した細胞とその中のウイルスを破壊するキラーT細胞によって認識される。クラスII MHC分子はB細胞およびマクロファージのような免疫系の分化した抗原提示細胞で主に発現する。これらの細胞は細胞外液から外来タンパク質を取り込み、細胞表面にMHC II/抗原複合体を発現する。この複合体はヘルパーT細胞を活性化し、次にヘルパーT細胞が、免疫反応を刺激するサイトカインおよび他の因子を分泌する。MHC分子はまた、臓器移植後の臓器拒絶反応において重要な役割を果たす。拒絶反応は、移植受容者のT細胞群が、移植された臓器の外来MHC分子群に対し、外来抗原に結合した自己MHC分子群に対するのと同様の応答をする場合に発生する(Alberts, B. 他 (1994) Molecular Biology of the Cell, Garland Publishing, New York, NY, 1229-1246ページの概説を参照)。

10

【0032】

抗体はIgスーパーファミリーの多量体メンバーであり、B細胞の表面で発現されるか、または、B細胞によって分泌されて血液循環に入る。抗体は、血液中や他の細胞外液中で外来抗原に結合し、それらを中和する。プロトタイプの抗体は、ジスルフィド結合によって連結された2つの同じポリペプチド重鎖(H鎖)と2つの同じポリペプチド軽鎖(L鎖)からなる四量体である。この配列は、抗体分子に対して特徴的なY型を形成する。抗体はH鎖組成に基づいて分類される。抗体の5つのクラスであるIgA, IgD, IgE, IgG およびIgMは、 α 、 δ 、 ϵ 、 γ および μ のH鎖型によって定義される。L鎖には κ と λ の2つのタイプがあり、どちらもH鎖対を持つペアとして会合している。血液循環内に見られる抗体の最も一般的なクラスであるIgGは四量体であるが、抗体の他のクラスのものは一般にこの基本的構造の変異体か、または多量体である。

20

【0033】

H鎖とL鎖は各々、N末端可変領域とC末端定常領域を有する。定常領域はL鎖の約110のアミノ酸と、H鎖の約330または440のアミノ酸から構成される。定常領域のアミノ酸配列は、或る特定クラスのH鎖またはL鎖群の内では、ほぼ同一である。可変領域は約110のアミノ酸からなり、H鎖とL鎖の両方にある。しかし、可変領域のアミノ酸配列は、特定クラスのH鎖またはL鎖群の中でも異なる。H鎖またはL鎖の可変領域のそれぞれに、広範な配列多様性を持つ3つの高頻度可変領域があり、各々約5~10のアミノ酸からなる。抗体分子において、H鎖およびL鎖の高頻度可変領域は1つになり、抗原認識部位を形成する(前出のAlberts, B. 他1206-1213ページと1216-1217ページの概説を参照)。

30

【0034】

H鎖とL鎖は共に、Igスーパーファミリーのメンバーの、反復したIgドメインを含む。例えば、典型的なH鎖は4つのIgドメインを含んでおり、そのうちの3つは定常領域内に発生し、1つは可変領域内に発生して抗原認識部位の形成に寄与している。同様にして、典型的なL鎖は2つのIgドメインを含み、そのうちの1つは定常領域内に発生し、他の1つは可変領域内に発生する。

40

【0035】

免疫系は体内に入る外来分子を認識し、それに応答する能力を持っている。したがって、免疫系はすべての可能性のある抗原に対する抗体の完全な蓄積で武装していなければならない。このような抗体の多様性は、可変領域と定常領域とをコードする遺伝子セグメント群の体細胞性再配列によって作られる。これらの遺伝子セグメント群は、各々の遺伝子セグメントが隣接する高度に保存されたDNA配列間で生じる、部位特異的組換えによって連結されている。何百という異なった遺伝子セグメントがあるため、何百万という独自の遺伝子が、組み合わせにより作成され得る。その上、これらのセグメントの不正確な連結と

50

これらのセグメント内での異常に高頻度の体細胞性突然変異が、多様な抗体集団の産生に更に寄与している。

【0036】

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

【0037】

乳癌

180,000例を超える新規発症の乳癌が毎年診断されており、乳癌による死亡率は45~54歳の女性の全死亡の10%に近い(K. Gish (1999) AWIS Magazine 28:7-10)。ただし、限局性乳癌の早期診断に基づく生存率はきわめて高い(97%)。これに対して、疾患の段階が進行し腫瘍が乳房外に進展した場合は22%である。現在の臨床乳癌検診の手法は感度と特異性とを欠いており、乳癌の包括的な遺伝子発現プロファイルを開発する努力がなされている。このプロファイルを従来のスクリーニング法とともに用いれば、乳癌の診断と予後診断とを向上し得る(Perou, C.M. 他(2000) Nature 406:747-752)。

20

【0038】

乳癌は細胞疾患における突然変異によって通常生じる遺伝疾患である。

2種の遺伝子、BRCA1とBRCA2との突然変異が、女性の乳癌の大きな素因であることが知られており、これらの変異は親から子へ受け渡されるようである(前出Gish)。しかし、このタイプの遺伝性乳癌はわずかに乳癌の約5%~9%であり、大多数の乳癌の原因は乳房上皮細胞で生じる非遺伝性突然変異である。

【0039】

乳癌に関連する特定の遺伝子の発現についてはかなりのことがすでにわかっている。例えば、上皮細胞成長因子(EGF)とその受容体(EGFR)と、ヒト乳癌との関係は特によく研究されている。(この領域の概説については前出のKhazaie 他およびそこに引用されている参考文献を参照されたい。)EGFRの過剰発現、特にエストロゲン受容体の下方調節と結びついた発現は、乳癌患者の不良な予後のマーカーである。また、乳房腫瘍転移におけるEGFR発現は、しばしば、原発性腫瘍に比べて上昇することから、EGFRは腫瘍の進行と転移とに関わることが示唆される。これを支持する証拠が蓄積されている。それは、EGFが、転移の潜在能に関する細胞機能に対する効果を持つという証拠である。この機能とは例えば、細胞運動性、走化性、分泌、および分化である。erbB受容体ファミリー(EGFRはその1つである)の他のメンバーの発現の変化も、乳癌に関わると考えられている。erbB受容体(例えばHER-2/neu, HER-3およびHER-4)およびそれらのリガンドが乳癌において豊富に存在するということは乳癌の病原におけるそれらの機能的な重要性を示唆しており、したがって乳癌治療の標的となり得る(Bacus, S. S. 他(1994) Am J Clin Pathol 102:S13-S24)。別の既知の乳癌マーカーとしては、或るヒトの分泌型frizzledタンパク質のmRNA(乳房腫瘍では下方調節される)、マトリクスG1aタンパク質(ヒト乳癌細胞で過剰発現される)、Drg1すなわちRTP(この遺伝子の発現は結腸癌、乳癌、前立腺腫瘍で低下する)、maspin(この腫瘍抑制遺伝子は浸潤性乳癌で下方調節)、およびCaN19(S100蛋白ファミリーのメンバーであり、このファミリーは全て、乳癌細胞では正常乳腺上皮細胞に比して下方調節される)を含む(Zhou, Z. 他(1998) Int J Cancer 78:95-99; Chen, L 他(1990) Oncogene 5:1391-1395; Ullrich W 他(1999) FEBS Lett 455:23-26; Sager, R 他(1996) Curr Top Microbiol Immunol 213:51-64; 及び Lee, SW 他(1992) Proc Natl Acad Sci USA 89:2504-2508)。

30

40

50

【 0 0 4 0 】

種々の病期のヒト乳癌の乳腺上皮細胞に由来する細胞株は、悪性転換と腫瘍進行とのプロセスの研究に有用なモデルを提供する。その理由は、これらの細胞株が、その親腫瘍の多くの特性を、長い培養期間にわたり保持することが示されているからである(Wistuba, I. I. 他 (1998) Clin Cancer Res 4:2931-2938).このようなモデルが特に有用なのは、ヒト乳腺上皮細胞の表現型と分子との特徴を、種々の段階の悪性転換について比較する場合である。

【 0 0 4 1 】

大腸癌

結腸直腸癌は米国の癌死亡で二番目に高い。結腸癌は、全症例の90%が55才以上に発症しているため老化に関連している。広く受け入れられている仮説は、いくつかの起因する遺伝子突然変異が長年にわたる罹患により蓄積されることである。結腸直腸癌における遺伝子的変化の本質を理解するために、遺伝性疾患を重視した多くの研究がなされている。最初に知られた遺伝性疾患である、家族性大腸ポリポーシス(FAP)は腺腫性ポリープ症Coli遺伝子(APC)における突然変異によって生じ、そのタンパク質の切断形態または不活性形態を生じる。この癌抑制遺伝子は染色体5qにマップされている。二番目によく知られている遺伝性疾患は遺伝性非ポリープ症結腸直腸癌(HNPCC)で、これはミスマッチの修復遺伝子における突然変異によって生じる。

【 0 0 4 2 】

遺伝性大腸癌の発症率は低く、大部分の結腸直腸癌は散発的であるが、遺伝性疾患の研究から理解されることが一般に適用され得る。例えば、APCにおける体細胞の突然変異は無差別の結腸腫瘍の少なくとも80%に発生する。APC突然変異は疾患における開始現象であると考えられている。引き続いて、他の突然変異が生じる。結腸直腸癌の約50%に、ラス(ras)癌遺伝子の活性化突然変異があり、また、85%にp53の不活性化突然変異がある。これらの遺伝子における変化によって、大腸癌における遺伝子発現の変化が生じる。これらの突然変異の下流の標的と、これらが癌の発症と進行において果たす役割についてはあまり理解されていない。

【 0 0 4 3 】

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群の発見により、新規の組成物群を提供することで、当分野の要望に応えることができる。これら新規の組成物は、細胞増殖異常、自己免疫/炎症疾患、心血管疾患、神経系疾患、および発達障害の、診断・治療・予防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群の効果についての算定にも有用である。

【 発明の開示 】

【 発明の効果 】

【 0 0 4 4 】

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-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」、および「SECP-24」および「SECP-25」呼ぶ分泌タンパク質である精製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:1-25を有する群から選択したアミノ酸配列からなるポリペプチド、(b) SEQ ID NO:1-25を有する群から選択したアミノ酸配列と少なくとも90%が同一性のある天然のアミノ酸配列を有するポリペプチド、(c) SEQ ID NO:1-25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、および(d) SEQ ID NO:1-25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免疫原性断片を含む群から選択した単離されたポリペプチドを提供する。一実施態様では、SEQ ID NO:1-25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。

【 0 0 4 5 】

また、本発明は (a) SEQ ID NO:1-25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:1-25からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列からなるポリペプチド、 (c) SEQ ID NO:1-25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:1-25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より選択されたポリペプチドをコードするような単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドはSEQ ID NO:1-25からなる群から選択したポリペプチドをコードする。別の実施態様では、ポリヌクレオチドはSEQ ID NO:26-50を有する群から選択される。

【0046】

本発明は更に、 (a) SEQ ID NO:1-25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、 (b) SEQ ID NO:1-25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持つポリペプチド、 (c) SEQ ID NO:1-25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:1-25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持つ組換えポリヌクレオチドを提供する。一実施態様で本発明は、この組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様で本発明は、この組換えポリヌクレオチドを持つ遺伝形質転換体を提供する。

【0047】

更に本発明は、 (a) SEQ ID NO:1-25からなる群から選択されたアミノ酸配列からなるポリペプチドと、 (b) SEQ ID NO:1-25からなる群から選択されたアミノ酸配列と90%以上の同一性を有する天然のアミノ酸配列からなるポリペプチドと、 (c) SEQ ID NO:1-25からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、 (d) SEQ ID NO:1-25とからなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択された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。

【0048】

本発明は更に、 (a) SEQ ID NO:1-25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:1-25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:1-25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:1-25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。

【0049】

本発明は更に、 (a) SEQ ID NO:26-50からなる群から選択したポリヌクレオチド配列を有するポリヌクレオチド、 (b) SEQ ID NO:26-50からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、 (c) (a) に相補的なポリヌクレオチド配列、 (d) (b) に相補的なポリヌクレオチド配列、および (e) (a) ~ (d) のRNA等価物からなる群から選択された単離されたポリヌクレオチドを提供する。一実施態様で該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群からなる。

【0050】

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:26-50からなる群から選択したポリヌクレオ

10

20

30

40

50

チド配列を有するポリヌクレオチド、(b) SEQ ID NO:26-50からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を有するポリヌクレオチド、(c)(a)に相補的なポリヌクレオチド、(d)(b)に相補的なポリヌクレオチド、および(e)(a)~(d)のRNA等価物からなる群から選択されたポリヌクレオチドの配列を有する。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群からなる。

10

【0051】

本発明はまた、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは、(a) SEQ ID NO:26-50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:26-50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%が同一の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) 20

20

【0052】

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:1-25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:1-25からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO:1-25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の生物学的活性断片、および(d) SEQ ID NO:1-25からなる群から選択したアミノ酸配列の免疫原性断片からなる群から選択される。一実施例では、その組成物は、SEQ ID NO:1-25からなる一群から選択されたアミノ酸配列を含む。更に、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。

30

【0053】

本発明はまた、(a) SEQ ID NO:1-25からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO:1-25からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を有するポリペプチド、(c) SEQ ID NO:1-25を有する群から選択したアミノ酸配列の生物学的活性断片、および(d) SEQ ID NO: 40

40

【0054】

本発明は更に、(a) SEQ ID NO:1-25からなる群から選択した或るアミノ酸配列を持つポ

50

リペプチド、(b) SEQ ID NO:1-25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持つポリペプチド、(c) SEQ ID NO:1-25からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO:1-25からなる群から選択した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。一実施態様で本発明は、この方法によって同定したアンタゴニスト化合物と薬剤として許容できる賦形剤とを含む組成物を提供する。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

【0055】

本発明は更に、(a) SEQ ID NO:1-25からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO:1-25からなる群から選択したアミノ酸配列と少なくとも90%が同一の天然のアミノ酸配列を有するポリペプチド、(c) SEQ ID NO:1-25を有する群から選択したアミノ酸配列の生物学的活性断片、または(d) SEQ ID NO:1-25からなる群から選択したアミノ酸配列の免疫原性断片を含む群から選択されたポリペプチドに特異結合する化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドを適切な条件下で少なくとも1つの試験化合物と混合させる過程と、(b) 試験化合物とのポリペプチドの結合を検出し、それによってポリペプチドに特異結合する化合物を同定する過程とを含む。

20

【0056】

本発明は更に、(a) SEQ ID NO:1-25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、(b) SEQ ID NO:1-25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持つポリペプチド、(c) SEQ ID NO:1-25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生物学的活性断片、および(d) SEQ ID NO:1-25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の活性をモジュレート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モジュレートする化合物を標示する。

30

【0057】

更に本発明は、SEQ ID NO:26-50からなる群から選択した或るポリヌクレオチド配列を持つ標的ポリヌクレオチドの発現を改変する効果につき、或る化合物をスクリーニングする一方法を提供する。この方法は、(a) この標的ポリヌクレオチドを有する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改変を検出する過程と、(c) 可変量のこの化合物の存在下でのこの標的ポリヌクレオチドの発現と、この化合物の不在下での発現とを比較する過程とからなる。

40

【0058】

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:26-50を有す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:26-50を有する群から選択したポリヌクレオチド配列と少なくとも90%の相同性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(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なポリヌクレオチド、(v) (i) ~ (iv) のRNA等価物を含む群から選択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から構成されるプローブを用いて

50

、処理した生物学的サンプルの核酸をハイブリダイズする過程とを含む試験化合物の毒性の算定方法を提供する。ハイブリダイゼーションは、上記プローブと生体サンプル中の標的ポリヌクレオチドとの間に特異的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生じる。上記標的ポリヌクレオチドは、(i) SEQ ID NO:26-50からなる群から選択した或るポリヌクレオチド配列を持つポリヌクレオチド、(ii) SEQ ID NO:26-50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持つポリヌクレオチド、(iii) (i)のポリヌクレオチドに相補的なポリヌクレオチド、(iv) (ii)のポリヌクレオチドに相補的なポリヌクレオチド、(v) (i) ~ (iv)のRNA等価物、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持つ。10

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。

【発明を実施するための最良の形態】

【0059】

(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態様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

【0060】

請求の範囲および明細書中で用いている単数形の「或る」および「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を指す場合もあることに注意されたい。したがって、例えば「或る宿主細胞」と記されている場合にはそのような宿主細胞が複数あることもあり、「或る抗体」と記されている場合には単数または複数の抗体、および、当業者に公知の抗体の等価物などについても言及している。

本明細書中で用いる全ての技術用語および科学用語は、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、材料およ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

【0061】

(定義)

用語「SECP」は、天然、合成、半合成或いは組換え体などからの、全ての種(特にウシ、ヒツジ、ブタ、ネズミ、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

【0062】

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。

【0063】

用語「対立遺伝子変異体/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、変容したmRNAまたはポリペプチドを作製し得る。その構造または機能は、変容することもし50

ないこともある。遺伝子は、天然の対立遺伝子変異体を全く有しないか、1個若しくは数個の天然の対立遺伝子変異体を有し得る。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。

【0064】

SECPをコードする「変容した/改変された」核酸配列は、様々なヌクレオチドの欠失、挿入、或いは置換を有し、その結果、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。コードされるタンパク質も「変容する/改変される」ことがあり、サイレント変化を生じ機能的に等価なSECPとなる、アミノ酸残基の欠失、挿入、或いは置換を含み得る。意図的なアミノ酸置換は、生物学的或いは免疫学的に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。親水性値が近似している非荷電極性側鎖を有するアミノ酸には、アスパラギンとグルタミン、セリンとトレオニンがある。親水性値が近似している非荷電側鎖を有するアミノ酸には、ロイシン、イソロイシンおよびバリン、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

【0065】

「アミノ酸」または「アミノ酸配列」の語は、オリゴペプチド、ペプチド、ポリペプチド若しくはタンパク質の配列またはその断片を指し、天然または合成分子を指す。ここで「アミノ酸配列」は天然のタンパク質分子のアミノ酸配列を指すものであり、「アミノ酸配列」及び類似の語は、アミノ酸配列を、列挙したタンパク質分子に関連する完全な本来のアミノ酸配列に限定しようとするものではない。

【0066】

「増幅」は、核酸配列の追加複製に関連する。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。

【0067】

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。

30

【0068】

「抗体」の語は、そのままの免疫グロブリン分子やその断片、例えばFab₁、F(ab')₂及びFv断片を指し、これらは抗原決定基と結合することができる。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、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、所望により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。結合その結合ペプチドは、動物を免疫化するために用いる。

40

【0069】

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

50

【0070】

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 SEL EX (Systematic Evolution of Ligands by EXponential Enrichment の略、試験管内選択法)）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、ヌクレオチド誘導体、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構成要素は修飾された糖基（例えば、リボヌクレオチドの2'-OH基が2'-Fまたは2'-NH₂で置換されている）を有することが可能で、これらの糖基は、例えば、ヌクレアーゼに対する耐性あるいは血中でのより長い寿命など、望む性質を改善しうる。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ー等の分子に抱合させることができる。アプタマー類は、例えば或る架橋剤の光活性化によって、それらの同種リガンド群と特異的に架橋させ得る（Brody, E.N. および L. Gold (2000) J. Biotechnol. 74:5-13等を参照）。

10

【0071】

「intramer」の用語は *in vivo* で発現されるアプタマーを意味するたとえば、ワクシニアウイルスに基づくRNA発現系は、白血球の細胞質で特定のRNAアプタマーを高度に発現させるために使用されている（Blind, M. 他 (1999) Proc. Natl Acad. Sci. USA 96:3606-3610）。

20

【0072】

「spiegelmer」の語はL-DNA、L-RNAその他の左旋性ヌクレオチド誘導体またはヌクレオチド様分子を含むアプタマーを指す。左旋性ヌクレオチドを含むアプタマーは、右旋性のヌクレオチドを含む基質に通常作用する天然の酵素による分解に耐性がある。

【0073】

「アンチセンス」の語は、特定の核酸配列の「センス」（コーディング）鎖と塩基対を形成することが可能な任意の組成物を指す。アンチセンス組成物の例には、DNAや、RNAや、ペプチド核酸（PNA）や、修飾されたバックボーン連鎖たとえばホスホロチオ酸、メチルホスホン酸またはベンジルホスホン酸などを有するオリゴヌクレオチドや、修飾された糖基たとえば2'-メトキシエチル糖または2'-メトキシエトキシ糖などを有するオリゴヌクレオチドや、あるいは修飾された塩基たとえば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などを有するオリゴヌクレオチドがあり得る。

30

【0074】

アンチセンス分子は、化学合成または転写を含む任意の方法で製造することができる。相補的アンチセンス分子は、ひとたび細胞に導入されたら、細胞が形成した天然の核酸配列と塩基対を形成し、転写または翻訳を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、或る参考DNA分子のアンチセンス鎖を指すことがあり、「正」または「プラス」という表現は、或る参考DNA分子のセンス鎖を意味し得る。

【0075】

「生物学的に活性」の語は、天然分子の構造的機能、調節機能または生化学的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

【0076】

「相補(的)」または「相補性」の語は、塩基対形成によってアニーリングする2つの一本鎖核酸の間の関係を指す。例えば、配列「5'-AGT-3'」は、相補配列「3'-TCA-5'」と対合する。

【0077】

「～のポリヌクレオチド配列を含む(有する)組成物」及び「～のアミノ酸配列を有する組成物」は、所定のポリヌクレオチド配列またはアミノ酸配列を含む広範囲の任意の成分を指す。この組成物には、乾燥製剤または水溶液が含まれ得る。SECP若しくはSECPの断片

50

をコードするポリヌクレオチド配列を含む組成物は、ハイブリダイゼーションプローブとして使用され得る。このプローブは、凍結乾燥状態で保存可能であり、糖質などの安定化剤と結合させることが可能である。ハイブリダイゼーションにおいては、塩（例えばNaCl）、界面活性剤（例えばドデシル硫酸ナトリウム；SDS）及びその他の構成要素（例えばデンハート液、粉乳、サケの精子のDNA等）を含む水溶液中にプローブを分散させることができる。

【0078】

「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断片構築システム（GCG, Madison, WI）か、あるいはPhrap（University of Washington, Seattle WA）等の断片構築用のコンピュータプログラムを用いて1つ或いはそれ以上の重複するcDNAやEST、またはゲノムDNA断片から構築された核酸配列を指す。伸長及び構築の両方を行ってコンセンサス配列を作製する配列もある。

10

【0079】

「保存的なアミノ酸置換」は、置換がなされた時に元のタンパク質の特性を殆ど損なわないと予測されるような置換、即ちタンパク質の構造と特に機能が保存され、そのような置換による大きな変化がない置換を指す。下表は、タンパク質中で元のアミノ酸と置換可能で、保存アミノ酸置換と認められるアミノ酸を示している。

20

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

【0080】

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、例えばシートやヘリックス構造、（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）側鎖の大部分が保持される。

【0081】

「欠失」は、結果的に1個若しくは数個のアミノ酸残基またはヌクレオチドが失われてなくなるようなアミノ酸またはヌクレオチド配列における変化を指す。

【0082】

「誘導体」の語は、ポリペプチドまたはポリヌクレオチドの化学修飾を指す。例えば、ア

50

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。

【0083】

「検出可能な標識」は、測定可能なシグナルを生成することができ、ポリヌクレオチドまたはポリペプチドに共有結合または非共有結合するようなレポーター分子または酵素を指す。

10

【0084】

「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)、または欠損遺伝子またはタンパク発現の欠損を指す。このような比較は例えば、治療後サンプルと未治療のサンプルまたは病態のサンプルと正常サンプルの間で行われ得る。

【0085】

「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、安定したサブストラクチャー群の新たな組み合わせによって、新しいタンパク質が組立てられることが可能であり、新しいタンパク質機能の進化を促進できる。

20

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(parent sequence)と同一であるがその配列より長さが短いものを指す。断片は、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最初の250または500アミノ酸(またはポリペプチドの最初の25%または50%)から選択された或る長さの連続したアミノ酸を有し得る。これらの長さは明らかに例として挙げているものであり、本発明の実施例では、配列表、表及び図面を含む明細書に裏付けされた任意の長さであってよい。

30

【0086】

SEQ ID NO:26-50の断片は、例えば、この断片を得たゲノム内の他の配列とは異なる、SEQ ID NO:26-50を明確に同定する固有のポリヌクレオチド配列の領域を含む。SEQ ID NO:26-50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:26-50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:26-50の断片の正確な長さ及び、断片が対応するSEQ ID NO:26-50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】

SEQ ID NO:1-25のある断片は、SEQ ID NO:26-50のある断片によってコードされる。SEQ ID NO:1-25の断片には、SEQ ID NO:1-25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:1-25の断片は、SEQ ID NO:1-25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:1-25の断片、及び断片が対応するSEQ ID NO:1-25の領域の正確な長さは、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

40

【0088】

「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。

50

【0089】

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

【0090】

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少なくとも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するので、2つの配列をより有意に比較できる。

【0091】

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。このプログラムは、LASERGENE ソフトウェアパッケージ(一組の分子生物学的分析プログラム)(DNASTAR, Madison WI)の一部である。このCLUSTAL Vは、Higgins, D. G. 及び P.M. Sharp (1989) CABIOS 5:151-153、Higgins, D.G. 他 (1992) CABIOS 8:189-191に記載されている。ポリヌクレオチド配列をペアワイズ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と設定される。「weighted(重み付けされた)」残基重み付け表が、デフォルトとして選択される。CLUSTAL Vは、アラインメントされたポリヌクレオチド配列対間の「percent similarity(類似率)」として一致率を報告する。

【0092】

或いは、米国国立バイオテクノロジー情報センター(NCBI)のBasic Local Alignment Search Tool (BLAST)が一般的に用いられ、且つ、無料で入手可能な配列比較アルゴリズム一式を提供している(Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410)。このアルゴリズムは、幾つかの情報源から入手可能であり、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/>)からも入手可能である。BLASTソフトウェア一式には様々な配列分析プログラムが含まれており、既知のポリヌクレオチド配列を種々のデータベースから得た別のポリヌクレオチド配列とアラインメントする「blastn」もその1つである。その他にも、2つのヌクレオチド配列をペアワイズで直接比較するために用いる「BLAST 2 Sequences」と称されるツールも利用可能である。「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することが可能である。「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に設定された「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。

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

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

10

20

30

40

50

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。

【0093】

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

【0094】

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

【0095】

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で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「類似率」として一致率を報告する。

【0096】

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をする場合、ある者は、デフォルトパラメータで設定された「BLAST 2 Sequences」ツールVersion 2.0.12 (Apr-21-2000)でblastpを使用しうる。デフォルトパラメータの設定例を以下に示す。

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

一致率は、定義された（例えば特定の配列番号で定義された）ポリペプチド配列全体の長さと比較して測定し得る。或いは、より短い長さ、例えばより大きな定義された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）の長さと比較して一致率を測定してもよい。ここに挙げた長さは単なる例示的なものに過ぎず、表、図及び配列リストを含めた本明細書に記載された配列に裏付けられた任意の配列長さの断片を用いて、或る長さであってその長さに対して一致率を測定し得る長さを説明し得ることを理解されたい。

【0097】

「ヒト人工染色体（HAC）」は、約6 kb（キロベース）～10 MbのサイズのDNA配列を含み得る、染色体の複製、分離及び維持に必要な全ての要素を含む直鎖状の微小染色体である。

【0098】

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。

【0099】

「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

10

20

30

40

50

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に対する許容条件は、当業者が慣例的に決定できる。許容条件は、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することができる。アニーリングが許容される条件は、例えば、温度が68 10
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。

【0100】

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における特定の配列の融点(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 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 20
を参照されたい。

【0101】

本発明の、ポリヌクレオチドとポリヌクレオチドのハイブリダイゼーションに対する高ストリンジェンシー条件には、約0.2 × SSC及び約1%のSDS存在下で約68 において1時間の洗浄条件が含まれる。或いは、65、60、55 または42 の温度で行ってもよい。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の変性サケ精子DNAがある。特定条件下で、例えばRNAとDNAのハイブリダイゼーションに有機溶剤、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 30
、当業者には自明であろう。ハイブリダイゼーションは、特に高ストリンジェント条件下では、ヌクレオチド間の進化的な類似性を示唆し得る。このような類似性は、ヌクレオチド及びヌクレオチドにコードされるポリペプチドに対する類似の役割を強く示唆している。

【0102】

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(C₀tまたはR₀t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 40
いるような2つの核酸配列間に形成され得る。

【0103】

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

【0104】

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

【0105】

用語「免疫原性断片」は、例えば哺乳動物などの生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。

【0106】

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。

【0107】

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

10

【0108】

用語「モジュレート」または「活性の調節」は、SECPの活性の変化を指す。例えば、調節によって、SECPのタンパク質活性の増減、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。

【0109】

「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

【0110】

「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、2つのタンパク質コード領域を接続するために必要な場合は同一リーディングフレーム内にある。

【0111】

「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のバックボーンに結合した、少なくとも約5ヌクレオチドの長さのオリゴヌクレオチドを含む、アンチセンス分子または抗遺伝子剤を指す。末端のリジンは、この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。

30

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当分野で既知の、その他の修飾が含まれ得る。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ることとなる。

【0112】

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

40

【0113】

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、8

50

0、90、100または150の連続したヌクレオチドからなるようなプローブ及びプライマーを用いてもよい。これよりもかなり長いプローブ及びプライマーもある。表、図面及び配列リストを含む本明細書に裏付けされた任意の長さのヌクレオチドを用いることができるものと理解されたい。

【0114】

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, Plain view NY、Ausubel, F.M. 他, (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY、Innis他 (1990) PCR Protocols, A Guide to Methods and Applications Academic Press, San Diego CA等を参照されたい。PCRプライマー対は、その目的のためのコンピュータプログラム、例えばPrimer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, 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。

10

【0115】

プライマーとして用いるオリゴヌクレオチドの選択は、そのような目的のために本技術分野でよく知られているソフトウェアを用いて行う。例えば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、オリゴヌクレオチド及び最大5,000までの大きめのポリヌクレオチドであって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得たもの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「非プライミングライブラリ(mispriming library)」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(英国ケンブリッジ市の英国ヒトゲノムマッピングプロジェクト-リソースセンターから一般向けに入手可能)は、多数の配列アラインメントに基づいてプライマーを設計し、それによって、アラインメントされた核酸配列の最大保存領域または最小保存領域の何れかとハイブリダイズするようなプライマーの選択を可能にする。従って、このプログラムは、固有であって保存されたオリゴヌクレオチド及びポリヌクレオチドの断片の同定に有用である。上記選択方法のいずれかによって同定したオリゴヌクレオチド及びポリヌクレオチドの断片は、ハイブリダイゼーション技術において、例えば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、或いは核酸のサンプルにおいて完全または部分的相補的ポリヌクレオチドを同定する特異プローブとして有用である。オリゴヌクレオチドの選択方法は、上記の方法に限定されるものではない。

20

30

【0116】

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

40

【0117】

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

50

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。

【0118】

「調節因子」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。調節因子は、転写、翻訳またはRNA安定性を調節する宿主タンパク質またはウイルスタンパク質と相互作用する。

【0119】

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

10

【0120】

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

【0121】

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、細胞や細胞から単離された染色体や細胞内小器官(オルガネラ)または膜からの抽出物と、細胞と、溶液中に存在するまたは基板に固定されたゲノムDNA、RNA、cDNAと、組織と、組織プリント等を含み得る。

20

【0122】

用語「特異的結合」及び「特異的に結合する」は、タンパク質若しくはペプチドと、アゴニスト、抗体、アンタゴニスト、小分子、若しくは任意の天然若しくは合成の結合組成物との間の相互作用を指す。この相互作用は、タンパク質の特定の構造(例えば抗原決定基即ちエピトープ)であって結合分子が認識するものが存在するか否かに依存している。例えば、抗体がエピトープ「A」に対して特異的である場合、遊離した標識A及びその抗体を含む反応において、エピトープA(つまり遊離した、標識されていないA)を含むポリペプチドの存在が、抗体に結合する標識されたAの量を低減させる。

30

【0123】

「実質的に精製された」の語は、自然環境から取り除かれ、単離または分離された核酸またはアミノ酸配列であって、自然に会合する他の構成要素の少なくとも約60%、好ましくは少なくとも約75%、最も好ましいのは少なくとも約90%が無いものを指す。

【0124】

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

用語「基板」は、任意の好適な固体或いは半固体の支持物を指し、膜及びフィルター、チップ、スライド、ウエハ、ファイバー、磁気または非磁気ビーズ、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、基板表面にはポリヌクレオチドやポリペプチドが結合する。

40

【0125】

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

【0126】

「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

50

よって選択する。限定するものではないが形質転換方法には、ウイルス感染、電気穿孔法（エレクトロポレーション）、熱ショック、リポフェクション及び微粒子銃を用いる方法がある。「形質転換された細胞」には、導入されたDNAが自律的に複製するプラスミドとして或いは宿主染色体の一部として複製可能である安定的に形質転換された細胞が含まれる。さらに、限られた時間に一過的に導入DNA若しくは導入RNAを発現する細胞も含まれる。

【0127】

ここで用いる「遺伝形質転換体(transgenic organism)」とは任意の生物であり、限定するものではないが動植物を含み、生物の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り、計画的な遺伝子操作によって、例えば微量注射法によって或いは組換えウイルスの感染によって行う。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). 遺伝子操作の語は、古典的な交雑育種或いは試験管内受精を指すものではなく、組換えDNA分子の導入を指すものである。本発明に基づいて予期される遺伝形質転換体には、バクテリア、シアノバクテリア、真菌及び動植物がある。本発明の単離されたDNAは、本技術分野で知られている方法、例えば感染、形質移入、形質転換またはトランス接合(transconjugation)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

【0128】

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

【0129】

特定のポリペプチド配列の「変異体」は、ポリペプチド配列の1本の長さ全体で特定のポリペプチド配列に対して少なくとも40%の相同性を有するポリペプチド配列として定義される。定義付けには、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%、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。

【0130】

(発明)

本発明は、新規のヒト分泌タンパク質群(SECP)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く。また、これらの組成物を利用した、細胞増殖異常、自己

10

20

30

40

50

免疫 / 炎症性障害、心血管障害、神経障害、および発達障害の、診断、治療、および予防の開発に基づく。

【0131】

表1は、本発明の完全長ポリヌクレオチド配列及びポリペプチド配列の命名の概略である。各ポリヌクレオチド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およびポリヌクレオチド配列に相当する物理的な、完全長クローンのIncyte ID 番号を示す。完全長のクローンは列3に示すポリペプチド配列に少なくとも95%の配列同一性を有するポリペプチドをコードする。

10

【0132】

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

【0133】

表3は、本発明のポリペプチドの様々な構造的特徴を示す。列1および列2はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号 (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、各ポリペプチドのアミノ酸残基数を示す。列4および列5はそれぞれ、GCG配列分析ソフトウェアパッケージのMOTIFSプログラム (Genetics Computer Group, Madison WI) によって決定された、リン酸化およびグリコシル化の可能性のある部位を示す。列6は、シグネチャ配列、ドメイン、およびモチーフを含むアミノ酸残基を示す。列7は、タンパク質の構造 / 機能の分析のための分析方法を示し、該当箇所には更に、分析方法に利用した検索可能なデータベースを示す。

30

【0134】

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。例えば、SEQ ID NO:2は、ヒトのSrCyp (イムノフィリン/サイクロフィリン) タンパク質 (GenBank ID g1770526) と残基F145から残基D308まで50%の同一性を有することが、Basic Local Alignment Search Tool (BLAST) によって示された (表2参照)。BLAST確率スコアは $1.6e-40$ であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO:2はまた、サイクロフィリンタイプのペプチジル-プロリルシストランスイソメラーゼ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ードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された (表3参照)。MOTIFSからのデータおよび PROFILESCAN解析により、SEQ ID NO:2 がサイクロフィリンタイプのペプチジルプロリルシストランスイソメラーゼである、更なる確証が提供される。

40

【0135】

別の例において、SEQ ID NO:3 はニワトリの新規なコラーゲンタンパク質 (GenBank ID g211610) に残基P223から残基G333まで53%の同一性を有することがBasic Local Alignment

50

Search Tool (BLAST)によって決定された(表2参照)。BLAST確率スコアは $1.9e-20$ であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO:3はまた、コラーゲン三重ヘリックスリピートとEGF様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ードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された(表3参照)。BLIMPSとMOTIFS解析からのデータはSEQ ID NO:3が分泌タンパク質である、さらに確証的な証拠を提供する(EGF様ドメインは分泌されるとされているタンパク質に存在する)。

【0136】

また他の例として、SEQ ID NO:17はM11残基からY60残基まで、ネコの主要な抗原I (GenBank ID g163825)と64%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 $1.5e-11$ 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:17はまた、ウテログロビン(uteroglobin)ファミリーのシグネチャを含むことが、タンパク質ファミリーとドメインのPrositeデータベースで定義されている構造的モチーフおよび配列的モチーフを検索するProfileScanアルゴリズムを用いて決定された。(表3参照)。BLIMPSとMOTIFS解析のデータはSEQ ID NO:17が分泌タンパク質である、さらに実証的な証拠を提供する。「ネコの主要抗原」(fel DI)はウテログロビンに構造的に関連していることがわかっているタンパク質であり、共に成熟タンパク質として分泌される)(Morgenster n, J.P., 他(1991) Proc. Natl. Acad. Sci. U.S.A. 88:9690-9694)。

【0137】

別の例として、SEQ ID NO:20は、タンパク質(アイソフォーム1)(GenBank ID g9930918)を含むヒトの新規な免疫グロブリンドメインと98%の同一性(残基M1から残基N541まで)を有することがBasic Local Alignment Search Tool (BLAST)によって決定された。(表2参照)。BLAST確率スコアは $3.9e-294$ であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO:20もまた、免疫グロブリン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ードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。(表3参照)。BLIMPS解析のデータは、SEQ ID NO:20が分泌タンパク質であるさらに実証的な証拠を提供する(「免疫グロブリンドメイン」は基質タンパク質の特性を示す)。

【0138】

また他の例として、SEQ ID NO:23はマウスの脂肪細胞特異的タンパク質3(GenBank ID g15777917)とM1残基からL154残基まで76%の、G47残基からG188残基まで39%の同一性を有することがBasic Local Alignment Search Tool (BLAST)によって示された(表2参照)。BLAST確率スコアは $3.1e-79$ と $1.3e-23$ 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、タンパク質間相互作用およびタンパク質とリガンドの相互作用に関与している、2つの免疫グロブリン(Ig)ドメインを含むヒトタンパク質に相同性を有することがPROTEOMEデータベースを用いたBLAST解析によって決定された。SEQ ID NO:23はまた、免疫グロブリンドメイン(e-値: $3.6e3$)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ードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。

【0139】

SEQ ID NO:1、SEQ ID NO:4、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8、SEQ ID NO:9、SEQ ID NO:10、SEQ ID NO:11、SEQ ID NO:12、SEQ ID NO:13、SEQ ID NO:14、SEQ ID NO:15、SEQ ID NO:16、SEQ ID NO:18、SEQ ID NO:19、SEQ ID NO:21、SEQ ID NO:22、SEQ ID NO:24およびSEQ ID NO:25は同様にして解析し、注釈付けをした。SEQ ID NO:1-25の解析のためのアルゴリズム及びパラメータが表7に記述されている。

10

20

30

40

50

【0140】

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で構築した。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の構築に使われたcDNA配列および/またはゲノム配列のヌクレオチド開始位置(5')と停止位置(3')、およびSEQ ID NO:26-50を同定するか、またはSEQ ID NO:26-50と、関連するポリヌクレオチド配列とを区別する技術(例えば、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ヌクレオチド開始位置(5')と終了位置(3')を示す。

【0141】

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構築に寄与した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(即ち「ENST」命名を含む配列)を同定し得る。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集合を意味する場合がある。例えば、FL_XXXXXX_N₁_N₂_YYY YY_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「縫合された(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構築物を指す場合もある。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識別番号(「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。

【0142】

あるいは、接頭コードは手作業で編集されたか、ゲノムDNA配列から予測されたか、または配列解析方法の組み合わせから由来してい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同じ配列の分析方法の例を列記する(実施例4と5を参照)。

接頭コード	解析のタイプおよび/またはプログラムの例
GNN 、 GFG 、 ENST	例えば GENSCAN (Stanford University, CA, USA) または FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK)を用いた、ゲノム配列からのエキソン予測
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	ゲノムへのEST配列のマッピングからの完全長転写物とエキソンとの予測。エキソン群と生じる転写物を予測するために、ゲノム位置とEST構成とのデータが組み合わされる。

10

【0143】

20

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。

【0144】

表5は Incyte cDNA配列を用いて構築された完全長ポリヌクレオチド配列のための代表的なcDNAライブラリを示している。代表的なcDNAライブラリは、上記のポリヌクレオチド配列を構築及び確認するために用いられた 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。

【0145】

30

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。

【0146】

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:26-50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:26-50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配列も含まれるが、そこでは窒素塩基チミンの出現はウラシルに置換され、糖のバックボーンはデオキシリボースの代わりにリボースで構成される。

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:26-50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:26-50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。

40

【0147】

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

50

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。

例えば、SEQ ID NO:48の配列を有する或るポリヌクレオチドは、SEQ ID NO:29の配列を有するポリヌクレオチドのスプライス変異体であり、SEQ ID NO:49の配列を有する或るポリヌクレオチドは、SEQ ID NO:43の配列を有するポリヌクレオチドのスプライス変異体であり、またSEQ ID NO:50の配列を有する或るポリヌクレオチドは、SEQ ID NO:47の配列を有するポリヌクレオチドのスプライス変異体である。上記の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つを有するアミノ酸配列をコードすることができる。

【0148】

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、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。したがって本発明は、可能コドン選択に基づく組合せの選択によって産出し得るようなありとあらゆる可能性のあるポリヌクレオチド配列のパ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。

【0149】

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ェントな条件下で、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含む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。

【0150】

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。

【0151】

更に本発明には、種々のストリンジェンシー条件下で、請求項に記載されたポリヌクレオチド配列、特に、SEQ ID NO:26-50及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる(例えば、Wahl, G.M.及びS.L. Berger (1987) *Methods Enzymol.* 152:399-407, Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511等を参照)。アニーリングおよ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」に記載した。

【0152】

DNAシーケンシングの方法は当分野では公知であり、本発明のいずれの実施例もDNAシーケンシング方法を用いて実施可能である。DNAシーケンシング方法には酵素を用いることができ、例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cle

10

20

30

40

50

veland 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, W atertown 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)。

10

【0153】

SECPをコードする核酸配列は当分野で周知の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 Application, 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鑄型から未知の配列を増幅する方法である。鑄型は、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。(Lagerstrom, M. 他 (1991) PCR Methods Application 1:111-119等を参照)。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。また、未知の配列を検索するために用い得る別の方法については当分野で知られている(Parker, J.D. 他 (1991) Nucleic Acids Res. 19:3055-3060等を参照)。更に、PCR、ネステッドプライマー類、およびPROMOTERFINDERライブラリ類 (Clontech, Palo Alto CA) を用いて、ゲノムDNAをウォーキングし得る。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で鑄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。

20

30

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

40

【0154】

市販のキャピラリー電気泳動システムを用いて、シーケンシングまたはPCR産物のサイズを分析し、またはそのヌクレオチド配列を確認することができる。具体的には、キャピラリーシーケンシングは、電気泳動による分離のための流動性ポリマーと、4つの異なるヌクレオチドに特異的であるような、レーザで活性化されるフルオロフォア (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

50

ブ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。

【0155】

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。

【0156】

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再構築による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。

10

【0157】

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作製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配列群を同定する、選択またはスクリーニングの手順を経る。続いて、これら好適な変異配列をプールし、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。従って、「人工的な」育種及び急速な分子の進化によって遺伝的多様性が生み出される。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み換えて、スクリーニングし、その後所望の特性が最適化されるまでシャッフリングし得る。或いは、所与の遺伝子を同種または異種のいずれかから得た同一遺伝子ファミリーの相同遺伝子と組み換え、それによって天然に存在する複数の遺伝子の遺伝多様性を、定方向の、制御可能な方法で最大化させることができる。

20

30

【0158】

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; 及びHorn, T.他 (1980) Nucl. Acids Res. Symp. Ser.225-232を参照)。別法として、化学的方法を用いて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(Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, 55-60ページ、Roberge, J.Y.他 (1995) Science 269:202-204等を参照)。自動合成はABI 431Aペプチドシンセサイザ(Applied Biosystems)を用いて達成し得る。更に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。

40

【0159】

このペプチドは、分取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.およびF.Z. Regnier (1990) Methods Enzymol. 182:392-421などを参照)。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)。

【0160】

50

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。必要なエレメント群は、特長およ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.* 20:125-162を参照)。

10

【0161】

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章及び8章, および16-17章; およびAusubel, F.M. 他. (1995) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York NY, ch. 9章、13章および16章を参照)。

20

【0162】

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や、酵母菌発現ベクターで形質転換させた酵母菌などの微生物や、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、動物細胞系がある。(前出のSambrook、前出のAusubel、Van Heeke, G. 及び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) *EMBOJ.* 6:307-311、『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, 191-196ページ、Logan, J. 他T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659、Harrington, J.J. 他 (1997) *Nat. Genet.* 15:345-355等を参照)レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ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. 及び 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。

30

40

【0163】

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

50

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(例えば、Van Heeke, G. および S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509を参照)。抗体の産生などのために多量のSECPが必要な場合は、SECPの発現を高レベルで誘導するベクターが使用し得る。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含むベクターが使用できる。

【0164】

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、Ausubel, 1995,前出、Bitter, G.A. 他 (1987) Methods Enzymol.153:516-544、及びScorer, C. A. 他 (1994) Bio/Technology 12:181-184を参照)。

【0165】

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)。或いは、RUBISCOの小サブユニット等の植物プロモーターまたは熱ショックプロモーターを用いてもよい(例えば、Coruzzi, G. 他 (1984) EMBO J. 3: 1671-1680; Broglie, R. 他 (1984) Science 224: 838-843; および Winter, J. 他 (1991) Results Probl. Cell Differ. 17 : 85-105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)。

哺乳動物細胞においては、多数のウイルスベースの発現系を利用し得る。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. および T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659を参照)。更に、ラウス肉腫ウイルス(RSV)エンハンサー等の転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させ得る。SV40またはEBVをベースにしたベクターを用いてタンパク質を高レベルで発現させることもできる。

【0166】

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6 kb~10 Mb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)で送達されている(例えば、Harrington, J.J. 他 (1997) Nat Genet.15:345-355.を参照)。

【0167】

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

10

20

30

40

50

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。

【0168】

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*⁻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単純ヘルペスウイルスのアデニンホスホリボシルトランスフェラーゼ遺伝子がある(例えば、Wigler, M. 他 (1977) *Cell* 11:223-232; 及びLowy, I. 他 (1980) *Cell* 22:817-823を参照)。また、選択の基礎として代謝拮抗物質、抗生物質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* 77:3567-3570; Colbere-Garapin, F. 他 (1981) *J. Mol. Biol.* 150:1-14等を参照)。この他の選択可能な遺伝子、例えば、代謝のための細胞の必要条件を変える*trpB*及び*hisD*が、文献に記載されている(例えば、Hartman, S.C. 及びR.C. Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:8047-8051参照)。可視マーカー、例えばアントシアニン、緑色蛍光タンパク質(GFP; Clontech)、グルクロニダーゼ及びその基質グルクロニド、またはルシフェラーゼ及びその基質ルシフェリン等を用いてもよい。これらのマーカーを用いて、トランスフォーマントを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)。

10

20

【0169】

マーカー遺伝子発現の存在/不存在によって目的の遺伝子の存在が示唆されても、その遺伝子の存在及び発現の確認が必要な場合もある。例えば、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。

【0170】

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。

30

【0171】

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。これらのアッセイ及びこれ以外のアッセイは、当分野で公知である(Hampton, R. 他 (1990) *Serological Methods, a Laboratory Manual*. APS Press, St Paul, MN, Sect. IV、Coligan, J. E. 他 (1997) *Current Protocols in Immunology*, Greene Pub. Associates and Wiley-Interscience, New York NY、Pound, J.D. (1998) *Immunochemical Protocols*, Humana Press, Totowa NJ等を参照)。

40

【0172】

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

50

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。

10

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。

【0173】

更に、宿主細胞株の選択は、挿入した配列の発現を調節する能力または発現したタンパク質を所望の形に処理する能力によって行い得る。限定するものではないがこのようなポリペプチドの修飾には、アセチル化、カルボキシル化、グリコシル化、リン酸化、脂質化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、特定の細胞装置および特徴的な機序を持つ、種々の宿主細胞（例えばCHO、HeLa、MDCK、HEK293、WI38など）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およびプロセッシングを確実にするように選択し得る。

20

【0174】

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及び異種ペプチド部分も、市販されている親和性基質を用いて融合タンパク質の精製を促進し得る。限定されるものではないがこのような部分には、グルタチオンSトランスフェラーゼ (GST)、マルトース結合タンパク質 (MBP)、チオレドキシン (Trx)、カルモジュリン結合ペプチド (CBP)、6-His、FLAG、c-myc、赤血球凝集素 (HA) がある。GSTは固定化グルタチオン上で、MBPはマルトース上で、Trxはフェニルアルシンオキシド上で、CBPはカルモジュリン上で、そして6-Hisは金属キレート樹脂上で、同族の融合タンパク質の精製を可能にする。FLAG、c-myc及び赤血球凝集素 (HA) は、これらのエピトープ標識を特異的に認識する市販されているモノクローナル抗体及びポリクローナル抗体を用いて、融合タンパク質の免疫親和性精製を可能にする。また、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。

30

40

【0175】

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

50

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。

【0176】

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。

【0177】

ある実施態様では、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に擬似のパートナーまたは天然の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): 5章などを参照）。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、また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。

10

【0178】

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、フルオロフォア、放射性同位体、酵素抱合体またはその他の検出可能な標識によりその結合を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、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およ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産物の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。

20

【0179】

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。

30

40

【0180】

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）等のマーカー遺伝子で破壊された目的の遺伝子を含むベクターで形質転換する。このベクターは、相同組換えにより、宿主ゲノムの対応する領域に組み込まれる。別法では、Cre-loxP系を用いて相同組換えを行い、組織特異的または発生段階

50

特異的に目的遺伝子をノックアウトする (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、可能性のある治療薬や毒性薬剤で検査することができる。

【0181】

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、造血系統及び心筋細胞に分化する (Thomson, J.A. 他 (1998) Science 282:1145-1147)。

10

【0182】

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「ノックイン」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。別の実施態様において、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)。

20

【0183】

(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。また、SECPを発現する組織の数例は、健常および癌性乳房組織であり、表6にも示す。したがって、SECPは細胞増殖異常、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。

30

【0184】

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、および発達障害が含まれ、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

40

50

筋または心膜炎、骨関節炎、骨粗しょう症、脾炎、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 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、トゥーレット病、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、デュシェンヌノ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。

10

20

30

【0185】

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。

40

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。

【0186】

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。

【0187】

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

50

、このような疾患の例には、上記した細胞増殖異常、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。

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。

【0188】

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み合わせて投与することもできる。併用療法で用いる好適な治療薬は、当業者が従来 of 医薬原理に従って選択し得る。治療薬と組み合わせることにより、上記した種々の疾患の治療または予防に相乗効果をもたらし得る。この方法により、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。

10

【0189】

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)。

20

【0190】

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 ゲラン桿菌)及びコリネバクテリウム パルバム(*Corynebacterium parvum*)が特に好ましい。

30

【0191】

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また、天然のタンパク質のアミノ酸配列の一部と同一であることが望ましい。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。

40

【0192】

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及びEBV-ハイブリドーマ技術がある(Kohler, G. 他. (1975) Nature 256:495-497、Kozbor, D. 他 (1985) J. Immunol. Methods 81:31-42、Cote, R.J. 他 (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030、Cole, S.P. 他 (1984) Mol. Cell Biol. 62:109-120等を参照)。

【0193】

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

50

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. 81:68516855; 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ィオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリ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）。

【0194】

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。（Orlandi, R. 他 (1989) Proc. Natl. Acad. Sci. U SA 86: 3833-3837、Winter, G. 他 (1991) Nature 349:293-299等を参照）。

10

【0195】

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 246:1275-1281等を参照）。

20

【0196】

種々のイムノアッセイを用いてスクリーニングし、所望の特異性を有する抗体を同定することができる。確立された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる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）。

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Kaで表すが、このKa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Ka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Kaは、親和性の真の測定値を表す。Ka値が $10^9 \sim 10^{12}$ liter/mo l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。Ka値が $10^6 \sim 10^7$ liter/mo l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)。

30

40

【0197】

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試薬の品質及び適性を決定することができる。例えば、少なくとも $1 \sim 2$ mg/ml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）。

【0198】

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

50

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（例えばAgrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJを参照。）

治療に用いる場合、アンチセンス配列を好適な標的細胞に導入するのに好適な任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を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（Slater, J.E. 他 (1998) *J. Allergy Clin. Immunol.* 102(3):469-475 及び Scanlon, K.J. 他 (1995)9(13):1288-1296等を参照）。アンチセンス配列はまた、例えばレトロウイルスやアデノ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Aubel, Uckert, W. and 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及び当分野で公知のその他のシステムが含まれる（Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J.他 (1998) *J. Pharm. Sci.* 87(11):1308-1315、Morris, M.C. 他 (1997) *Nucleic Acids Res.* 25(14):2730-2736. 等を参照）。

【0199】

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染色体連鎖遺伝（Cavazzana-Calvo, M.他 (2000) *Science* 288:669-672）により特徴付けられる重症複合型免疫不全(SCID)-X1病の場合）、先天性アデノシンデアミナーゼ（ADA）欠損症に関連する重症複合型免疫不全症候群（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. and 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）、(iii) 細胞内の寄生生物（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）などヒトレトロウイルス、B型若しくはC型肝炎ウイルス（HBV、HCV）、Candida albicans及びParacoccidioides brasiliensis等の真菌寄生菌、並びにPlasmodium falciparum及びTrypanosoma cruzi等の原動物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。

【0200】

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。

【0201】

10

20

30

40

50

SECPの発現に影響を及ぼし得る発現ベクターには、限定するものではないが、PCDNA 3.1、EPITAG、PRCCMV2、PREP、PVAX、PCR2-TOPOTAベクター (Invitrogen, Carlsbad CA)、PCM V-SCRIPT、PCMV-TAG、PEGSH/PERV (Stratagene, La Jolla CA)、PTET-OFF、PTET-ON、PTRE2、PTRE2-LUC、PTK-HYG (Clontech, Palo Alto CA)が含まれる。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inkin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 268:1766-1769; 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 (iii) 正常な個体に由来するSECPをコードする内因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。

10

20

30

40

50

【0202】

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。

【0203】

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。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、VPCLは、標的細胞上の受容体に対する親和性を有するエンベロープ遺伝子またはVSVg等の汎親和性エンベロープタンパク質を発現する (Armentano, D. 他 (1987) J. Virol. 61:1647-1650、Bender, M.A. 他 (1987) J. Virol. 61:1639-1646、Adam, M.A. および A.D. Miller (1988) J. Virol. 62:3802-3806、Dull, T. 他 (1998) J. Virol. 72:8463-8471、Zufferey, R. 他 (1998) J. Virol. 72:9873-9880)。Riggに付与された米国特許第5,910,434号(「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」)において、レトロウイルスパッケージング細胞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の増殖、細胞集団(例えばCD4⁺T細胞)の形質導入、及び形質導入した細胞の患者への戻しは、遺伝子治療の分野では当業者に公知の方法であり、多数の文献に記載されている (Ranga, U. 他 (1997) J. Virol. 71:7020-7029、Bauer, G. 他 (1997) Blood 89:2259-2267、Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716、Ranga, U. 他 (1998) Proc. Natl. Acad. Sci. U.S.A. 95:1201-1206、Su, L. (1997) Blood 89:2283-2290)。

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つ以上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の作製及びパッケージングについては、当業者に公知である。複製欠

損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を臍臓の無損傷の臍島内に導入するためにも用途が広いことが証明された (Csete, M.E. 他 (1995) *Transplantation* 27:263-268)。使用できる可能性のあるアデノウイルスベクターは、Armentanoに付与された米国特許第5,707,618号 (「Adenovirus vectors for gene therapy」) に記載されており、引用することをもって本明細書の一部とする。アデノウイルスベクターについては、Antinozzi, P.A. 他 (1999) *Annu. Rev. Nutr.* 19:511-544 及び Verma, I.M. 及び N. Somia (1997) *Nature* 18:389:239-242も参照されたい。両文献は、引用することをもって本明細書の一部とする。

【0204】

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つ以上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の作製及びパッケージングは、当業者に公知である。複製適格性単純ヘルペスウイルス(HSV)I型系のベクターは、レポーター遺伝子を霊長類の眼に送達するために用いられてきた (Liu, X. 他 (1999) *Exp. Eye Res.* 169:385-395)。HSV-1ウイルスベクターの作製についても、DeLucaに付与された米国特許第5,804,413号 (「Herpes simplex virus swains for gene transfer」) に開示されており、該特許の引用をもって本明細書の一部とする。米国特許第5,804,413号には、ヒト遺伝子治療を含む目的のために好適なプロモーターの制御下において細胞に導入される少なくとも1つの外因性遺伝子を有するゲノムを含む組換えHSVd92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 及び Xu, H. 他 (1994) *Dev. Biol.* 163:152-161も参照されたい。両文献は、引用をもって本明細書の一部とする。クローン化ヘルペスウイルス配列の操作、巨大ヘルペスウイルスのゲノムの異なった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。

【0205】

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(Semliki Forest Virus, SFV)の生物学的研究が広範に行われており、遺伝子導入ベクターは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)。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、ウイルスゲノムに、キャプシドをコードする領域の代わりにSECPのコード配列を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。

【0206】

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)

10

20

30

40

50

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頁等を参照)。相補配列またはアンチセンス分子もまた、転写物がリボソームに結合するのを阻止することによって mRNA の翻訳を阻止するように設計することができる。

【0207】

酵素的 RNA 分子であるリボザイムは、RNA の特異的切断を触媒するために用いることもできる。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関与している。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。任意の RNA 標的内の特異的リボザイム切断部位は、GUA、GUU、GUC 配列を含めたリボザイム切断部位に対して標的分子をスキャンすることによって先ず同定される。一旦同定されると、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。

【0208】

本発明の相補リボ核酸分子及びリボザイムは、核酸分子合成のために当分野でよく知られている任意の方法を用いて作製し得る。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、RNA 分子は、SECP をコードする DNA 配列の *in vitro* 及び *in vivo* 転写によって産出し得る。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、細胞株、細胞または組織内に導入することができる。

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' 末端または 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リボチオエートまたは 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。それには、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の他、非従来型塩基、例えばイノシン、クエオシン (queosine)、ワイプトシン (wybutosine) 等を含める。

【0209】

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の改変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改変し得る。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。

【0210】

特異ポリヌクレオチドの発現を改変する際の有効性に対して、少なくとも1個以上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容させる場合と、既存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、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。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性を調べるためのオリゴヌクレオチド (デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾オリゴヌクレオチド) の組み合わせライブラリをスクリーニングすることに関与している (Bruce, T.W. 他 (1997) 米国特許第5,686,242号、Bruce, T.W. 他 (2000) 米国特許第6,022,691号)。

【0211】

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、リポソーム注入またはポリカチオンアミノポリマーによる送達は、当分野でよく知られている方法を用いて実行することができる (Goldman, C.K. 他 (1997) *Nat. Biotechnol.* 15:462-466. 等を参照)。

【0212】

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

【0213】

本発明のさらなる実施例は、薬物として許容できる或る賦形剤と共に製剤される或る活性成分を一般に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。

【0214】

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。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 (例えば従来の低分子量有機薬) 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

ブチド及びタンパク質)の場合には、当該分野において肺の肺泡領域を介しての肺送達が一番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。

【0215】

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。

【0216】

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)。

10

【0217】

任意の化合物に対して、細胞培養アッセイ、例えば新生物性細胞の細胞培養アッセイにおいて、或いは、動物モデル、例えばマウス、ウサギ、イヌまたはブタ等において、先ず治療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。

20

【0218】

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。治療有効性及び毒性は、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の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。

30

【0219】

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、患者の全身健康状態、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する。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。

【0220】

通常投与量は、投与の経路にもよるが約0.1~100,000 µgであり、合計で約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、ヌクレオチドの処方では、タンパク質またはそれらのインヒビター類とは異な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。

40

【0221】

(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

50

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、上記の治療の箇所で記載した方法と同じ方法で作成される。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。

【0222】

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。

10

【0223】

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。用いられることができるポリヌクレオチドには、オリゴヌクレオチド配列、相補的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、疾患と相関し得る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。

20

【0224】

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）から作られているか、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が決まるであろう。

30

【0225】

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:26-50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。

【0226】

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

【0227】

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、および発達障害が含まれ、細胞増殖異常の中には

50

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（MC TD）、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 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。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

10

20

30

40

50

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。

【0228】

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

【0229】

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。

20

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験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。

【0230】

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、実際に臨床的症狀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

【0231】

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シー条件下で、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。

40

【0232】

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) 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

50

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。fSCCPでは、オリゴヌクレオチドプライマーを蛍光性に標識する。それによってDNAシーケンシング機などの高処理機器でアンプリマー(amplimer)の検出が可能になる。更に、インシリコSNP(in silico SNP, isSNP)と呼ばれる配列データベース分析法は、一般的なコンセンサス配列へ構築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。

10

【0233】

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はまた、嚢胞性線維症、鎌状赤血球貧血、慢性肉芽腫性疾患等の単一遺伝子病の転帰の違いを研究するために有用である。例えば、マンノース結合レクチンでの変異体(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わかっている。SNPにはまた薬理ゲノミクスで有用性がある。薬理ゲノミクスは、生命を脅かす毒性のように、ある薬剤への患者の応答に影響する遺伝的変異体を同定する。例えば、N-アセチルトランスフェラーゼにおける或る変異は抗結核剤、イソニアジドに反応した末梢神経障害の発生率が高くなるが、ALOX5遺伝子のコアプロモータの変異は5-リポキシゲナーゼ経路を標的とする抗喘息薬での治療に対する臨床的反応を減少する。異なった集団でのSNPの分布についての分析は遺伝的浮動、突然変異、組み換えおよび選択の研究に有用であると共に、集団の起源と移動の調査にも有用である(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)。

20

【0234】

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 159:235-244; Duplaa, C. 他(1993) Anal. Biochem. 212:229-236を参照)。目的のオリゴマーが種々の希釈液中に存在し、分光光度法または比色反応によって定量が迅速になるような高処理フォーマットのアッセイを行うことによって、複数のサンプルの定量速度を加速することができる。

30

【0235】

更に別の実施例では、本明細書に記載した任意のポリヌクレオチド配列由来のオリゴヌクレオチドまたはより長い断片を、マイクロアレイにおけるエレメントとして用いることができる。多数の遺伝子の関連発現レベルを同時にモニターする転写イメージング技術にマイクロアレイを用いることが可能である。これについては、以下に記載する。マイクロアレイはまた、遺伝変異体、突然変異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有効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。

40

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。

【0236】

50

或る実施例は、或る組織または細胞タイプの転写イメージを生成するような本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞タイプによ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、所定の条件下で所定の時間に発現した遺伝子の数及び相対存在量を定量することにより分析される（Seilhamer 他、米国特許第5,840,484号の「Comparative Gene Transcript Analysis」を参照。これらを引用することを以って本明細書の一部とする）。従って、特定の組織または細胞タイプの転写物または逆転写物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。 10

【0237】

転写イメージは、組織、株化細胞、生検またはその生物学的サンプルから単離した転写物を用いて生成し得る。転写イメージはしたがって、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。

【0238】

本発明のポリヌクレオチドの発現のプロファイルを作製する転写イメージはまた、工業的または天然の環境化合物の毒性試験のみならず、*in vitro* モデル系及び薬剤の前臨床評価と併せて使用し得る。全ての化合物は、作用及び毒性のメカニズムを示す、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.* 24:15 3-159、Steiner, S. 及び N.L. Anderson (2000) *Toxicol. Lett.* 112-113:467-471、該文献は特に引用することを以って本明細書の一部となす)。試験化合物が、既知の毒性を有する化合物のシグネチャと同様のシグネチャを有する場合には、毒性特性を共有している可能性がある。+フィンガープリントまたはシグネチャは、多数の遺伝子及び遺伝子ファミリーからの発現情報を含んでいる場合には、最も有用且つ正確である。理想的には、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。毒性シグネチャの要素に遺伝子機能を割り当てることで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、シグネチャ群の統計的マッチングには、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。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。 20 30

【0239】

或る実施例では、核酸を含有する生物学的サンプルを試験化合物で処理することにより試験化合物の毒性を算定する。処理した生物学的サンプル中で発現した核酸は、本発明のポリヌクレオチドに特異的な 1 つ以上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物レベルを、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

【0240】

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 50

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。或る実施例では、1次元等電点電気泳動によりサンプルからタンパク質を分離し、2次元ドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に応じて分離するような2次元ゲル電気泳動により、分離が達成される（前出のSteiner および Anderson）。タンパク質は、通常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された、または未処理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クエンシングする。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

【0241】

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。或る実施例では、マイクロアレイ上でエレメントとして抗体を用い、マイクロアレイをサンプルに曝して各アレイ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、各アレイのエレメントにおける蛍光結合の量を検出し得る。

20

【0242】

プロテオームレベルでの毒性シグネチャも中毒学的スクリーニングに有用であり、転写レベルでの毒性シグネチャと並行して分析するべきである。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の存在量の相関が乏しいこともあるので（Anderson, N. L. 及び J. Seilhamer (1997) Electrophoresis 18:533-537）、転写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ァイルを変化させるような化合物の分析においてプロテオーム毒性シグネチャは有用たり得る。更に、体液中の転写物の分析は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

【0243】

別の実施例では、タンパク質を含有する生物学的サンプルを試験化合物で処理することにより試験化合物の毒性を算定する。処理された生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、未処理生物学的サンプル中の対応するタンパク質の量と比較する。両サンプルのタンパク質量の差は、処理サンプル中の試験化合物に対する毒性反応を示す。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クエンシングし、これら部分配列を本発明のポリペプチドと比較することにより同定する。

40

【0244】

別の実施例では、タンパク質を含有する生物学的サンプルを試験化合物で処理することにより試験化合物の毒性を算定する。生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、未処理生物学的サンプル中のタンパク質の量と比較する。両サンプルのタンパク質量の差は、処理サンプル中の試験化合物に対する毒性反応を示す。

【0245】

50

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、Sчена, M. 他 (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619、Baldeschweiler 他 (1995) PCT出願第W095/251116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。

【0246】

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、コード配列より非コード配列の方が好ましい。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。核酸配列は、特定の染色体、染色体の特定領域または人工形成の染色体、例えば、ヒト人工染色体 (HAC)、酵母人工染色体 (YAC)、細菌人工染色体 (BAC)、細菌P1産物、或いは単一染色体cDNAライブラリに対してマッピングされる (例えば、Harrington, J.J. 他 (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and 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) と関連するような遺伝子連鎖地図を作製できる。 (例えば、Lander, E.S. 及び D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照)。

【0247】

蛍光 *in situ* ハイブリッド形成法 (FISH) は、他の物理的及び遺伝地図データと関連し得る (前出のHeinz-Ulrich, 他 (1995) in Meyers, 965-968頁等を参照)。遺伝地図データの例は、種々の科学雑誌あるいはOnline Mendelian Inheritance in Man (OMIM) のウェブサイトに見ることができる。物理的な染色体地図上のSECPをコードする遺伝子の位置と、特定の疾患または特定の疾患の素因との相関性は、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。

【0248】

確定した染色体マーカーを用いた連鎖分析等の物理的マッピング技術及び染色体標本原位ハイブリッド形成法を用いて、遺伝地図を拡張することができる。例えばマウスなど別の哺乳類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、または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領域等、特定の遺伝子領域への遺伝的結合によって大まかに位置決めがなされると、該領域にマップされる任意の配列は、更なる調査のための関連遺伝子あるいは調節遺伝子を提示している可能性がある (Gatti, R.A.他 (1988) Nature 336:577-580等を参照)。転座、反転等に起因する、健

【0249】

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。

【0250】

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 application W084/03564等を参照)。この方法においては、多数の様々な低分子の試験用化合物を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。

【0251】

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

【0252】

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。

【0253】

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。したがって、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

20

かようにも本発明を限定するものではない。
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は、米国特許出願第60/282.112号、第60/280.527号、第60/282.702号、第60/283.855号、第60/357.002号、第60/339.236号および第60/343.718を含め、言及することをもって特に本明細書の一部となす。

【実施例】

【0254】

1 cDNAライブラリの作製

Incyte cDNA群は、LIFESQ GOLDデータベース (Incyte Genomics, Palo Alto CA) に記載されたcDNAライブラリ群に由来する。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、塩化セシウムのクッション液の上に重層して遠心分離するかクロロホルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方、或いは別の方法を用いて、溶解物からRNAを沈殿させた。

30

【0255】

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、DNA分解酵素で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

【0256】

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

50

、好適な制限酵素または酵素群でcDNAを消化した。殆どのライブラリに対して、cDNAのサイズの選択(300~1000 bp)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)、或いは分取用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に結合させた。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)、あるいはその誘導体である。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な大腸菌細胞に形質転換した。

10

【0257】

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

【0258】

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穴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 2 蛍光スキャナ(Labsystems Oy, Helsinki, Finland)を用いて蛍光分析的に定量した。

【0259】

30

3 シークエンシング及び分析

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

40

【0260】

Incyte cDNA配列に由来するポリヌクレオチド配列は、ベクター、リンカー及びポリ(A)配列を除去し、あいまいな塩基対をマスクすることによって有効性を確認した。その際、BLAST、動的プログラミング及び隣接ジヌクレオチド頻度分析に基づくアルゴリズム及びブ

50

プログラムを用いた。次に、IncyteのcDNA配列、またはその翻訳を公共のデータベース（例えばGenBankの霊長類及びげっ歯類、哺乳動物、脊椎動物、真核生物のデータベースと、BLOCKS、PRINTS、DOMO、PRODOM）と、ヒト、ラット、マウス、線虫、出芽酵母（*Saccharomyces cerevisiae*）、分裂酵母（*Schizosaccharomyces pombe*）、および驚口瘡カンジダ（*Candida albicans*）からの配列を含むPROTEOMEデータベース（Incyte Genomics, Palo Alto CA）、及びPFAM等隠れマルコフモデル（HMM）に基づいたタンパク質ファミリーデータベース、INCYおよびTIGRFAM(Haft, D.H. 他(2001) *Nucleic Acids Res.* 29:41-43)並びに、SMART(Schultz 他(1998) *Proc. Natl. Acad. Sci. USA* 95:5857-5864; Letunic, I. 他(2002) *Nucleic Acids Res.* 30:242-244)のようなHMMに基づいたタンパク質ドメインデータベースから選択したデータベースに対して問い合わせた。（HMMは、遺伝子ファミリーのコンセンサス1次構造を分析する確率的アプローチである。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構築した。あるいは、GenBank cDNA群、GenBank EST群、ステッチされた配列群、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の集団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の集団を、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、PFAM等の隠れマルコフモデル（HMM）ベースのタンパク質ファミリーデータベース群、INCY、およびTIGRFAM、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間の一致率も計算するMEGALIGNマルチシーケンサ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指定されるデフォルトパラメータを用いて作製する。

10

20

30

【0261】

Incyte cDNA及び完全長配列の分析及び構築に利用したツール、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は好適な参照文献であり、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、または確率値が低いほど、2配列間の相同性が高くなる）。

40

【0262】

完全長ポリヌクレオチド配列およびポリペプチド配列の組み立て及び分析に用いる上記のプログラムは、SEQ ID NO:26-50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。

【0263】

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、様々な生物からのゲノム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. および S. Karlin (1997) *J. Mol. Biol.* 268:78-94 及びBurge, C. 及び S. Karlin (1998) *Cuf*

50

f. Opin. Struct. Biol. 8:346-354参照)。プログラムは予測されたエキソンを連結し、メチオニンから停止コドンに及ぶ構築されたcDNA配列を形成する。Genscanの出力は、ポリヌクレオチド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、30 kb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もまた、分泌タンパク質として注釈を付けたインサイトcDNA配列に対する相同性を基に同定した。こうして選択されたGenscan予測配列は、次にBLAST分析により公共データベース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に記載した構築プロセスを用いて、Incyte cDNA配列および/または公共cDNA配列でGenscan予測コード配列を構築して得た。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

【0264】

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

20

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構築された部分cDNAは、ゲノムDNAにマッピングし、関連するcDNA及び1つ以上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。区間の全長が、2つ以上の配列に在るような配列区間群をクラスター内で同定し、そのように同定された区間群は、推移性により、等しいと考えた。例えば、1つのcDNA及び2つのゲノム配列に或る区間が存在する場合、3つの区間は全て等しいと考えられた。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ステッチアルゴリズムで縫い合わせ、可能な最も長い配列および変異配列を作製する。1種類の親配列に沿って発生した区間間の連鎖 (cDNA - cDNAまたはゲノム配列 - ゲノム配列) は、親の種類を変える連鎖 (cDNA - ゲノム配列) に優先した。結果として得たステッチ配列群を翻訳し、BLAST分析で公共データベースgenpeptおよびgbpriと比較した。Genscanが予測した不正確なエキソン群は、genpeptからのトップBLASTヒットとの比較により補正した。必要な場合には、追加cDNA配列を用いるかゲノムDNAの検査により配列を更に伸長させた。

30

【0265】

ストレッチ配列 (Stretched Sequence)

40

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。先ず、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に記載したように構築された部分cDNAを問い合わせた。次に、最も近いGenBankタンパク質相同体を、BLAST分析により、Incyte cDNA配列または実施例 4に記載のGenScanエキソン予測配列のいずれかと比較した。結果として得られる高スコアリングセグメント対 (HSP) を用いてキメラタンパク質を産出し、翻訳した配列をGenBankタンパク質相同体上にマッピングした。元のGenBankタンパク質相同体に対し、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

50

、相同ゲノム配列の付加によりストレッチすなわち伸長した。結果として得られるストレッチ配列を検査し、完全遺伝子を含んでいるか否かを判定した。

【0266】

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:26-50を構築するために用いた配列を、BLAST及び他のSmith-Watermanアルゴリズムの実装を用いて、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:26-50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組み入れた(表7)。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確認した。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

【0267】

地図上の位置は、ヒト染色体の範囲または間隔として表される。センチモルガン単位での或る区間の地図上の位置は、染色体のpアームの末端に関して測定する(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。もっとも、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よっ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

20

【0268】

この方法で、SEQ ID NO:29は、97.20~115.70センチモルガンの間隔内で3番染色体にマッピングされた。SEQ ID NO:46は、111.10~123.50センチモルガン以内の間隔で染色体4にマッピングする。SEQ ID NO:47は、61.40~65.80センチモルガン以内で染色体19に、また、62.00~64.80センチモルガン以内で染色体19にマッピングされた。

30

【0269】

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が結合されている膜への標識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)。

【0270】

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一致、或いは類似的な一致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。検索の基準は積スコアであり、次式で定義される。

40

【0271】

【数1】

(BLASTスコア×配列一致率)

5 × (長さ(配列1), 長さ(配列2))の最小値

【0272】

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

50

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。BLASTスコアを計算するには、或る高スコアリングセグメント対 (HSP) 内の一致する各塩基に + 5 のスコアを割り当て、各不一致塩基に - 4 を割り当てる。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。

10

【0273】

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配列群を用いて構築される (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。すなわち心血管系、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、血液および免疫系、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。

20

【0274】

8 SECPをコードするポリヌクレオチドの伸長

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

30

【0275】

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。

【0276】

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穴プレート内で行った。反応混合液は、鋳型DNA及び200 nmolの各プライマーを有する。また、 Mg^{2+} と $(NH_4)_2SO_4$ と2-メルカプトエタノールを含むバッファー、Taq DNAポリメラーゼ (Amersham Pharmacia Biotech)、ELONGASE酵素 (Life Technologies)、Pfu DNAポリメラーゼ (Stratagene) を含む。プライマーの組、PCI AとPCI Bに対して以下のパラメータで増幅を行った。ステップ1: 94, 3分、ステップ2: 94, 15秒、ステップ3: 60, 1分、ステップ4: 68, 2分、ステップ5: ステップ2、3および4を20回反復する。ステップ6: 68, 5分、ステップ7: 4で保存する。別法では、プライマーの組、T7とSK+に対して以下のパラメータで増幅を行った。ステップ1: 94, 3分、ステップ2: 94

40

50

、15秒、ステップ3: 57、1分、ステップ4: 68、2分、ステップ5: ステップ2、3および4を20回反復する。

ステップ6: 68、5分、ステップ7: 4で保存する。

【0277】

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が試薬と結合できるようにして測定する。サンプルの蛍光を計測してDNAの濃度を定量するためにプレートをFluoroskan II (Labsystems Oy, Helsinki, Finland)でスキャンした。反応混合物のアリコート5~10µl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。

10

【0278】

伸長したヌクレオチドは、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。伸長したクローンを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部分を満たし、大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りとってLB/2Xカルベニシリン培養液の384ウェルプレートに37で一晚培養した。

20

【0279】

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、3分、ステップ2: 94、15秒、ステップ3: 60、1分、ステップ4: 72、2分、ステップ5: ステップ2、3および4を29回反復する。

ステップ6: 72、5分、ステップ7: 4で保存する。上記のようにPICOGREEN試薬(Molecular Probes)でDNAを定量した。DNAの回収率が低いサンプルは、上記と同一の条件を用いて再増幅した。サンプルは20%ジメチルスルホキシド(1:2, v/v)で希釈し、DYENAMIC MIC 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。

30

【0280】

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

【0281】

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

40

一塩基多型(SNP)として知られる一般的なDNA配列変異体がLIFESEQ データベース(Incyte Genomics)を用いてSEQ ID NO:26-50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を共にクラスターにして構築し、これによって遺伝子内のすべての配列変異体の同定ができた。一連のフィルタからなるアルゴリズムを使って、SNPを他の配列変異体から区別した。前段フィルタは、最小限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、配列アライメントエラーや、ベクター配列、キメラおよびスプライス変異体の不適当なトリミングにより生じるエラーを取り除いた。染色体の高度解析の自動化手順により、推定SNPの近傍におけるオリジナルのクロマトグラムファイルが解析された。クローンエラーフィルタ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、または体細胞突然変異

50

によって引き起こされるエラーのような、実験処理時に導入されるエラーを識別した。クラスターエラーフィルターは、統計的に生み出されたアルゴリズムを用いて、近縁の相同体または偽遺伝子のクラスター化に起因するエラー、または非ヒト配列によるコンタミネーションにより生じたエラーを同定した。最後のフィルター群によって、免疫グロブリンまたはT細胞受容体に存在する重複(duplicates)とSNPが除去された。

【0282】

高速処理MASSARRAY システム(Sequenom, Inc.) を用いて質量分析によってさらに特徴付けるために数種のSNPを選択して、四つの異なったヒト母集団中のSNP部位における対立遺伝子発生頻度を分析した。白人母集団は、ユタ州の83人、フランス人4人、ベネズエラ3人およびアーミッシュ派2人を含む92人(男性46人、女性46人)で構成された。アフリカ人母集団はすべてアフリカ系アメリカ人である194人(男性97人、女性97人)からなる。ヒスパニック母集団はすべてメキシコ系ヒスパニックの324人(男性162人、女性162人)からなる。アジア人母集団は126人(男性64人、女性62人)からなり、親の内訳は中国人43%、日本人31%、コリアン13%、ベトナム人5%およびその他のアジア人8%と報告されている。対立形質の発生頻度は最初に白人母集団において分析し、いくつかの例において、この母集団で対立形質分散を示さなかったSNPは他の三つの母集団においてさらに検査しなかった。

10

【0283】

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:26-50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は、OLIGO 4.06ソフトウェア(National Biosciences)等の最新技術のソフトウェアを用いて設計し、各オリゴマー50 pmolと、[^{-32}P]アデノシン3リン酸(Amersham Pharmacia Biotech) 250 μCi と、T4ポリヌクレオチドキナーゼ(DuPont NEN, Boston MA)を混合することにより標識する。標識したオリゴヌクレオチドは、SEPHADEX G-25超細繊維分子サイズ排除デキストラン ビーズカラム(Amersham Pharmacia Biotech)を用いて実質的に精製する。Ase I、Bgl II、Eco RI、Pst I、Xba IまたはPvu II(DuPont NEN)のいずれか1つのエンドヌクレアーゼで消化したヒトゲノムDNAの典型的な膜ベースのハイブリダイゼーション解析において、毎分 10^7 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。

20

30

【0284】

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°Cで16時間行う。非特異的シグナルを除去するため、例えば0.1Mクエン酸ナトリウム食塩水及び0.5%ドデシル硫酸ナトリウムに一致する条件下で、プロットを室温で順次洗浄する。オートラジオグラフィまたはそれに代わるイメージング手段を用いてハイブリダイゼーションパターンを視覚化し、比較する。

【0285】

11 マイクロアレイ

マイクロアレイ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、均一且つ、無孔の表面を持つ固体とするべきである(Schena (1999) 前出)。推奨する基板には、シリコン、シリカ、スライドガラス、ガラスチップおよびシリコンウエハがある。あるいは、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、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およ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、Mars

40

50

hall, A.及びJ. Hodgson (1998) Nat. Biotechnol. 16:27-31等を参照)。

【0286】

完全長cDNA、発現配列タグ(EST)、またはその断片またはオリゴマーは、マイクロアレイのエLEMENT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)など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ELEMENT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ELEMENT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ELEMENT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。一実施態様におけるマイクロアレイの調製および使用について、以下に詳述する。

10

【0287】

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サンプルは、MMLV逆転写酵素、0.05 pg/ μ lのオリゴ(dT)プライマー(21mer)、1X 第1鎖バッファ、0.03 unit/ μ lのRNアーゼ阻害剤、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、200 ngのポリ(A)⁺RNAを含有する体積25 ml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、2.5 mlの0.5 M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。混合した後、2つの反応サンプルは、1 mlのグリコーゲン(1 mg/ml)、60 mlの酢酸ナトリウム及び300 mlの100%エタノールを用いてエタノール沈殿させる。サンプルは次に、SpeedVAC(Savant Instruments Inc., Holbrook NY)を用いて完全に乾燥させ、14 μ lの5 \times SSC/0.2% SDS中で再懸濁する。

20

30

【0288】

マイクロアレイの調製

本発明の配列を用いて、アレイELEMENTを作製する。各アレイELEMENTは、クローン化cDNAインサートを持つベクター含有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1~2 ngの初期量から5 μ gより大きい最終量までアレイELEMENTを増幅する。増幅したアレイELEMENTは、SEPHACRYL-400(Amersham Pharmacia Biotech)を用いて精製する。

40

【0289】

精製したアレイELEMENT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。

【0290】

米国特許第5,807,522号に記載されている方法を用いて、コーティングしたガラス基板に

50

アレイエレメントを付加する。この特許は引用することを以って本明細書の一部とする。平均濃度が100 ng/ 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5 nlのアレイエレメントサンプルを加える。

【0291】

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

10

【0292】

ハイブリダイゼーション

ハイブリダイゼーション反応液は、5 \times SSC、0.2% SDSハイブリダイゼーション緩衝液にCy3及びCy5標識したcDNA合成産物を各0.2 μ g含む9 μ lのサンプル混合体を含めたものである。サンプル混合体は、65 まで5分間加熱し、マイクロアレイ表面上で等分して1.8 cm^2 のカバーガラスで覆う。アレイは、顕微鏡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。チェンバー内を湿度100に保つため、140 μ lの5 \times SSCをチェンバーの1コーナーに加える。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 \times 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 \times SSC) において各々45 で10分間3

20

【0293】

検出

レポーター標識されたハイブリダイゼーション複合体は、Cy3の励起のためには488 nm、Cy5の励起のためには632 nmでスペクトル線を生じ得るInnova 70混合ガス10 Wレーザー (Coherent, 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 \times 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1.8 $\text{cm} \times 1.8 \text{cm}$ のアレイは、解像度20 μm でスキャンする。

30

【0294】

2回の異なるスキャンで、混合ガスマルチラインレーザーは2つのフルオロフォア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565 nm、Cy5では650 nmである。装置は両方のフルオロフォアからのスペクトルを同時に記録し得るが、レーザー源において好適なフィルタを用いて、フルオロフォア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。

【0295】

スキャンの感度は通常、既知濃度でサンプル混合体に添加されるcDNA対照種により生成されるシグナル強度を用いて較正する。アレイ上の特定の位置には相補的DNA配列が含まれ、その位置におけるシグナルの強度を重量比1:100,000でハイブリダイゼーション種と相関させる。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、各々異なるフルオロフォアで標識し、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、2つのフルオロフォアで較正するcDNAのサンプルを標識し、ハイブリダイゼーション混合液に各々等量を加えることによって較正を行う。

40

【0296】

光電子増倍管の出力は、IBM互換性PCコンピュータにインストールされた12ビットRTI-8

50

35H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直線的 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つの異なるフルオロフォアを同時に励起及び測定する場合には、各フルオロフォアの発光スペクトルを用いて、データは先ずフルオロフォア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。

【0297】

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、GEMT00LS遺伝子発現分析プログラム (Incyte) である。

10

【0298】

発現

このようにして、SEQ ID NO:46が5つのヒト乳癌細胞系の中4つに、また6つのヒト肺腺癌細胞および肺扁平上皮細胞悪性腫瘍の中5つに少なくとも2.2倍下方制御されていることが示された。これらの実験はSEQ ID NO:21 と SEQ ID NO:46 が疾患進展度診断において有用性があり、あるいは癌および腫瘍の潜在的治療標的としてSEQ ID NO:21 を使うことに有用性があることを示す。

【0299】

ヒト上皮細胞系の増殖

HMECは健常乳房組織 (Clonetics San Diego, CA) に由来するヒト初代乳腺上皮細胞株である。次の細胞系はATCC (Manassus, VA) から得られた。Sk-BR3 は43才の女性の悪性胸水から単離された乳腺癌細胞系である。MCF7は69才の女性の胸水から由来する乳腺癌細胞系である。T47Dは乳房浸潤性管癌を有する54才女性の胸水から由来する乳癌細胞系である。BT20は74才の女性から単離された腫瘍塊の薄片から *in vitro* で遊出される細胞群由来の乳癌細胞系である。MDA-mb-435S と MDA-mb-231 は転移性乳癌を有する女性 (51才) の胸水から由来する転移性乳房腫瘍細胞系である。全ての細胞培養はATCCの推薦にしたがった培養液内で増殖させた。また、RNA単離の前に70-80% コンフルエンスに増殖させた。

20

【0300】

乳房腫瘍進行の種々の段階 (BT20、MCF7、MDA-mb-231、MDA-mb-435S、SKBr3およびT47D) を示す6つの腫瘍細胞系からのcDNAの発現を非悪性乳腺上皮細胞系 (HMEC) の発現と比較した。発現の変化が少なくとも2倍、また250単位以上のシグナル強度、シグナルとバックグラウンドの割合が少なくとも2.5、および少なくとも40%のエレメントのスポットサイズを示すアレイエレメントは、差次的発現されるものとGEMTOOLS プログラム (Incyte Genomics) を用いて同定された。SEQ ID NO:28の発現はT47D細胞に対して2.75倍低く、MD-am-B435S細胞に対して3.7倍低く、BT20細胞に対して4.6倍低く、SKBr3細胞に対して4.7倍低く、MCF7細胞に対して5.1倍低く、MD-am-B231細胞に対して1.9倍低かった。これらの実験によって、SEQ ID NO:28 は試験した乳癌細胞系の6種のうち5種において発現の有意な低下が示された、さらに、SEQ ID NO:28 の乳癌の診断的マーカーとして、或いは治療標的としての利用性が確立された。

30

40

【0301】

例えば、SEQ ID NO:48 は炎症性応答において差次的発現を示すことがマイクロアレイ分析によって確認された。SEQ ID NO:48 の発現は、100 nM PMA (phorbol 12myristate 13acetate) で48時間刺激しておいたTHP-1ヒト前単球系において、未処理のTHP-1細胞に比べて少なくとも2倍増加した。THP-1は、急性単球性白血病の男児 (1才) の末梢血に由来する前単球系である。この細胞系はPMAで刺激されると単球の特性を獲得する。単球は炎症性免疫応答の開始および保持において重要な役割を果たす。

【0302】

他の例において、SEQ ID NO:48 の発現はインターロイキン-10 (IL-10) で処理した末梢血

50

単核細胞(PBMC、12% Bリンパ球、40% Tリンパ球、20% NK細胞、25% 単球および3%の種々の細胞(樹状突起細胞と前駆体細胞を含む))において少なくとも2倍減少していた。IL-10は種々の細胞タイプに免疫賦活的作用か、または免疫抑制的作用のどちらかを及ぼす多面的サイトカインである。IL-10は免疫賦活性サイトカインとしてBリンパ球に作用して、Bリンパ球の生存度、細胞増殖、Ig分泌およびクラスII MHCの発現を亢進し得る。Bリンパ球の他に、IL-10はまた、細胞毒性T細胞発生のエンハンサーであるとともに、胸腺細胞および肥満細胞の増殖共刺激物質でもある。IL-10は免疫抑制物質サイトカインとしては、単球またはマクロファージの活性およびその結果生じる細胞毒性作用の強力なインヒビターである。これはTNF- α 、IL-1、IL-6およびIL-10などの多くのサイトカインの産生を抑制し、また活性化単球/マクロファージによるスーパーオキシドアニオン、活性酸素反応中間体および活性窒素反応中間体の合成を抑制することができる。

10

【0303】

なお別の実施例において、SEQ ID NO:48の発現は未処理のPBMC群に比べて、混合リンパ球反応(MLR)のヒトPBMCにおいて少なくとも2倍減少した。同種の遺伝的に異なった動物からのリンパ球を組織培養と混合すると、混合リンパ球反応(MLR)が生じる。MLRはT細胞活性化の認められたモデルであり、移植片対宿主病や移植拒絶(宿主対移植片病)を含む、このような活性化が起こる多くの症候群を研究するのに関連する。この実験において、ヒトPBMCは標準の勾配分離法を用いて6人の健常ボランティアドナーの血液から採取された。

【0304】

上記の三つの実験によってSEQ ID NO:48が炎症応答の診断試験において有用であることがわかる。

20

【0305】

別の実施例において、SEQ ID NO:48は、初代乳腺上皮細胞に対して特定の乳癌および乳腺細胞系において示差的な発現を示すことがマイクロアレイ分析によって確認された。この乳癌細胞系には、BT20(74才女性から単離された腫瘍塊の薄切片から *in vitro*で遊出される細胞群由来の乳癌細胞系)、BT474(60才女性乳房の充実性で浸潤性の乳管癌から単離された乳管癌細胞系)、BY483(23才、正常月経、経産女性からの乳頭浸潤性管腫瘍から単離された乳管癌細胞系)、HS578T(乳癌の74才女性から単離された乳管癌細胞系)、MCF7(69才女性の胸水由来の乳腺癌細胞系)、およびMDA-mb-468(乳房の転移性腺癌を有する51才女性の胸水から単離された乳腺癌細胞系)が含まれる。乳房線維嚢胞病(fibrocystic breast disease)の36才女性から単離された乳房乳腺(管腔管(luminal ductal)の特性)細胞系であるMCF10Aもまた比較した。初代乳房上皮細胞系のHMECは健常乳房組織(Clonetics San Diego, CA)から得た。マイクロアレイ実験によると、SEQ ID NO:48の発現は初代乳房上皮細胞系HMECからの細胞に比べて、HS578T乳癌系で少なくとも4倍減少し、MDA-mb-468乳癌系とMCF10A乳房乳腺細胞系で少なくとも2倍減少した。したがって、SEQ ID NO:48は乳癌および乳房線維嚢胞病の診断的マーカーとして、あるいは、潜在的治療標的として有用である。

30

【0306】

更なる実施例において、SEQ ID NO:48の発現は健常結腸組織に比べて腫瘍性結腸組織において少なくとも2倍増加することがマイクロアレイ分析によって確認された。腫瘍性結腸組織と健常結腸組織は共に、粘液性腺癌と診断された58才女性から単離した。したがって、SEQ ID NO:48は結腸癌の診断的マーカーとして、あるいは潜在的治療標的として有用である。

40

【0307】

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

50

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。

【0308】

13 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としては、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、および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レベルのcDNA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)。

10

20

【0309】

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(QIAGEN)上での精製を可能にする。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。

30

【0310】

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。リボソーム製剤あるいは電気穿孔法を用いて、5~10µgの組換えベクターをヒト細胞株、例えば内皮細胞系由来または造血細胞系由来のヒト細胞系に、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1~2µgのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入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-GF

40

50

P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、前方散乱光と90°側方散乱光によって計測される細胞サイズと粒度の変化、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容、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容とがある。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。

10

【0311】

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) の保存された複数の領域に結合する。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か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。

20

【0312】

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; 例えば、Harrington, M.G. (1990) *Methods Enzymol.* 182: 488-495を参照) または他の精製技術で実質的に精製されたSECPを用いて、標準的なプロトコールで動物 (例えば、ウサギやマウス等) を免疫化して抗体を作り出す。

【0313】

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)。

30

【0314】

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。

40

【0315】

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。

【0316】

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

50

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。

【0317】

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、 ^{125}I 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。

10

【0318】

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。

【0319】

SECPはまた、高処理型の酵母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つの大きなライブラリによってコードされるタンパク質間の全ての相互作用を決定することができる(Nandabalan, K. 他 (2000) 米国特許第6,057,101号)。

20

【0320】

18 SECP活性の実証

SECPのペプチジルプロリルシス/トランスイソメラーゼ活性は、Rahfeld, J.U., 他 (1994) (FEBS Lett. 352: 180-184)によって記載されている酵素アッセイによって測定できる。このアッセイは、キモトリプシン(0.5 mg/ml)と種々の濃度のSECPを含む35 mMのHEPESバッファー(pH 7.8)で10分に行う。これらのアッセイの条件下では、基質であるSuc-Ala-Xaa-Pro-Phe-4-NAはプロリル結合に関して平衡になり、トランス結合が80～95%であり、シス結合が5～20%である。ジメチルスルフォキシド(10 mg/ml)中に溶解した基質のアリコット(2 ul)を上記の反応混合液に加える。トランスからシスへの変換は、キモトリプシンによるシス結合体の加水分解によって4-ニトロアニリドが生成され、これは390 nmでのその吸光度によって検出される。4-ニトロアニリドは時間依存的に生じ、アッセイにおけるSECPの量に比例する。

30

【0321】

あるいは、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., eds. (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 $[\text{}^3\text{H}]$ 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 $[\text{}^3\text{H}]$ 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 $[\text{}^3\text{H}]$ チミジンの最大の組み込みを表す。

40

【0322】

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

50

。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、S E C Pの神経調節効果を示す (Morimoto, T.他 (1995) Neuron 15:689-696)。

【0323】

別法では、S E C P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S E C P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、小さい膜で囲まれた小胞、およびその他の分泌細胞小器官のような、細胞内区画の単離を可能とする。分画し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、S E C P特異抗体を用いて実行され、また免疫沈降サンプルは、SDS-PAGEおよび免疫ブロット技術を用いて解析される。全細胞ライセート中のS E C Pに対する分泌性細胞小器官中のS E C P濃度は、分泌経路を通るS E C Pの量に比例する。

10

【0324】

別法では、AMP結合活性の測定を、S E C P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、S E C Pの活性に比例する。

【0325】

19 免疫グロブリン活性の実証

S E C P活性の或るアッセイでは、S E C P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)。S E C Pを、当分野で既知の方法で同位体標識する。一定量の標識S E C Pに種々の血清濃度を加える。S E C P-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S E C P-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S E C P-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S E C P-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S E C P-抗原複合体の量は、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S E C P活性の測定量である。

20

【0326】

別法として、S E C P活性の或るアッセイは、細胞表面におけるS E C Pの発現を測定する。S E C P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、ビオチンで標識される (de la Fuente, M.A. 他 (1997) Blood 90:2398-2405)。S E C P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S E C Pの量に比例する。

30

【0327】

別法で、S E C P活性の或るアッセイは、S E C P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S E C P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、S E C Pの活性の直接の測定値となる。

40

【0328】

当業者は、本発明の範囲及び精神から逸脱することなく本発明に記載した方法及びシステムの種々の改変を行い得る。本発明について説明するにあたり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、特許請求の範囲内にあ

50

るものとする。

【0329】

(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。

【0330】

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GenBank相同体が一致する確率スコアも併せて示す。

【0331】

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

10

【0332】

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

【0333】

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。

【0334】

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。

20

【0335】

表7は、本発明のポリヌクレオチドおよび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、参照文献及び閾値パラメータと共に示す。

【0336】

【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	ポリスクレオチド SEQ ID NO:	Incyte ポリスクレオチド ID	CA2 試験
7757335	1	7757335CD1	26	7757335CB1	
1482539	2	1482539CD1	27	1482539CB1	
2234213	3	2234213CD1	28	2234213CB1	
1345785	4	1345785CD1	29	1345785CB1	90078841CA2
3807190	5	3807190CD1	30	3807190CB1	
4856078	6	4856078CD1	31	4856078CB1	4856078CA2, 90079645CA2
6106886	7	6106886CD1	32	6106886CB1	56050011CA2, 6106886CA2, 90079304CA2, 90079312CA2, 90079320CA2, 90079344CA2, 90079412CA2, 90079420CA2, 90079436CA2, 90079444CA2, 90079582CA2, 90079690CA2
5627037	8	5627037CD1	33	5627037CB1	
3688835	9	3688835CD1	34	3688835CB1	3688835CA2
2295419	10	2295419CD1	35	2295419CB1	5391296CA2
2437896	11	2437896CD1	36	2437896CB1	6519813CA2
2641482	12	2641482CD1	37	2641482CB1	
7494292	13	7494292CD1	38	7494292CB1	90067354CA2
7636288	14	7636288CD1	39	7636288CB1	90089678CA2, 90089694CA2
8095391	15	8095391CD1	40	8095391CB1	6053256CA2, 7644258CA2, 8095391CA2, 8324949CA2

10

20

30

40

【 0 3 3 7 】

【 表 1 - 2 】

表1-2

Incyte サイト ID	プロシ NO:	ポリペプチドSEQ ID	Incyte ポリペプチド ID	ポリスクレオチド SEQ ID NO:	Incyte ポリスクレオチド ID	CA2 試薬
532018	16		532018CD1	41	532018CBI	90090702CA2, 90090710CA2, 90090842CA2
7481063	17		7481063CD1	42	7481063CBI	
2044436	18		2044436CD1	43	2044436CBI	
4091564	19		4091564CD1	44	4091564CBI	4091540CA2, 4091564CA2, 4092602CA2
8039739	20		8039739CD1	45	8039739CBI	
1265837	21		1265837CD1	46	1265837CBI	8704942CA2
5568527	22		5568527CD1	47	5568527CBI	1358190CA2
7503641	23		7503641CD1	48	7503641CBI	
7503458	24		7503458CD1	49	7503458CBI	
7500925	25		7500925CD1	50	7500925CBI	

10

20

30

40

【 0 3 3 8 】

【 表 2 - 1 】

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	7757335CD1	g2138330	0	[ヒト]アセチル CoA カルボキシラーゼ Abu-Elheiga, L. 他 (1997) J. Biol. Chem. 272:10669-10677
2	1482539CD1	g1770526	1.6E-40	[ヒト]SRcyp タンパク質 Bourquin, J.P. (1997) Nucleic Acids Res. 25:2055-2061
3	2234213CD1	g211610	1.9E-20	[ニワトリ]新規のコラーゲンタンパク質 Marchant, J.K., 他 (1991) Proc. Natl. Acad. Sci. U.S.A. 88:1560-1564
4	1345785CD1	g15777917	0.0E+00	脂肪細胞特異的タンパク質 3[マウス]
16	532018CD1	g15778556	0.0E+00	α -1-B 糖タンパク質前駆体[ヒト]
		g11877348	2.3E-90	[ラット]推定 α 1B 糖タンパク質 Ishioka, N., 他 (1986) Proc. Natl. Acad. Sci. U.S.A. 83:2363-7
17	7481063CD1	g163825	1.5E-11	[イエネコ]主要抗原 I retbindin [ヒト]
18	2044436CD1	g13488611	3.0E-79	しおり糸シルクタンパク質スピロイン(spidroin)2[ジヨロウモ(Nephila clavata)]
19	4091564CD1	g16974791	4.0E-05	ヒト]bA367J7.2.1 (新規な免疫グロブリンドメイン含有タンパク質(アイソゾーム 1))
20	8039739CD1	g9930918	3.9E-294	SH2 ドメイン含有ホスファターゼアンカータンパク質 2c[ヒト]
20	8039739CD1	g16033594	0.0E+00	免疫グロブリンスーパーファミリー受容体転座関連タンパク質 3[ヒト]
20	8039739CD1	g18092655	0.0E+00	[マウス] 脂肪細胞特異的タンパク質 3
23	7503641CD1	g15777917	4.0E-56	[ヒト]タンパク質間およびタンパク質とリガンド間の相互作用に関する 2 個の免疫グロブリン(Ig)ドメインを含むタンパク質
24	7503458CD1	716615 MGC 3047	2.6E-80	[ヒト]retbindin
25	7500925CD1	g13488611	1.9E-43	[分裂酵母]DUF28 ドメインを含む保存されたタンパク質
		377422 pi0 53	8.4E-27	[出芽酵母]mitコンドリアミトコンドリアタンパク質合成に関与すると思われる。
		6910 YGR02 1W	8.2E-20	

【 0 3 3 9 】

【 表 2 - 2 】

10

20

30

40

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
		634850;orf 6.2546	2.9E-17	[驚口瘡カンジダ]特徴化されていないドメインを含むタンパク質は出芽酵母 Ygr021 に対する類似性が低い。Ygr021 はミトコンドリアタンパク質合成におそらく関与している。

10

20

30

40

【 0 3 4 0 】

【 表 3 - 1 】

表3-1

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
1	77573335CD1	2487	S35 S72 S78 S91 S105 S136 S148 S169 S174 S175 S183 S187 S192 S246 S278 S423 S469 S676 S705 S725 S876 S921 S998 S1018 S1060 S1226 S1242 S1278 S1437 S1443 S1477 S1613 S1639 S2099 S2276 S2297 S2467 S2479 T188 T206 T216 T300 T412 T668 T732 T909 T1031 T1086 T1301 T1371 T1491 T1505 T1773 T1825 T1894 T1961 T2024 T2034 T2054 T2085 T2113 T2350 T2427 T2443 Y1642 Y1774 Y2284	N118 N883 N1002 N1071 N1153 N1453 N1610 N1611 N1682 N2431	シグナルペプチド: M1-T16	SPSCAN
1					シグナルペプチド: M1-W19 カルボモイルリン酸シンターゼ(CPSase) K262-S685	HMMER HMMER_PFAM

10

20

30

40

50

表 3 - 2

SEQ ID NO:	inocyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					カルボキシトランスフェラーゼドメイン: G1794-L2363 ピオチンを必要とする酵素 T895-L961	HMMER_PPFAM HMMER_PPFAM
					膜貫通セグメント: L4-I22 Q528-K556 V822-H850 I1954-R1973	TMAP
					N 末端はサイトソル内にな ピオチンを必要とする酵素 BL00188: P900-I945	BLIMPS_BLOCKS BLIMPS_BLOCKS
					カルバモイルリン酸シクターゼサブドメイン BL00866: F490-P535, A1969-Y2003	PROFILESAN
					ピオチンを必要とする酵素 接着部位 biotin.prf: L905-A954	BLAST_PRODUM
					カルボキシラーゼアセチル CoA リガーゼ ピオチン ATP 結合 PD002092: H1177-D1632, V1629-G1793, S1070-F1144, W351-N367, I2384-W2404	BLAST_PRODUM
					カルボキシラーゼ リガーゼ ベータトランスフェラーゼ サブユニット 脂肪酸生成 PD001372: I1957-R2361, G1814-I1864, D241-L264	BLAST_PRODUM
					アセチル CoA カルボキシラーゼは以下を含む: ピオチン 脂肪酸生成リガーゼ PD011672: L2362-S2486 PD037401: M1-A155	BLAST_PRODUM
1 cont					ピオチンを必要とする酵素 接着部位 DM01562 P11029 1510-2277: T1681-L2441 Q00955 1445-2194: N1682-D2422 P32874 1491-2243: T1681-Q2416 DM01850 P11029 619-1213: E762-F1144 細胞接着配列 R727-D729	BLAST_DOMO
						MOTIFS

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	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					カルバモイルリン酸シグネチャ1: F449-G463	MOTIFS
					カルバモイルリン酸シグネチャ2: F578-L585	MOTIFS
2	1482539CD1	311	S18 S42 S53 S132 S185 T139 T170 T179 T227 T292	N81 N232 N253	Signal_peptide: M1-A40	SPSCAN
					サイクロフィリンタイプ ペプチジルプロリンシグネチャ1: F143-Y310	HMMER_PFAM
					サイクロフィリンタイプのペプチジルプロリンシグネチャ2: BL00170: G156-A182, Y193-N232, R240-L284	BLIMPS_BLOCKS
					サイクロフィリンタイプのペプチジルプロリンシグネチャ3: D165-D230	PROFILESSCAN
					サイクロフィリンのペプチジルプロリンシグネチャ4: PR00153: Q256-D268, R269-L284, L162-L177, F198-G210, G241-Q256	BLIMPS_PRINTS
					イソメラゼ ROTAMASE サイクロフィリン シグネチャ5: PP1ase シグネチャ6: PD000341: V144-D308	BLAST_PRODROM
2	cont				サイクロフィリンタイプのペプチジルプロリンシグネチャ7: DM00129 B47328 58-229: D147-G307 P30414 58-229: D147-G307 Q08752 13-182: V144-G307 P52011 1-170: V144-G307	BLAST_DOMC
					角質コラーゲンに類似 PD067228: P248-G333	BLAST_PRODROM

10

20

30

40

50

【 0 3 4 3 】

【 表 3 - 4 】

表3-4

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					前駆体シグナル コラーゲン α 31X 鎖 細胞外基質結合組 織 PD028299: G247-R326	BLAST_PRODROM
					繊維性コラーゲンカルボキシル末端 DM00019 P3906 455-643: S241-G333	BLAST_DOMO
					繊維性コラーゲンカルボキシル末端 DM00019 Q02388 1373-1547: L245-G333	BLAST_DOMO
					繊維性コラーゲンカルボキシル末端 DM00019 P29400 705-886: P246-G333	BLAST_DOMO
					繊維性コラーゲンカルボキシル末端 DM00019 P08572 778-947: P246-G333	BLAST_DOMO
					アスパラギン酸とアスパラギンヒドロキシ化部位 C150-C161	MOTIFS
					EGF 様ドメイン シグネチャ 2 C159-C174	MOTIFS
3	2234213CD1	406	S62 S156 S316 S363 S364 S376 S385 T34 T54 T175 T184 T201 Y131	N142 N182	シグナルペプチド: M1-R36	HMMER
					Signal_cleavage: M1-T34	SPSCAN
					コラーゲントリプルヘリックスリピート (20 コピ ー) : G271-P329	HMMER_PFAM
					EGF 様ドメイン C138-C174, C93-K126	HMMER_PFAM
					膜貫通ドメイン: L16-W33, E319-N347 N 末端は細胞質ゾル内	TMAP
					カルシウム結合 EGF 様ドメインタンパク質パターンタンパ ク質 BL01187: C132-G143, C150-Y165	BLIMPS_BLOCKS
					タイプ III EGF 様シグネチャ PR00010: D89-Q100, G155-Y165, A195-T201	BLIMPS_PRINTS

10

20

30

40

50

表3-5

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					コラーゲン α 前駆体鎖 リピート シグナル 結合 組織 細胞外 基質 PD000007: P246-G333	BLAST_PROD0M
					プレコラーゲン前駆体シグナル PD072959: N242-S334	BLAST_PROD0M
					角質コラーゲンに類似 PD067228: P248-G333	BLAST_PROD0M
					前駆体シグナル コラーゲン α 31X 鎖 細胞外基質結合組織 PD028299: G247-R326	BLAST_PROD0M
3					繊維性コラーゲンカルボキシル末端 DM00019 P3906 1455-643: S241-G333	BLAST_DOMO
cont					繊維性コラーゲンカルボキシル末端 DM00019 Q02388 1373-1547: L245-G333	BLAST_DOMO
					繊維性コラーゲンカルボキシル末端 DM00019 P29400 705-886: P246-G333	BLAST_DOMO
					繊維性コラーゲンカルボキシル末端 DM00019 P08572 778-947: P246-G333	BLAST_DOMO
					アスパラギン酸とアスパラギンヒドロキシル化部位 C150-C161	MOTIFS
					EGF 様ドメイン シグネチャ 2 C159-C174	MOTIFS
					カルシウム結合 EGF 様ドメインパターン シグネチャ:	MOTIFS
4	1345785CD1	442	S56 S93 S113 S256 S377 T194 T290	N119 N306	シグナルペプチド: M1-A30, M1-S26	HMMER
					Signal_cleavage: M1-S24	SPSCAN
					免疫グロブリン ドメイン: G179-L274, G47-L139	HMMER_FFAM
					膜貫通ドメイン: S334-A362	TMAP
					N 末端は細胞質ゾル内にはない	
5	3807190CD1	146			シグナルペプチド: M1-H20, M1-P19, M1-K25, M1-Y24, M1-P26	HMMER
					Signal_cleavage: M29-A89	SPSCAN

【 0 3 4 5 】

【 表 3 - 6 】

表3-6

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
6	4856078CD1	109	T98		シグナルペプチド: M1-G18	HMMER
6					Signal_cleavage: M1-T34	SPSCAN
					膜貫通ドメイン: M1-G33 N 末端は細胞質ゾル内にはない	TMAP
7	6106886CD1	78	T29		シグナルペプチド: M1-P27, M1-25	HMMER
					Signal_cleavage: M1-G24	SPSCAN
					膜貫通ドメイン: P4-I120N-末端は細胞質内にはない。	TMAP
8	5627037CD1	721	S57 S150 S157 S218 S294 S303 S491 S516 S538 S580 S602 S678 T36 T71 T90 T109 T128 T138 T147 T161 T181 T222 T265 T290 T326 T334 T496 T556 T667 T683	N88 N245 N435	免疫グロブリン ドメイン: G270-A337, G84-S150, R5-A50, G185-G209, G370-A438	HMMER_PPFAM
					膜貫通ドメイン: E455-F483 N 末端は細胞質ゾル内にはない	TMAP
					糖タンパク質抗原 PRE PD02327: L62-I73, T90-Q111	BLIMPS_PRODOM
					不規則交叉 CROUGHEST タンパク質前駆帯 IRREC 膜貫通免疫グロブリン フォールド糖タンパク質シグナル細胞接着 PD124347: G7-I177	BLAST_PRODOM
9	3688835CD1	144	S10 S79 S106 T36 T133	N82	Signal_cleavage: M1-G15	SPSCAN
10	2295419CD1	124	S102		signal_cleavage: M1-S17	SPSCAN
10					シグナルペプチド: M1-R30	HMMER

【表 3 - 7】

表 3 - 7

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
11	2437896CD1	142	S87 S99 T97		膜貫通ドメイン: S84-F107 N 末端は細胞質ソル内 signal cleavage: M1-G21 シグナルペプチド: M1-G21	TMAP SPSCAN HMMEER
12	2641482CD1	103	S83		膜貫通ドメイン: L109-I132 N 末端は細胞質ソル内 ATP/GTP 結合部位モチーフ A (P-loop) G80-S87 signal cleavage: M1-V24 シグナルペプチド: M1-V24, M1-Q26	MOTIFS SPSCAN HMMEER
13	7494292CD1	670	S47 S121 S129 S402 S417 S424 S441 S445 S491 S514 S584 T61 T106 T113 T156 T179 T222 T317 T327 T454 T494 T538 Y279	N57 N189	signal cleavage: M1-A27	SPSCAN
14	7636288CD1	215	S127 S181 T99 T130		シグナルペプチド: M1-V21, M1-C25, M1-P29, M1-A27 膜貫通ドメイン: H8-E36 N 末端は細胞質ソル内 リカバリンファミリーシグネチャ PR00450: L155-F169, P171-Q190 signal cleavage: M1-S40	HMMEER TMAP BLIMPS_PRINTS SPSCAN
15	8095391CD1	86	S51 S63 S78	N36	シグナルペプチド: M1-L24, M1-C26, M1-A30 signal cleavage: M1-A25 シグナルペプチド: M1-A25	HMMEER SPSCAN HMMEER

【表 3 - 8】

表3-8

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
16	532018CD1	495	S183 S237 S317 S365 S389 T186 T200 T257 T278 T445	N44 N179 N363 N371	signal_cleavage: M1-G16 シグナルペプチド: M1-V18, M1-A21, M1-A22 免疫グロブリンドメイン: L42-S95, G225-Y281, G416-Y472, G318-Y376 膜貫通ドメイン: L4-E20 N末端は細胞質ゾル内にはない $\alpha 1$ B-糖タンパク質免疫グロブリンフォールド糖タンパク質血症漿 PD138678: A22-A387 $\alpha 1$ B-糖タンパク質免疫グロブリンフォールド糖タンパク質血症漿 KIAA0364 PD101369: P388-E489 免疫グロブリン DM00001 P04217 192-268: M213-W290 DM00001 P04217 10-82: L31-L104 DM00001 P04217 100-170: L121-L192 signal_cleavage: M1-C32	SPSCAN HMMER HMMER_PFAM TMAP BLAST_PRODOM BLAST_PRODOM BLAST_DOMO
17	7481063CD1	60	T10 T49		シグナルペプチド: M1-I34, M11-C32, M11-R29, M11-I34 膜貫通ドメイン: T8-R29 N末端は細胞質ゾル内にはない Uteroglobin シグネチャ PR00486: K12-L26, S28-D43, D43-Y60 Uteroglobin ファミリングネチャ: L13-Y60 UTEROGLOBIN FAMILY DM02636 P30438 1-91: M11-Y60 Uteroglobin ファミリングネチャ I A30-L46	SPSCAN HMMER TMAP BLIMPS_PRINTS PROFILESCAN BLAST_DOMO MOTIFS

10

20

30

40

50

【 0 3 4 8 】

【 表 3 - 9 】

表3-9

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
18	2044436CD1	241	S62 S88 S142 S143		signal_cleavage: M1-G30, M7-A27, M7-G29, M1-G29, M7-A36 膜貫通ドメイン: M7-E26 N末端は細胞質ゾル内がない	SPSCAN TMAP
19	4091564CD1	174	S50 S62 S70 S142 T41		signal_cleavage: M1-G23	SPSCAN
20	8039739CD1	727	S81 S235 S261 S266 S399 S459 S492 S608 S617 S621 S696 S701 T170 T193 T328 T340 T368 T379 T472 T558 T691 Y349	N419 N431 N441 N512 N524 N534 N554 N561	signal_cleavage: M1-L19 免疫グロブリン ドメイン: G295-A353, G197-V255, G483-A539, G390-A446, G106-A158, G30-T77 膜貫通ドメイン: N561-Y589 N末端は細胞質ゾル内がない	HMMER_PPFAM TMAP
					受容体 FC 免疫グロブリン PD01270: P14-H53, I65-HI01, D107-N135 細胞表面糖タンパク質 GP42 前駆体シグナル GPI アンカ ー膜 PD116497: P376-S559, Q177-N355, M1-P23, F313-N448	BLIMPS_PRODROM BLAST_PRODROM
20 cont					受容体 FC ガンマ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: M1-K131 血小板内皮細胞接着前駆体シグナル分子 PECAM1 CD31 抗原 PD150932: S362-D694 ミエリン結合糖タンパク質 DM00682P123198-196: I9-V178	BLAST_PRODROM BLAST_PRODROM BLAST_PRODROM

10

20

30

40

50

表 3 - 1 0

SEQ ID NO:	inocyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
21	1265837CD1	132	S3 S38 S49 T31	N58	ミエリン結合糖タンパク質 DM00682[S42209]1-172: V12-V178 ミエリン結合糖タンパク質 DM00682[S40204]28-216: M1-E180 ミエリン結合糖タンパク質 DM00682[P27645]19-208: P14-Q179 N-6 アデニン特異的 DNA メチラーゼシグネチャ L18-W24 膜貫通ドメイン: T88-F111 N 末端は細胞質ゾル内がない	BLAST_DOMO BLAST_DOMO BLAST_DOMO MOTIFS TMAP
22	5568527CD1	297	S35 S120 S157 S284 T221	N155	HCP リポーターシグネチャ: D19-K70 signal_cleavage: M1-A19 シグナルペプチド: M1-A26	PROFILESCAN SPSCAN HMNER
23	7503641CD1	218	S56 S93 S113 S207	N119	signal_cleavage: M1-S24 シグナルペプチド: M1-A19, M1-V20, M1-L22, M1-G25, M1-A30, M1-S24, M1-S27 免疫グロブリンドメイン: G47-L139 E 値: 3.6E-3	SPSCAN HMNER HMNER_PPFAM
24	7503458CD1	266	S87 S113 S167 S168		signal_cleavage: M1-G30	SPSCAN
24 cont					シグナルペプチド: M7-A27, M7-G29, M7-R32, M1-G30, M1-G29, M7-A36	HMNER
25	7500925CD1	246	S35 S120 S157 T221	N155	Signal_cleavage: M1-A19 シグナルペプチド: M1-A26	SPSCAN HMNER

10

20

30

40

【表 3 - 1 1】

表3-11

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					COAT 部分の保存されたタンパク質遺伝子間領域 MG332 VMA7RPS31A F17H15.14 PD004186: Y138-P236, PD004323: K62-K127	BLAST_PRODOM
					HI0315; RUVC; 26.4; ASP5; DM02856 P24237 20-238: E76-D223 G64147 21-238: R77-D223 P44634 21-238: R77-D223 P53212 45-289: E76-G241	BLAST_DOMO

10

20

30

40

50

【 0 3 5 1 】
【 表 4 - 1 】

表4-1

ポリヌクレオチド SEQ ID NO/ Incyte 求 リヌクレオチド ID/ 配列長	配列断片
26/7757335CBI/ 7633	1-512, 1-663, 1-836, 5-516, 6-509, 10-772, 29-702, 160-334, 160-391, 160-555, 160-685, 160-711, 160-760, 160-819, 162-713, 168-684, 182- 335, 219-945, 227-891, 227-897, 230-878, 258-956, 260-411, 265-823, 289-636, 318-843, 341-934, 344-805, 346-741, 355-1028, 376-943, 457- 1037, 466-745, 477-1064, 485-1099, 493-1131, 501-1189, 505-1235, 512-1205, 525-1255, 539-853, 549-1121, 558-1158, 568-1215, 573-1256, 592-1096, 592-1241, 604-1195, 606-950, 606-1124, 611-1053, 622-1168, 626-1286, 633-1286, 636-832, 645-1138, 653-939, 657-1248, 675-1265, 705-899, 722-1286, 724-1287, 733-1094, 736-1286, 746-1094, 750-1094, 754-1085, 756-1286, 759-1201, 766-1286, 777-1286, 782-1286, 786- 1286, 789-1286, 802-1286, 811-1286, 823-897, 823-930, 823-947, 823- 1085, 823-1098, 823-1167, 823-1218, 823-1286, 836-1286, 837-1286, 837-1383, 844-1286, 857-1286, 858-1082, 875-1286, 882-1286, 892- 1286, 901-1094, 901-1286, 903-1287, 908-1286, 910-1287, 912-1286, 916-1286, 928-1094, 953-2377, 954-1111, 954-1258, 954-1274, 954- 1285, 954-1287, 957-1286, 963-1286, 964-1087, 968-1094, 969-1286, 1004-1286, 1008-1513, 1016-1150, 1017-1286, 1018-1100, 1030-1286, 1033-1286, 1053-1287, 1057-1286, 1076-1286, 1086-1286, 1086-1287, 1100-1287, 1123-1680, 1127-1287, 1139-1286, 1147-1286, 1151-1287, 1164-1286, 1169-1286, 1169-1287, 1175-1287, 1181-1208, 1182-1287, 1183-1286, 1191-1286, 1193-1857, 1196-1287, 1204-1287, 1206-1286, 1214-1633, 1219-1287, 1225-1286, 1225-1287, 1227-1287, 1307-1978, 1457-2058, 1500-2092, 1550-2096, 1613-7633, 1655-2091, 1729-2361, 1865-2058, 1941-1982, 2198-2740, 2296-2786, 2348-2786, 2415-2569, 2485-2788, 2513-3053, 2519-3045, 2520-3044, 2567-2652, 2568-2691, 2569-2652, 2592-3183, 2756-3428, 2764-2850, 2932-2972, 2974-3456, 3228-3519, 3400-3541, 3766-3807, 3965-4105, 4665-5170, 5927-6285, 6473-6971, 7199-7369, 7232-7267
27/1482539CBI /992	1-135, 63-135, 63-338, 63-518, 65-135, 65-583, 81-414, 100-367, 108- 532, 109-523, 109-610, 111-545, 124-463, 204-791, 225-791, 231-528, 231-571, 231-585, 231-642, 231-644, 231-647, 239-759, 279-568, 319- 759, 358-759, 484-630, 561-989, 687-992, 718-992

10

20

30

40

50

【 0 3 5 2 】
【 表 4 - 2 】

表4-2

ポリヌクレオチド SEQ ID NO/ Incyte ポリヌクレオチド ID/ 配列長	配列断片
28/2234213CB1/ 3788	1-700, 38-473, 85-497, 94-505, 106-505, 113-506, 234-501, 498-1105, 499-680, 499-683, 502-680, 502-683, 507-683, 508-700, 515-683, 523-671, 523-998, 528-1123, 529-671, 532-678, 532-683, 536-713, 547-1138, 735-981, 735-1172, 735-1242, 735-1247, 735-1260, 735-1270, 735-1363, 735-1389, 735-1405, 735-1415, 735-1463, 739-782, 746-1619, 764-872, 764-1455, 782-1255, 826-1549, 906-1604, 984-1738, 992-1846, 1030-1836, 1034-1790, 1083-1777, 1286-1940, 1377-2087, 1494-2130, 1501-2063, 1574-1808, 1579-2268, 1604-1839, 1606-1836, 1606-1848, 1606-1936, 1606-2019, 1606-2060, 1606-2111, 1606-2123, 1606-2136, 1606-2239, 1624-2188, 1626-2276, 1655-2229, 1683-2273, 1693-2436, 1721-2303, 1721-2336, 1822-2383, 1832-2066, 1889-2409, 1889-2455, 1889-2469, 1915-2395, 1918-2539, 1920-2157, 1935-2523, 2053-2577, 2143-2560, 2241-2810, 2246-2838, 2255-2868, 2293-2746, 2305-2934, 2313-2910, 2335-2942, 2385-2997, 2413-2676, 2465-3019, 2497-3076, 2503-3066, 2507-3127, 2510-3056, 2515-3096, 2566-2749, 2626-2904, 2632-3239, 2689-3309, 2697-3239, 2787-3445, 2812-3401, 2920-3502, 2923-3500, 2962-3345, 2984-3589, 3016-3652, 3034-3638, 3065-3332, 3142-3750, 3163-3679, 3167-3358, 3221-3788, 3230-3765, 3257-3788, 3265-3508
29/1345785CB1/ 2443	1-665, 2-44, 3-43, 153-418, 166-486, 166-515, 175-515, 178-516, 179-414, 180-821, 184-497, 184-515, 209-788, 244-634, 248-475, 388-943, 504-1004, 508-716, 508-722, 508-815, 508-834, 508-894, 508-906, 508-909, 508-945, 508-947, 508-956, 508-965, 508-1009, 508-1018, 508-1029, 508-1030, 510-956, 512-867, 513-1090, 513-1164, 526-829, 526-941, 526-981, 535-954, 536-772, 536-947, 536-995, 543-759, 543-1002, 555-952, 555-988, 557-1009, 604-653, 611-1262, 634-859, 699-951, 775-1044, 842-1134, 875-1085, 911-1112, 927-1154, 931-1143, 938-1162, 948-1249, 950-1418, 1012-1207, 1036-1300, 1044-1402, 1057-1328, 1065-1773, 1073-1343, 1073-1773, 1077-1450, 1088-1324, 1088-1498, 1089-1670, 1124-1366, 1150-1416, 1151-1189, 1179-1432, 1179-1461, 1179-1478, 1185-1430, 1185-1467, 1187-1447, 1189-1436, 1189-1449, 1195-1438, 1196-1426, 1197-1397, 1197-1425, 1197-1435, 1197-1461, 1197-1466, 1197-1602, 1202-1785, 1205-1486, 1221-1508, 1226-1425, 1247-1503, 1261-1472, 1267-1562, 1274-1432, 1274-1496, 1275-1535, 1281-1512, 1293-1543, 1304-1584, 1310-1619, 1324-1492, 1326-1605, 1331-1610, 1333-1562, 1333-1600, 1344-1665, 1344-1674, 1346-1571, 1346-1587, 1347-1634, 1361-1648, 1384-1603, 1387-1614, 1387-

【 0 3 5 3 】
【 表 4 - 3 】

表4-3

ホリスクレオチド SEQ ID NO/ Incyte 求	配列断片
リスクレオチド ID/ 配列長	
1646,	1387-1922,
1687,	1411-1683,
1649,	1439-1697,
1736,	1445-1693,
1731,	1449-1699,
1712,	1505-1800,
1970,	1530-1776,
1786,	1554-1926,
2031,	1665-1899,
2212,	1700-1932,
2292,	1733-2014,
2254,	1762-2056,
2090,	1782-2104,
2362,	1825-2080,
1896,	1837-1900,
1905,	1851-1912,
2008,	1869-2399,
2159,	1889-1912,
2169,	1896-2424,
2397,	1939-2021,
1976,	1947-2007,
2020,	1964-2020,
2020,	1987-2435,
2279,	2011-2433,
2306,	2048-2435,
2337,	2062-2353,
2435,	2079-2322,
2430,	2100-2292,
2438,	2119-2360,
2443,	2149-2399,
2420,	2171-2443,
2443,	2208-2443,
2443,	2251-2442,
1407-1683,	1403-1617,
1439-1634,	1423-1763,
1444-1446,	1440-1708,
1446-1498,	1445-1712,
1491-2066,	1445-1708,
1515-1515,	1485-1789,
1554-1554,	1508-1747,
1652-1680,	1531-1677,
1680-1719,	1620-1838,
1719-1761,	1673-2147,
1761-1782,	1708-1994,
1810-1837,	1745-2004,
1837-1840,	1765-2042,
1840-1863,	1798-2063,
1863-1885,	1837-1857,
1885-1896,	1837-1912,
1896-1934,	1854-1912,
1934-1947,	1873-2138,
1947-1962,	1896-2144,
1962-1986,	1910-2180,
1986-2008,	1931-2180,
2008-2046,	1946-2190,
2046-2062,	1947-2022,
2062-2078,	1951-2020,
2078-2092,	1967-2423,
2092-2111,	1984-2020,
2111-2146,	1994-2261,
2146-2171,	1997-2020,
2171-2198,	2019-2435,
2198-2249,	2029-2286,
2249-2443,	2058-2307,
2443-2443,	2061-2317,
2443-2443,	2075-2435,
2443-2443,	2077-2436,
2443-2443,	2085-2443,
2443-2443,	2083-2437,
2443-2443,	2085-2443,
2443-2443,	2091-2443,
2443-2443,	2105-2366,
2443-2443,	2106-2365,
2443-2443,	2111-2422,
2443-2443,	2131-2441,
2443-2443,	2132-2440,
2443-2443,	2141-2401,
2443-2443,	2157-2443,
2443-2443,	2159-2435,
2443-2443,	2187-2413,
2443-2443,	2188-2324,
2443-2443,	2189-2435,
2443-2443,	2228-2436,
2443-2443,	2231-2443,
2443-2443,	2235-2393,
2443-2443,	2269-2443,
2443-2443,	2273-2443,
2443-2443,	2301-2443,
2443-2443,	2302-2443

10

20

30

40

【 0 3 5 4 】

【 表 4 - 4 】

表4-4

ホリスクレオチド SEQ ID NO/ Incyte ホリスクレオチド ID/ 配列長	配列断片
30/3807190CB1/ 957	1-226, 1-288, 1-379, 1-380, 1-518, 1-537, 1-555, 1-573, 1-577, 1-582, 1-609, 1-621, 1-632, 1-691, 1-708, 1-709, 1-710, 1-776, 1-792, 2-694, 3-541, 12-697, 20-503, 164-423, 167-799, 195-769, 215-923, 268-949, 314-640, 330-585, 333-948, 350-888, 380-954, 408-957, 411-957, 439-893, 440-952, 457-957, 475-957, 480-510, 480-957, 503-957, 523-955, 534-655, 534-662, 536-662, 537-662, 539-662, 551-957, 553-662, 615-655, 615-662, 726-932, 729-957, 823-950, 825-872, 825-934, 825-950, 825-951, 831-951, 832-862, 832-951, 834-860, 834-867, 834-872, 836-865, 836-951, 844-872, 847-951, 858-935, 885-950, 887-941, 888-951, 914-950, 914-951, 915-940, 915-950
31/4856078CB1/ 674	1-345, 28-559, 84-523, 305-559, 315-568, 315-674
32/6106886CB1/ 559	1-157, 1-264, 1-379, 66-559, 162-402
33/5627037CB1/ 2600	1-104, 1-469, 31-705, 37-650, 245-520, 300-963, 355-983, 456-1000, 604-882, 620-1000, 777-936, 850-1506, 939-1067, 941-1242, 942-1067, 944-1067, 945-1067, 1066-1280, 1066-1419, 1067-1242, 1180-1831, 1345-1541, 1460-2245, 1497-2080, 1618-2263, 1634-2263, 1676-2263, 1827-2149, 1827-2570, 1829-2528, 1964-2126, 1968-2600, 1982-2485, 1984-2198, 2396-2580
34/3688835CB1/ 788	1-278, 1-526, 165-455, 166-739, 166-788, 179-775, 185-775, 187-782, 576-754
35/2295415CB1/ 702	1-274, 1-597, 1-608, 8-248, 118-639, 188-640, 261-670, 272-411, 272-678, 272-702
36/2437896CB1/ 806	1-549, 1-640, 275-806, 379-631
37/2641482CB1/ 1112	1-214, 34-286, 105-342, 109-701, 128-399, 137-382, 161-412, 296-456, 379-1086, 456-526, 456-578, 456-646, 456-970, 456-1086, 477-578, 485-714, 486-569, 520-781, 566-1090, 677-801, 691-1105, 732-1024, 766-1033, 964-1112
38/7494292CB1/ 2401	1-455, 2-455, 16-2401, 59-531, 72-591, 124-649, 124-676, 124-811, 151-647, 151-776, 324-1133, 430-1266, 513-821, 535-734, 567-1464, 666-1226, 720-1585, 721-1071, 731-1457, 753-1474, 766-1725, 785-1562, 789-1320, 802-1652, 810-1543, 822-1337, 822-1342, 823-1749, 832-1449, 850-1513, 865-1603, 1790-2059, 2104-2396, 2108-2401

【 0 3 5 5 】
【 表 4 - 5 】

表4-5

ホリスクレオチド SEQ ID NO/ Incyte ポリヌクレオチド ID/ 配列長	配列断片
39/7636288CB1/1575	1-512, 267-703, 341-950, 357-735, 379-858, 422-858, 428-735, 492-1061, 518-942, 769-1518, 769-1575
40/8095391CB1/998	1-625, 44-625, 56-330, 58-333, 92-650, 200-686, 254-549, 300-696, 341-578, 396-998
41/532018CB1/1697	1-92, 1-310, 16-256, 16-333, 18-534, 21-351, 21-366, 21-461, 25-182, 25-537, 27-264, 27-286, 27-328, 27-339, 27-341, 27-485, 37-371, 37-560, 45-441, 127-440, 293-967, 293-976, 293-1043, 295-483, 298-917, 372-862, 390-712, 475-837, 475-962, 648-832, 706-1051, 752-1359, 986-1653, 993-1564, 1037-1638, 1092-1220, 1109-1248, 1135-1681, 1137-1393, 1139-1651, 1169-1467, 1198-1679, 1227-1598, 1227-1654, 1230-1654, 1231-1654, 1232-1654, 1233-1654, 1234-1654, 1237-1650, 1243-1654, 1244-1654, 1246-1654, 1249-1485, 1249-1543, 1254-1418, 1255-1540, 1256-1503, 1256-1646, 1259-1503, 1259-1654, 1260-1654, 1261-1491, 1261-1654, 1262-1464, 1263-1624, 1266-1654, 1267-1675, 1268-1505, 1269-1644, 1275-1416, 1275-1502, 1275-1654, 1275-1688, 1278-1558, 1280-1654, 1284-1654, 1296-1654, 1298-1551, 1299-1558, 1300-1554, 1300-1654, 1303-1691, 1303-1695, 1305-1654, 1306-1676, 1307-1371, 1310-1694, 1331-1571, 1338-1654, 1343-1505, 1352-1575, 1361-1691, 1369-1603, 1369-1669, 1369-1678, 1375-1654, 1376-1685, 1379-1646, 1380-1654, 1389-1688, 1397-1570, 1401-1685, 1408-1654, 1422-1685, 1443-1691, 1481-1643, 1481-1675, 1488-1683, 1532-1692, 1577-1697, 1581-1654, 1617-1683
42/7481063CB1/592	1-91, 1-393, 31-183, 92-393, 153-506, 153-507, 153-508, 153-523, 153-529, 153-532, 153-566, 153-592
43/2044436CB1/1031	1-642, 19-259, 24-277, 24-487, 69-270, 69-379, 69-380, 70-374, 70-379, 71-171, 71-176, 189-677, 256-388, 313-483, 365-487, 533-957, 556-1012, 556-1025, 562-596, 576-1021, 580-1020, 583-1005, 588-1031, 593-1013, 593-1031, 602-877, 611-1023, 618-1031, 642-1022, 658-1031, 662-1004, 680-1023, 745-1011, 767-1021, 802-1022, 851-1006, 940-1031
44/4091564CB1/916	1-306, 1-566, 1-635, 217-907, 228-916
45/8039739CB1/2577	1-1789, 1-1817, 32-1110, 827-918, 827-1037, 827-1197, 895-1151, 994-1797, 1075-1729, 1112-1364, 1112-1365, 1148-1884, 1258-1365, 1289-1328, 1298-1365, 1383-2115, 1391-1476, 1391-1644, 1447-1644, 1539-2026, 1577-2081, 1883-2303, 1889-2158, 1890-2299, 2027-2301, 2044-2577, 2165-2533

【 0 3 5 6 】
【 表 4 - 6 】

表4-6

ホリスクレオチド SEQ ID NO/ Incyte ポ リスクレオチド ID/ 配列長	配列断片
46/1265837CB1/ 2999	1-577, 287-810, 317-548, 345-976, 573-826, 573-837, 575-835, 678-1226, 679-922, 680-900, 829-1081, 859-1093, 859-1342, 889-1168, 955-1212, 1212-1476, 1212-1697, 1283-1551, 1327-1588, 1333-1472, 1436-1706, 1466-1624, 1482-1751, 1482-2033, 1512-1714, 1536-1769, 1547-1752, 1610-1925, 1631-1882, 1637-1907, 1663-1946, 1785-2064, 1792-2061, 1817-2078, 1836-2073, 1844-2065, 1844-2332, 1857-2099, 1901-2064, 1944-2171, 1964-2197, 1990-2235, 1990-2572, 1990-2580, 2009-2273, 2083-2309, 2112-2411, 2119-2346, 2119-2663, 2127-2460, 2161-2775, 2168-2765, 2193-2418, 2243-2484, 2275-2548, 2275-2775, 2341-2615, 2350-2561, 2493-2711, 2493-2999, 2494-2921, 2575-2853, 2576-2811
47/5568527CB1/ 1314	1-42, 1-236, 1-255, 1-346, 1-394, 1-423, 1-427, 1-442, 1-444, 1-497, 4-150, 6-160, 6-339, 13-162, 124-431, 217-509, 245-568, 259-441, 259-506, 329-437, 332-593, 345-609, 351-518, 384-624, 384-883, 384-990, 393-649, 400-879, 412-630, 412-739, 427-613, 450-762, 450-1044, 470-744, 478-564, 526-762, 564-681, 564-726, 564-1027, 569-783, 569-921, 571-716, 603-896, 629-905, 640-1294, 646-866, 648-1287, 651-1286, 657-1293, 658-1303, 664-1296, 670-885, 687-1286, 747-1271, 761-1299, 764-1013, 771-1056, 775-983, 775-1296, 778-1314, 783-1076, 784-1023, 792-1057, 822-1077, 847-998, 862-1108, 897-1314, 905-1118, 905-1296, 905-1314, 944-1259, 945-1130, 945-1131, 995-1305, 999-1314, 1001-1305, 1023-1314, 1046-1314, 1073-1290, 1124-1314, 1157-1314
48/7503641CB1/ 2619	1-2348, 149-507, -53-418, 166-486, 166-515, 169-506, 175-515, 178-506, 179-414, 184-497, 184-515, 187-270, 188-479, 210-453, 248-475, 279-506, 604-1048, 653-1297, 656-948, 658-1309, 668-1363, 680-949, 727-1247, 751-1039, 751-1390, 760-1230, 780-990, 785-1247, 798-945, 802-1065, 802-1361, 802-1379, 802-1384, 806-1178, 806-1371, 816-1017, 832-1059, 836-1041, 836-1048, 843-1067, 843-1282, 843-1417, 843-1566, 850-1366, 854-1334, 855-1154, 855-1323, 889-1454, 917-1112, 941-1205, 949-1394, 962-1233, 970-1616, 978-1248, 978-1678, 982-1355, 987-1289, 991-1578, 993-1229, 994-1575, 996-1289, 997-1586, 997-1659, 1029-1271, 1041-1539, 1041-1559, 1055-1321, 1058-1784, 1079-1639, 1084-1366, 1085-1337, 1090-1335, 1090-1372, 1092-1352, 1094-1341, 1094-1354, 1100-1331, 1101-1343, 1102-1302, 1102-1340, 1102-1366, 1102-1507, 1102-1577, 1106-1330, 1106-1371, 1107-1660, 1107-1690, 1110-1391, 1110-1684, 1127-1413, 1131-1330, 1132-

【 0 3 5 7 】
【 表 4 - 7 】

表4-8

ホリヌクレオチド リヌクレオチド	SEQ ID NO/ ID/	Incyte 示 配列長	配列断片
1817,	1742-1817,	1742-1817,	1742-1817, 1742-1817, 1742-1817, 1742-
1817,	1742-1817,	1742-1817,	1742-1817, 1742-1817, 1742-1817, 1742-1817, 1742-
1817,	1742-1817,	1742-1817,	1742-1817, 1742-1817, 1742-1817, 1742-1817, 1742-
1817,	1742-1817,	1742-1817,	1742-1817, 1742-1817, 1742-1817, 1742-1817, 1742-
1819,	1742-2059,	1743-1813,	1743-1976, 1743-2010, 1745-1810, 1745-
1817,	1745-1817,	1746-1817,	1751-2356, 1753-2355, 1755-1817, 1755-
1817,	1756-1817,	1757-1986,	1758-2354, 1759-1817, 1759-1996, 1759-
2013,	1762-2117,	1766-2267,	1766-2274, 1766-2328, 1770-1913, 1774-
2304,	1774-2316,	1778-2043,	1779-1817, 1779-1817, 1779-
1817,	1779-1817,	1780-1817,	1783-2064, 1784-2291, 1789-2101, 1789-
2276,	1790-2064,	1791-2353,	1794-1817, 1794-1817, 1794-1817, 1794-
1817,	1801-2072,	1801-2074,	1801-2329, 1807-2120, 1812-2303, 1815-
2085,	1829-2362,	1836-2085,	1837-2113, 1837-2366, 1843-
2268,	1844-1926,	1845-2111,	1845-2228, 1849-2399, 1851-2095, 1852-
1876,	1852-1878,	1852-1881,	1852-1881, 1517 1712, 1525-1743, 1529-
1559,	1533-2261,	1535-1763,	1536-1804, 1558-1936, 1564-2080, 1570-
1804,	1578-2052,	1585-1766,	1585-1781, 1585-2117, 1595-2296, 1606-
1837,	1611-1872,	1613-1899,	1624-1838, 1624-2197, 1624-2283, 1624-
2364,	1624-2376,	1630-2149,	1630-2242, 1632-2282, 1638-2052, 1640-
2244,	1647-2264,	1650-2340,	1651-1909, 1657-2354, 1660-1904, 1660-
2176,	1661-2317,	1666-2159,	1666-2159, 1852-1921, 1852-1925, 1852-1931, 1852-
2273,	1852-2339,	1853-1925,	1855-1921, 1855-1925, 1855-2161, 1855-
2328,	1856-1925,	1861-1925,	1865-2344, 1865-2353, 1866-2340, 1867-
1925,	1868-1925,	1869-1925,	1869-2100, 1872-2206, 1872-2328, 1875-
1981,	1883-2333,	1883-2355,	1889-1921, 1889-1925, 1890-1925, 1890-
2363,	1891-1925,	1892-2340,	1894-1925, 1897-2147, 1898-2338, 1898-
2360,	1899-2166,	1899-2177,	1899-2385, 1901-2149, 1902-1925, 1910-
2369,	1913-2184,	1916-2338,	1916-2340, 1918-2347, 1923-2338, 1924-
2340,	1925-2324,	1927-2341,	1931-2374, 1933-2340, 1933-2416, 1934-
2191,	1934-2346,	1935-2254,	1937-2245, 1943-2351, 1948-
2358,	1951-2211,	1953-2340,	1961-2204, 1963-2211, 1966-2222, 1967-
2242,	1967-2258,	1967-2340,	1978-2228, 1979-2340, 1982-2341, 1983-
2340,	1984-2227,	1989-2341,	1990-2338, 1994-2343, 1996-
2338,	1997-2338,	1998-2355,	2005-2091, 2005-2197, 2010-2271, 2011-
2266,	2011-2270,	2016-2343,	2024-2265, 2034-2340, 2036-2346, 2037-
2345,	2037-2459,	2037-2619,	2048-2304, 2051-2367, 2054-2304, 2054-
2328,	2054-2339,	2055-2322,	2060-2281, 2062-2356, 2064-2334, 2064-

10

20

30

40

【 0 3 5 9 】

【 表 4 - 9 】

表4-9

ポリヌクレオチド リヌクレオチド	SEQ ID NO/ ID/	Incyte 本 配列長	配列断片
49/7503458CB1/ 1370			2339, 2076-2325, 2076-2350, 2081-2334, 2085-2340, 2093-2318, 2093-2355, 2094-2340, 2095-2386, 2096-2227, 2103-2366, 2113-2369, 2115-2300, 2122-2334, 2129-2373, 2133-2390, 2136-2331, 2136-2363, 2149-2341, 2154-2355, 2156-2383, 2179-2394, 2192-2232, 2197-2383, 2206-2423, 2209-2324, 2223-2361, 2232-2386, 2238-2354, 2239-2357, 2240-2334, 2247-2373, 2252-2406
			1-1097, 1-1106, 19-463, 24-130, 361-480, 388-558, 406-561, 504-1097, 608-739, 631-740, 631-743, 631-1087, 631-1100, 651-1096, 655-1095, 656-1056, 658-1080, 663-1108, 663-1116, 668-1088, 668-1115, 677-952, 686-1098, 693-1116, 704-799, 711-877, 730-997, 733-1116, 737-1114, 755-1098, 758-1111, 842-1096, 847-1091, 877-1097, 936-1081, 946-1107, 1015-1370
50/7500925CB1/ 1201			1-537, 11-467, 11-475, 22-190, 22-285, 22-295, 22-299, 22-400, 22-458, 22-468, 22-475, 22-478, 22-484, 23-386, 23-482, 23-484, 26-379, 27-276, 27-434, 38-468, 40-684, 41-1112, 69-355, 140-651, 164-471, 165-744, 257-549, 280-469, 299-481, 299-546, 372-633, 385-649, 389-693, 391-558, 424-664, 433-689, 453-670, 453-772, 453-786, 467-653, 490-626, 490-778, 510-782, 610-692, 611-756, 734-1020, 734-1065, 801-1111, 807-1111, 830-1111, 833-1201, 842-1113, 990-1111

10

20

30

40

表 5

ポリヌクレオチド SEQ ID NO:	Incyte プロジェクトID	代表的ライブラリ
26	7757335CB1	LIVRTUE01
27	1482539CB1	COLNNO05
28	2234213CB1	PANCTUT02
29	1345785CB1	LATRTUT02
30	3807190CB1	CONTTUT01
31	4856078CB1	BRSTTUT22
32	6106886CB1	MCLRNOC01
33	5627037CB1	BONEUNR01
34	3688835CB1	THYRNOT10
35	2295419CB1	KIDNNOT32
36	2437896CB1	BRAFTDT02
37	2641482CB1	LUNGTUT08
38	7494292CB1	BRAIFER06
39	7636288CB1	SINTDIE01
40	8095391CB1	BRAINYO02
41	532018CB1	LIVRTUT04
43	2044436CB1	PTHYNOT03
44	4091564CB1	BSCNSZT01
45	8039739CB1	SPLNNOE01
46	1265837CB1	ENDANOT01
47	5568527CB1	BRAITUT01
48	7503641CB1	FIBRUNT02
49	7503458CB1	BRANDIT04
50	7500925CB1	MONOTXN05

10

20

30

40

【 0 3 6 1 】

【 表 6 - 1 】

表6-1

ライブラリ	ベクター	ライブラリの説明
BONEUNR01	PCDNA2.1	このランダムにプライムされたライブラリは、2名のドナーからプールしたcDNAを使って作製された。cDNAは、14才白人男子(ドナーA)から切除した骨肉腫由来の未処置MG-63細胞株から、および診断的開腹および軟部組織切除時に18才白人女性(ドナーB)から摘出した仙骨腫瘍組織から単離されたmRNAを使って作製された。病理検査は、ドナーBに仙骨の巨細胞腫を示した。ドナーBの病歴には、骨盤関節痛、便秘、尿失禁、不特定の腹部/骨盤症状および骨盤軟部組織悪性新生物がある。家族歴にはドナーBに前立腺癌がある。
BRAFTDT02	pINCY	ライブラリは、77才男子から採取した前頭皮質組織から単離したRNAを用いて作製した。患者の病歴には、肺癌がある。
BRAIFER06	PCDNA2.1	このランダムにプライムされたライブラリは、妊娠23週間で死産の左心低形成症胎児(白人男子)から除去された脳組織から単離したRNAを用いて作製した。血清検査は陰性だった。
BRAIN0Y02	pINCY	この大型分画されノーマライズされたライブラリは、心不全で死亡した35歳の白人男性から採取した、中脳、下側頭皮質、髄質、および後頭頂皮質の組織から単離したmRNAを用いて産生した、プールしたcDNAを用いて構築された。病理検査では、中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、中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髄体ニューロンの形跡があった。大脳皮質全体に散発して、周囲にグリオシスを伴う、空洞化の複数の小さな微視的領域があった。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。このサイズ選択されたライブラリからの28万の独立クローンは、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) 91:9228-9232 および Bonaldo 他, Genome Research 6 (1996):791 を適用した条件を用いて2回、ノーマライズした。
BRAIN0Y02 cont.		
BRAITUT01	PSPORT1	ライブラリは50才の白人女性の前頭葉切除時に除去された脳腫瘍組織から単離したRNAを用いて作製した。病理は再発性グレード3の乏突起星細胞腫(オリゴアストロサイトーマ)を示し、限局性壊死と広汎にわたる石灰化があった。患者の病歴には言語障害と癲癇が含まれる。患者の脳はまた全投与量、5,082 cyg[分画8]の照射を受けた。家族歴には脳腫瘍がある。
BRANDIT04	pINCY	ライブラリはうつ血性心不全で死亡した68才の白人女性から採取した松果体組織から単離したRNAを用いて作製した。神経病理は、軽度から中等度のアルツハイマー病、アテローム性動脈硬化症および多発性梗塞を示した。顕微鏡的には、大脳皮質全体にわたってびまん性かつ神経突起状アミロイド斑があった。側頭葉特に嗅内皮質に神経原線維変化があった。前頭葉には散在性、膨張した神経細胞が含まれていた。扁桃体には著しい神経膠症、神経突起状斑と細胞内神経原線維変化が見られた。海馬には、神経突起および散在性斑、また神経原線維変化が見られた。視床にはびまん性かつ限局性神経突起状アミロイド斑および散在性神経原線維変化が見られた。左淡蒼球にグリオシスで囲まれた囊胞性空洞化部分があった。淡蒼球

【 0 3 6 2 】
【 表 6 - 2 】

表6-2

ライブラリ	ベクター	ライブラリの説明
BRANDIT04 cont.		ライブラリには散在性細胞内神経原線維変化が含まれていた。尾状核、被殻および中隔側坐核にはびまん性斑が含まれていた。右小脳半球に脂質含有マクロファージを有する囊胞性空洞化部分があった。患者の病歴には、真性糖尿病、慢性関節リウマチ、甲状腺機能亢進、アミロイド心疾患および痲果が含まれる。
BRSTTUT22	pINCY	ライブラリは59才の白人女性の片側性拡大単純乳房切除術時に採取した右乳房腫瘍組織から単離したRNAを用いて作製した。病理は浸潤性小葉癌を示し、右乳房生検空洞部位に位置する不明確な塊を形成して、管にまで及んでいた。非腫瘍性乳房実質は乳頭腫症を示す。前の右乳房生検は浸潤度3、核グレード3の浸潤性乳管癌かつ非浸潤性乳管癌を示した。腫瘍性細胞におけるエストロゲンおよびプロゲステロンの免疫染色は陽性であった。患者の病歴には肝硬変、食道潰瘍、高脂血症および不特定疾患によって生じた神経障害がある。前に行った手術には分節性肺切除および肝臓移植がある。患者の服用薬剤には Prograf、プレドニゾロン、イムラン、Lozol、ザンタック、Estraderm パッチおよび Provera が含まれる。
BSCNSZT01	pINCY	ライブラリは49才の男性の脳から採取した罹患尾状核組織から単離したRNAを用いて作製した。患者の病歴には精神分裂症が含まれる。
COLNON05	pINCY	この標準化された結腸および結腸上皮層組織のライブラリは、結腸および結腸上皮層組織のライブラリからの7百万の独立型クローンから作製した。開始RNAは2人のドナーのプールのcDNAから作成した。cDNAは自動車事故で死亡した16才の黒人男子(ドナーA)から採取した結腸組織と、自動車事故で死亡した13才の白人女性(ドナーB)から採取した結腸上皮組織から単離されたmRNAを用いて作製した。ドナーAの場合、血清学はサイトメガロウィルスに陽性であり、残りの血清学検査項目は陰性であった。患者の病歴には時折、アルコール(ビール)の飲用が含まれる。患者は自動車事故後胸部チューブ挿入で治療された。この患者が服用していた薬剤は無い。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 および Ronaldo 他, Genome Research (1996) 6:791 を適用した条件を用いて2回にわたり標準化した。
CONTTUT01	pINCY	ライブラリは、34才白人女性の軟組織切除術時に左外側大腿の腫瘍性軟部組織から単離されたRNAを用いて作製した。病理は転移度2の粘液様脂肪肉腫を示した。この粘液様脂肪肉腫は、皮下脂肪組織に存在する複数に分化した、限局性塊を形成した。患者の病歴には、肢の悪性軟組織腫瘍が含まれていた。家族歴には良性高血圧、急性白血病、およびII型糖尿病がある。
ENDANOT01	PBLUESCRIPT	ライブラリは、心臓移植時に或る男性から切除した外植心臓からの大動脈内皮細胞組織から単離されたRNAを用いて作製された。
FIBRUNT02	pINCY	ライブラリは14才の白人男子から摘出した骨肉腫由来する未処理のMC-63細胞系から単離したRNAを用いて作製した。

表6-3

ライブラリ	ベクター	ライブラリの説明
KIDNNOT32	pINCY	ライブラリは頭蓋内出血および脳血管事故で死亡した49才の白人男子から抽出した腎組織から単離したRNAを用いて作製した。患者の病歴にはタバコ濫用がある。
LATRITUT02	pINCY	ライブラリは43才白人男性の弁輪形成時に左心房から採取した粘液腫から単離したRNAを用いて作成した。病理は心房性粘液腫を示した。患者の病歴には肺不全、急性心筋梗塞、アテローム性冠動脈疾患、高脂血症および喫煙がある。家族歴には良性高血圧症、急性心筋梗塞、アテローム性冠動脈疾患およびII型糖尿病がある。
LIVRTUE01	PCDNA2.1	この5'パイアスのランダムプライムされたライブラリは72歳白人男性の部分肝切除時に摘出された肝腫瘍組織から単離したRNAを用いて作製した。病理は転移度2(4の中)の神経内分泌癌を示し、塊を形成していた。患者は転移性肝臓癌を示した。患者の病歴には良性高血圧、I型糖尿病、前立腺肥大、前立腺癌、アルコールおよびタバコ濫用(共に寛解期)がある。前に行われた手術には膀胱癌変部除去、非観血的前立腺生検(closed prostatic biopsy)、経尿道的前立腺切除、両側の精巣摘出および脾臓全摘が含まれる。患者の服用薬剤には、Eulexin、Hytrin、Proscar、エコトリンおよびインスリンがある。家族歴には母親にアテローム性冠動脈疾患と急性心筋梗塞、父親にアテローム性冠動脈疾患とII型糖尿病がある。
LIVRTUT04	pINCY	ライブラリは50才の白人男性の部分肝切除時に採取した肝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3-4の肝癌で塊を形成を示した。患者の病歴には良性高血圧と肝炎が含まれる。B型肝炎コア抗原およびB型肝炎表面抗原がこの患者に存在した。
LUNGTUT08	pINCY	ライブラリは、右上肺葉切除(光ファイバー気管支鏡検査を併行)時に63才白人男性から除去した肺腫瘍組織から単離されたRNAを用いて作製された。病理はグレード3の腺癌を示した。家族歴には、アテローム性冠動脈疾患、急性心筋梗塞、結腸癌、無症候性腹部大動脈瘤、タバコ乱用、および不整脈がある。家族歴には、鬱血性心不全、胃癌、肺癌、II型糖尿病、アテローム硬化型冠動脈疾患、および急性心筋梗塞がある。
MCLRNOC01	pINCY	この大きいサイズに分類されたライブラリは、年令の異なる、複数の男女の臍帯コードから得た単核細胞から単離されたRNAを用いて作製した。細胞はG-CSFで処理した。
MONOTXN05	pINCY	このノーマライズした処理済み単核細胞組織ライブラリは、或る単核細胞ライブラリからの103万個の独立クローンから作製した。開始RNAは、42才の女性から採取した末梢血からの処理済みの単核細胞から単離したRNAから作製した。細胞を、インターロイキン-10(IL-10)とリポ多糖(LPS)で処理した。このライブラリは、極めて長時間(48時間/一回)の再アニーリングによるハイブリダイゼーションを用いた他は、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research 6 (1996):791 を適用した条件を用いて2回にわたってノーマライズした。
PANCTUT02	pINCY	ライブラリは、根治的膵十二指腸切除時に45歳の白人女性から採取した膵臓腫瘍組織から単離したRNAを使用して作製された。病理学検査では、グレード4の未分化癌が見られた。家族の病歴には、良性高血圧、高脂血症およびアテローム硬

表6-4

ライブラリ	ベクター	ライブラリの説明
PTHYNOT03	pINCY	ライブラリは、69才白人女性の副甲状腺の部分切除時に左副甲状腺組織から単離された RNA を用いて作製された。病理学検査では過形成を示した。患者は、原発性副甲状腺亢進症を示した。
SINTDIE01	PCDNA2.1	この、5'に偏向してランダムプライムしたライブラリは、49才の白人女性の、胃腸吻合、診断的開腹、および迷走神経切除時に摘出された小腸組織から単離した RNA を用いて作製した。患者は、閉塞を伴う急性胃潰瘍、悪心および嘔吐、並びに異常な体重減少を示した。患者の病歴としては、腰痛、穿孔を伴う急性胃潰瘍、および正常な出産があった。過去の手術には、扁桃腺アデノイド切除および腹式子宮全摘出がある。患者の使用薬剤にはプレマリン(Premarin)がある。家族歴には、良性高血圧、II型糖尿病、および鬱血性心不全が父親にある。
SPLNNOE01	PCDNA2.1	この5'に偏向してランダムプライムされたライブラリは、脳無酸素症で死亡した2才ヒスパニック系男子の脾臓組織から単離された RNA を用いて作製された。過去の病歴は無く、血清検査は陰生であった。
THYRN0T10	pINCY	ライブラリは 30才白人女性の片側性甲状腺葉切除、および副甲状腺再移植時に摘出された病変左甲状腺組織から単離した RNA を用いて作成した。病理検査ではリンパ球性甲状腺炎が示された。

表7-1

プログラム	説明	参照文献	パラメータ閾値
ABI FACTURA	核酸配列においてベクター配列を除去して、あいまいな塩基をマスクするプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finder。アミノ酸配列または核酸配列の比較および注釈付けに有用である。	Applied Biosystems, Foster City, CA: Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler	核酸配列を構築するプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	Basic Local Alignment Search Tool。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る。BLASTにはblastp、blastn、blastx、tblastnおよびtblastxの5つの機能がある。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: 確率値=1.0E-8 以下 完全長配列: 確率値1.0E-10 以下
FASTA	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には最少5つの機能(fasta、tfasta、fastx、tfastxおよびsssearch)がある。	Pearson, W. R. and D. J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W. R. (1990) Methods Enzymol. 183:63-98; Smith, T. F. 及び M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E 値=1.06E-6 構築されたESTs: fasta 同一性=95%以上 一致した長さ=200塩基以上 fastx E 値=1.0E-8 以下 完全長配列: fastx スコア=100以上
BLIMPS	配列をBLOCKS、PRINTS、DOMO、PRODOMおよびPFAM データベースの配列と対応させて遺伝子ファミリー、配列相同性および構造的フィンガープリント領域を検索するBLocks IMProved Searcher。	Henikoff, S. 及び J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. 及び S. Henikoff (1996) Methods Enzymol. 266:88-105; 及び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頁	PFAM ヒット: 確率値=1.0E-3 以下 シグナルペプチドヒット: スコア=0 以上

【 0 3 6 6 】
 【 表 7 - 2 】

表7-2

プログラム	説明	参照文献	パラメータ関値
ProfileScan	Prositeで定義された配列パターンと一致するタンパク質配列内の構造的モチーフおよび配列モチーフを検索するアルゴリズム。	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スコアとその特定のPrositeモチーフに対するGCG指定「HIGH」値通常、スコア=1.4-2.1.
Phred	高い感度と確率で自動配列決定機トレースを調べベースコーティングアルゴリズム。	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なインプリメンテーションに基づくプログラムであるSWATやCrossMatchを含むPhils Revised Assembly プログラムで、配列相同性の検索やDNA配列の構築に有用である。	Smith, T. F. 及び M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T. F. 及び M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=56以上
Conscd	Phrap/センブリの表示および編集用グラフィカルツール。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキャンして、分泌シグナルペプチドの存在を調べる重み行列解析プログラム。	Nielson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J.M. 及び S. Audic (1997) CABIOS 12:431-439.	スコア=3.5以上
TMAP	重み行列を用いて蛋白配列で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. 及び 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, eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, 175-182ページ.	
Motifs	Prosite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

【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
10 October 2002 (10.10.2002)

PCT

(10) International Publication Number
WO 02/079441 A2

(51) International Patent Classification: C12N

(21) International Application Number: PCT/US02/09820

(22) International Filing Date: 29 March 2002 (29.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/280,527	30 March 2001 (30.03.2001)	US
60/282,112	6 April 2001 (06.04.2001)	US
60/282,702	9 April 2001 (09.04.2001)	US
60/283,855	13 April 2001 (13.04.2001)	US
60/343,718	19 October 2001 (19.10.2001)	US
60/339,236	7 December 2001 (07.12.2001)	US
60/357,002	13 February 2002 (13.02.2002)	US

Drive, San Jose, CA 95123 (US). **MASON, Patricia, M.** [US/US]; 360 Clarke Lane, Morgan Hill, CA 95037 (US). **SANJANWALA, Madhu, M.** [US/US]; 210 Silvia Court, Los Altos, CA 94024 (US). **SWARNAKAR, Anita** [CA/US]; 8 Locksley Avenue #5D, San Francisco, CA 94122 (US). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **THANGAVELU, Kavitha** [IN/US]; 1950 Montecito Avenue #23, Mountain View, CA 94043 (US). **TRAN, Uyen, K.** [US/US]; 2638 Mabury Square, San Jose, CA 95133 (US). **WALLA, Narinder, K.** [US/US]; 890 Davis Street A # 205, San Leandro, CA 94577 (US). **WARREN, Bridget, A.** [US/US]; 2250 Homestead Court # 2, Los Altos, CA 94024 (US). **VAO, Monique, G.** [US/US]; 1189 Woodgate Drive, Carmel, IN 46033 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US).

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

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

(72) Inventors: and
(75) Inventors/Applicants (for US only): **BAUGHN, Mariah, R.** [US/US]; 1424H Santiago Road, San Leandro, CA 94577 (US). **BURFORD, Neil** [GB/US]; 105 Willowood Circle, Durham, CT 06422 (US). **DING, Li** [CN/US]; 3353 Alma Street # 146, Palo Alto, CA 94306 (US). **DUGGAN, Brendan, M.** [AU/US]; 243 Buena Vista Avenue # 306, Sunnyvale, CA 94086 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **FORSYTHE, Ian, J.** [US/US]; 308 Roble Avenue, Redwood City, CA 94061 (US). **GANDHI, Ameena, R.** [US/US]; 705 5th Avenue, San Francisco, CA 94118 (US). **GIETZEN, Kimberly, J.** [US/US]; 691 Los Iluecos Drive, San Jose, CA 95123 (US). **GRIFFIN, Jennifer, A.** [US/US]; 33691 Mello Way, Fremont, CA 94555 (US). **HE, Ann** [CN/US]; 4601 Catalina Drive, San Jose, CA 95129 (US). **HONCHELL, Cynthia, D.** [US/US]; 400 Laurel Street #303, San Carlos, CA 94070 (US). **ISON, Craig, H.** [US/US]; 1242 Weathersfield Way, San Jose, CA 95118 (US). **LAL, Preeti, G.** [US/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **LEE, Ernestine, A.** [US/US]; 624 Kains Street, Albany, CA 94706 (US). **LEE, Sally** [US/US]; 825 East Evelyn #425, Sunnyvale, CA 94086 (US). **LI, Dyung Aina, M.** [US/US]; 233 Coy

(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, GL, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LL, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SI, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

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

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

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

(54) Title: 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 02/079441 A2

WO 02/079441

PCT/US02/09820

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 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 comprised 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, vasomediators, cell surface markers, and antigen recognition molecules. (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 "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs

WO 02/079441

PCT/US02/09820

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 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.)

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 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 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, 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 (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.

WO 02/079441

PCT/US02/09820

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
5 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
10 family with 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. 76:41-50). 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
15 (Yokoyama, M. et al. (1996) DNA Res. 3:311-320). Neuron-specific olfactomedin-related glycoproteins isolated from rat brain 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), a secreted glycoprotein isolated from
20 both the human breast carcinoma cell line SK-BR-3, and human breast milk. It 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
25 al. (1993) J. Biol. Chem. 268:14245-14249). 90K is elevated in the serum of 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).

Semaphorins are a large group of axonal guidance molecules consisting of at least 30 different
30 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

WO 02/079441

PCT/US02/09820

neuropilins consists of three different domains: CUB, discoidin, and MAM domains. The CUB and the MAM motifs of neuropilin have been suggested to have roles in protein-protein interactions and are thought 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
5 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
10 and binding partners for semaphorins (Winberg, M.L. et al (1998) *Cell* 95:903-916).

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, J.Y. (1991) *Endocrinology* 129:950-958) Different subpopulations of PSG have been
15 found to be produced by the trophoblasts of the human placenta, and the amniotic and chorionic membranes (Plouzek, C.A. et al. (1993) *Placenta* 14:277-285).

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
20 progression in thymoma (Ohta Y. et al. (2000) *Int. J. Oncol.* 17:259-264). AMFR is a cell surface glycoprotein of molecular weight 78KDa.

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
25 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 function by binding to cell surface receptors that transduce signals across
30 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.

WO 02/079441

PCT/US02/09820

(1995) Molecular Cell Biology, Scientific American Books Inc., New York, NY, pp. 856-864.)

Pro-opiomelanocortin (POMC) is the precursor polypeptide 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 polypeptide of the hormone beta-lipotropin (beta-LPH). Each
5 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 results in an endocrine disorder characterized by early-
10 onset obesity, adrenal insufficiency, and red hair pigmentation (Chretien, M. et al. (1979) *Can. J. Biochem.* 57:1111-1121; Krude, H. et al. (1998) *Nat. Genet.* 19:155-157; *Online Mendelian Inheritance in Man (OMIM)* 176830).

Growth and differentiation factors are secreted proteins which function in intercellular communication. Some factors require oligomerization or association with membrane proteins for
15 activity. Complex interactions among these factors and their receptors trigger intracellular signal 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
20 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,
25 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
30 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.

WO 02/079441

PCT/US02/09820

(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. ((1998) Brain Res. Mol. Brain Res. 62:175-186) have identified mammalian homologues of the slit gene
5 (human Slit-1, Slit-2, Slit-3 and rat Slit-1). The encoded proteins are putative secreted proteins containing EGF-like motifs and leucine-rich repeats, both of which 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., supra). The Slit family of proteins are indicated to be functional ligands of glypican-1 in nervous tissue and it is suggested that their interactions may be critical in certain stages
10 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
15 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. (Reviewed in Martin, C.R. et al. (1985) Endocrine Physiology, Oxford University Press,
20 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
25 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. 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
30 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).

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

WO 02/079441

PCT/US02/09820

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).

Other proteins that contain signal peptides include secreted proteins with enzymatic activity.

- 5 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. 275:26458-26466). 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 (Toh, H. (1991) Protein Seq. Data Anal. 4:111-117; and Iwai, N. et al. (1994) Hypertension 23:375-380).

- 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). 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

WO 02/079441

PCT/US02/09820

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).

Most normal eukaryotic cells, after a certain number of divisions, enter a state of senescence in which cells remain viable and metabolically active but no longer replicate. A number of phenotypic changes such as increased cell size and pH-dependent beta-galactosidase activity, and molecular changes such as the upregulation of particular genes, occur in senescent cells (Shelton (1999) Current Biology 9:939-945). When senescent cells are exposed to mitogens, a number of genes are upregulated, but the cells do not proliferate. Evidence indicates that senescent cells accumulate with age *in vivo*, contributing to the aging of an organism. In addition, senescence suppresses tumorigenesis, and many genes necessary for senescence also function as tumor suppressor genes, such as p53 and the retinoblastoma susceptibility gene. Most tumors contain cells that have surpassed their replicative limit, i.e. they are immortalized. Many oncogenes immortalize cells as a first step toward tumor formation.

A variety of challenges, such as oxidative stress, radiation, activated oncoproteins, and cell cycle inhibitors, induce a senescent phenotype, indicating that senescence is influenced by a number of proliferative and anti-proliferative signals (Shelton *supra*). Senescence is correlated with the progressive shortening of telomeres that occurs with each cell division. Expression of the catalytic component of telomerase in cells prevents telomere shortening and immortalizes cells such as fibroblasts and epithelial cells, but not other types of cells, such as CD8+ T cells (Migliaccio *et al.* (2000) J Immunol 165:4978-4984). Thus, senescence is controlled by telomere shortening as well as other mechanisms depending on the type of cell.

A number of genes that are differentially expressed between senescent and presenescent cells have been identified as part of ongoing studies to understand the role of senescence in aging and tumorigenesis. Most senescent cells are growth arrested in the G1 stage of the cell cycle. While expression of many cell cycle genes is similar in senescent and presenescent cells (Cristofalo (1992) Ann N Y Acad Sci 663:187-194), expression of others genes such as cyclin-dependent kinases p21 and p16, which inhibit proliferation, and cyclins D1 and E is elevated in senescent cells. Other genes that are not directly involved in the cell cycle are also upregulated such as extracellular matrix proteins fibronectin, procollagen, and osteonectin; and proteases such as collagenase, stromelysin, and

WO 02/079441

PCT/US02/09820

cathepsin B (Chen (2000) *Ann NY Acad Sci* 908:111-125). Genes underexpressed in senescent cells include those that encode heat shock proteins, *c-fos*, and *cdc-2* (Chen *supra*).

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
5 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. USA* 98:1370-1375). The process of post-translational modification of glutamic residues to form Gla is Vitamin K-
10 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).

15

Immunoglobulins

Antigen recognition molecules are key players in the sophisticated and complex immune systems which all vertebrates have developed to provide protection from viral, bacterial, fungal, and parasitic infections. A key feature of the immune system is its ability to distinguish foreign molecules,
20 or antigens, from "self" molecules. This ability is mediated primarily by secreted and transmembrane proteins expressed by leukocytes (white blood cells) such as lymphocytes, granulocytes, and monocytes. Most of these proteins belong to the immunoglobulin (Ig) superfamily, members of which contain one or more repeats of a conserved structural domain. This Ig domain is comprised of antiparallel β sheets joined by a disulfide bond in an arrangement called the Ig fold. The criteria for a
25 protein to be a member of the Ig superfamily is to have one or more Ig domains, which are regions of 70-110 amino acid residues in length homologous to either Ig variable-like (V) or Ig constant-like (C) domains. Members of the Ig superfamily include antibodies (Ab), T cell receptors (TCRs), class I and II major histocompatibility (MHC) proteins and immune cell-specific surface markers such as the "cluster of differentiation" or CD antigens, CD2, CD3, CD4, CD8, poly-Ig receptors, Fc receptors,
30 neural cell-adhesion molecule (NCAM) and platelet-derived growth factor receptor (PDGFR).

Ig domains (V and C) are regions of conserved amino acid residues that give a polypeptide a globular tertiary structure called an immunoglobulin (or antibody) fold, which consists of two approximately parallel layers of β -sheets. Conserved cysteine residues form an intrachain disulfide-

WO 02/079441

PCT/US02/09820

bonded loop, 55-75 amino acid residues in length, which connects the two layers of β -sheets. Each β -sheet has three or four anti-parallel β -strands of 5-10 amino acid residues. Hydrophobic and hydrophilic interactions of amino acid residues within the β -strands stabilize the Ig fold (hydrophobic on inward facing amino acid residues and hydrophilic on the amino acid residues in the outward facing portion of the strands). A V domain consists of a longer polypeptide than a C domain, with an additional pair of β -strands in the Ig fold.

A consistent feature of Ig superfamily genes is that each sequence of an Ig domain is encoded by a single exon. It is possible that the superfamily evolved from a gene coding for a single Ig domain involved in mediating cell-cell interactions. New members of the superfamily then arose by exon and gene duplications. Modern Ig superfamily proteins contain different numbers of V and/or C domains. Another evolutionary feature of this superfamily is the ability to undergo DNA rearrangements, a unique feature retained by the antigen receptor members of the family.

Many members of the Ig superfamily are integral plasma membrane proteins with extracellular Ig domains. The hydrophobic amino acid residues of their transmembrane domains and their cytoplasmic tails are very diverse, with little or no homology among Ig family members or to known signal-transducing structures. There are exceptions to this general superfamily description. For example, the cytoplasmic tail of PDGFR has tyrosine kinase activity. In addition Thy-1 is a glycoprotein found on thymocytes and T cells. This protein has no cytoplasmic tail, but is instead attached to the plasma membrane by a covalent glycosphatidylinositol linkage.

Another common feature of many Ig superfamily proteins is the interactions between Ig domains which are essential for the function of these molecules. Interactions between Ig domains of a multimeric protein can be either homophilic or heterophilic (i.e., between the same or different Ig domains). Antibodies are multimeric proteins which have both homophilic and heterophilic interactions between Ig domains. Pairing of constant regions of heavy chains forms the Fc region of an antibody and pairing of variable regions of light and heavy chains form the antigen binding site of an antibody. Heterophilic interactions also occur between Ig domains of different molecules. These interactions provide adhesion between cells for significant cell-cell interactions in the immune system and in the developing and mature nervous system. (Reviewed in Abbas, A.K. et al. (1991) Cellular and Molecular Immunology, W.B. Saunders Company, Philadelphia, PA, pp.142-145.)

Antibodies

MHC proteins are cell surface markers that bind to and present foreign antigens to T cells. MHC molecules are classified as either class I or class II. Class I MHC molecules (MHC I) are expressed on the surface of almost all cells and are involved in the presentation of antigen to cytotoxic

WO 02/079441

PCT/US02/09820

T cells. For example, a cell infected with virus will degrade intracellular viral proteins and express the protein fragments bound to MHC I molecules on the cell surface. The MHC I/antigen complex is recognized by cytotoxic T-cells which destroy the infected cell and the virus within. Class II MHC molecules are expressed primarily on specialized antigen-presenting cells of the immune system, such as B-cells and macrophages. These cells ingest foreign proteins from the extracellular fluid and express MHC II/antigen complex on the cell surface. This complex activates helper T-cells, which then secrete cytokines and other factors that stimulate the immune response. MHC molecules also play an important role in organ rejection following transplantation. Rejection occurs when the recipient's T-cells respond to foreign MHC molecules on the transplanted organ in the same way as to self MHC molecules bound to foreign antigen. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of the Cell, Garland Publishing, New York, NY, pp. 1229-1246.)

Antibodies are multimeric members of the Ig superfamily which are either expressed on the surface of B-cells or secreted by B-cells into the circulation. Antibodies bind and neutralize foreign antigens in the blood and other extracellular fluids. The prototypical antibody is a tetramer consisting of two identical heavy polypeptide chains (H-chains) and two identical light polypeptide chains (L-chains) interlinked by disulfide bonds. This arrangement confers the characteristic Y-shape to antibody molecules. Antibodies are classified based on their H-chain composition. The five antibody classes, IgA, IgD, IgE, IgG and IgM, are defined by the α , δ , ϵ , γ , and μ H-chain types. There are two types of L-chains, κ and λ , either of which may associate as a pair with any H-chain pair. IgG, the most common class of antibody found in the circulation, is tetrameric, while the other classes of antibodies are generally variants or multimers of this basic structure.

H-chains and L-chains each contain an N-terminal variable region and a C-terminal constant region. The constant region consists of about 110 amino acids in L-chains and about 330 or 440 amino acids in H-chains. The amino acid sequence of the constant region is nearly identical among H- or L-chains of a particular class. The variable region consists of about 110 amino acids in both H- and L-chains. However, the amino acid sequence of the variable region differs among H- or L-chains of a particular class. Within each H- or L-chain variable region are three hypervariable regions of extensive sequence diversity, each consisting of about 5 to 10 amino acids. In the antibody molecule, the H- and L-chain hypervariable regions come together to form the antigen recognition site. (Reviewed in Alberts, B. et al. supra, pp. 1206-1213 and 1216-1217.)

Both H-chains and L-chains contain the repeated Ig domains of members of the Ig superfamily. For example, a typical H-chain contains four Ig domains, three of which occur within the constant region and one of which occurs within the variable region and contributes to the formation of

WO 02/079441

PCT/US02/09820

the antigen recognition site. Likewise, a typical L-chain contains two Ig domains, one of which occurs within the constant region and one of which occurs within the variable region.

The immune system is capable of recognizing and responding to any foreign molecule that enters the body. Therefore, the immune system must be armed with a full repertoire of antibodies
5 against all potential antigens. Such antibody diversity is generated by somatic rearrangement of gene segments encoding variable and constant regions. These gene segments are joined together by site-specific recombination which occurs between highly conserved DNA sequences that flank each gene segment. Because there are hundreds of different gene segments, millions of unique genes can be generated combinatorially. In addition, imprecise joining of these segments and an unusually high rate
10 of somatic mutation within these segments further contribute to the generation of a diverse antibody population.

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
15 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.

Breast Cancer

There are more than 180,000 new cases of breast cancer diagnosed each year, and the mortality rate for breast cancer approaches 10% of all deaths in females between the ages of 45-54
25 (K. Gish (1999) AWIS Magazine 28:7-10). However the survival rate based on early diagnosis of localized breast cancer is extremely high (97%), compared with the advanced stage of the disease in which the tumor has spread beyond the breast (22%). Current procedures for clinical breast examination are lacking in sensitivity and specificity, and efforts are underway to develop comprehensive gene expression profiles for breast cancer that may be used in conjunction with conventional screening methods to improve diagnosis and prognosis of this disease (Perou CM et al. (2000) Nature 406:747-752).

30 Breast cancer is a genetic disease commonly caused by mutations in cellular disease. Mutations in two genes, BRCA1 and BRCA2, are known to greatly predispose a woman to breast cancer and may be passed on from parents to children (Gish, supra). However, this type of hereditary breast cancer accounts for only about 5% to 9% of breast cancers, while the vast majority of breast

WO 02/079441

PCT/US02/09820

cancer is due to noninherited mutations that occur in breast epithelial cells.

A good deal is already known about the expression of specific genes associated with breast cancer. For example, the relationship between expression of epidermal growth factor (EGF) and its receptor, EGFR, to human mammary carcinoma has been particularly well studied. (See Khazaie et al., supra, and references cited therein for a review of this area.) Overexpression of EGFR, particularly coupled with down-regulation of the estrogen receptor, is a marker of poor prognosis in breast cancer patients. In addition, EGFR expression in breast tumor metastases is frequently elevated relative to the primary tumor, suggesting that EGFR is involved in tumor progression and metastasis. This is supported by accumulating evidence that EGF has effects on cell functions related to metastatic potential, such as cell motility, chemotaxis, secretion and differentiation. Changes in expression of other members of the erbB receptor family, of which EGFR is one, have also been implicated in breast cancer. The abundance of erbB receptors, such as HER-2/neu, HER-3, and HER-4, and their ligands in breast cancer points to their functional importance in the pathogenesis of the disease, and may therefore provide targets for therapy of the disease (Bacus, SS et al. (1994) Am J Clin Pathol 102:S13-S24). Other known markers of breast cancer include a human secreted frizzled protein mRNA that is downregulated in breast tumors; the matrix G1a protein which is overexpressed in human breast carcinoma cells; Drg1 or RTP, a gene whose expression is diminished in colon, breast, and prostate tumors; maspin, a tumor suppressor gene downregulated in invasive breast carcinomas; and CaNI9, a member of the S100 protein family, all of which are down regulated in mammary carcinoma cells relative to normal mammary epithelial cells (Zhou Z et al. (1998) Int J Cancer 78:95-99; Chen, L et al. (1990) Oncogene 5:1391-1395; Ulrich W et al (1999) FEBS Lett 455:23-26; Sager, R et al. (1996) Curr Top Microbiol Immunol 213:51-64; and Lee, SW et al. (1992) Proc Natl Acad Sci USA 89:2504-2508).

Cell lines derived from human mammary epithelial cells at various stages of breast cancer provide a useful model to study the process of malignant transformation and tumor progression as it has been shown that these cell lines retain many of the properties of their parental tumors for lengthy culture periods (Wistuba II et al. (1998) Clin Cancer Res 4:2931-2938). Such a model is particularly useful for comparing phenotypic and molecular characteristics of human mammary epithelial cells at various stages of malignant transformation.

30 Colon Cancer

Colorectal cancer is the second leading cause of cancer deaths in the United States. Colon cancer is associated with aging, since 90% of the total cases occur in individuals over the age of 55. A widely accepted hypothesis is that several contributing genetic mutations must accumulate over time

WO 02/079441

PCT/US02/09820

in an individual who develops the disease. To understand the nature of genetic alterations in colorectal cancer, a number of studies have focused on the inherited syndromes. The first known inherited syndrome, Familial Adenomatous Polyposis (FAP), is caused by mutations in the Adenomatous Polyposis Coli gene (APC), resulting in truncated or inactive forms of the protein. This tumor suppressor gene has been mapped to chromosome 5q. The second known inherited syndrome is hereditary nonpolyposis colorectal cancer (HNPCC), which is caused by mutations in mismatch repair genes.

Although hereditary colon cancer syndromes occur in a small percentage of the population and most colorectal cancers are considered sporadic, knowledge from studies of the hereditary syndromes can be generally applied. For instance, somatic mutations in APC occur in at least 80% of indiscriminate colon tumors. APC mutations are thought to be the initiating event in the disease. Other mutations occur subsequently. Approximately 50% of colorectal cancers contain activating mutations in ras, while 85% contain inactivating mutations in p53. Changes in these genes lead to gene expression changes in colon cancer. Less is understood about downstream targets of these mutations and the role they may play in cancer development and progression.

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 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.

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," and "SECP-25." 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:1-25, 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:1-25, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ

WO 02/079441

PCT/US02/09820

ID NO:1-25. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-25.

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:1-25, 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:1-25, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-25. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:26-50.

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:1-25, 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:1-25, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25. 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:1-25, 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:1-25, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25. 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

WO 02/079441

PCT/US02/09820

selected from the group consisting of SEQ ID NO:1-25, 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:1-25, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25, and d) an immunogenic
5 fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25.

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:26-50, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least
10 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:26-50, 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
15 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:26-50, 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:26-50, 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
20 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
25 said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID
30 NO:26-50, 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:26-50, 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

WO 02/079441

PCT/US02/09820

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
5 selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-25, 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:1-25, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25, and d) an immunogenic fragment of
10 a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-25. 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.

15 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:1-25, 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:1-25, c) a biologically active fragment of a polypeptide
20 having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25. 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
25 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
30 amino acid sequence selected from the group consisting of SEQ ID NO:1-25, 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:1-25, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25,

WO 02/079441

PCT/US02/09820

and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a
5 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.

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
10 sequence selected from the group consisting of SEQ ID NO:1-25, 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:1-25, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group
15 consisting of SEQ ID NO:1-25. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino
20 acid sequence selected from the group consisting of SEQ ID NO:1-25, 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:1-25, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-25, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group
25 consisting of SEQ ID NO:1-25. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test
30 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 sequence selected from the group consisting of SEQ ID NO:26-50, the method

WO 02/079441

PCT/US02/09820

comprising a) exposing a sample comprising the target polynucleotide to a compound, 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.

5 The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:26-50, ii) a
10 polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:26-50, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological
15 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:26-50, 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:26-50, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of
20 ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of
25 toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

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

30 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 each polypeptide and its homolog(s) are also shown.

WO 02/079441

PCT/US02/09820

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

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

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

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

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

DESCRIPTION OF THE INVENTION

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

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

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

DEFINITIONS

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

WO 02/079441

PCT/US02/09820

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

"Altered" nucleic acid sequences encoding SECP include those sequences with deletions,
15 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 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
20 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 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
25 include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

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

WO 02/079441

PCT/US02/09820

to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence.

Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

5 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 directly interacting with SECP or by acting on components of the biological pathway in which SECP participates.

10 The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind SECP 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, 15 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 20 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 25 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 30 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.

WO 02/079441

PCT/US02/09820

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
5 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 groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having
15 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
20 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
25 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
30 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

WO 02/079441

PCT/US02/09820

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
 5 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
 10 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
 15 conservative amino acid substitutions.

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
20	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
25	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
30	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
35	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
 40 the side chain.

WO 02/079441

PCT/US02/09820

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

A fragment of SEQ ID NO:26-50 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:26-50, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:26-50 is useful, for

WO 02/079441

PCT/US02/09820

example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:26-50 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:26-50 and the region of SEQ ID NO:26-50 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

5 A fragment of SEQ ID NO:1-25 is encoded by a fragment of SEQ ID NO:26-50. A fragment of SEQ ID NO:1-25 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-25. For example, a fragment of SEQ ID NO:1-25 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-25. The precise length of a fragment of SEQ ID NO:1-25 and the region of SEQ ID NO:1-25 to which
10 the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

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

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

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

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of
25 molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent
30 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

WO 02/079441

PCT/US02/09820

several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62
Reward for match: 1
Penalty for mismatch: -2
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 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,

WO 02/079441

PCT/US02/09820

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
5 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.

10 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

15 *Open Gap: 11 and Extension Gap: 1 penalties*

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

20 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
25 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.

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

"Hybridization" refers to the process by which a polynucleotide strand anneals with a

WO 02/079441

PCT/US02/09820

complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity.

10 Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

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

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

30 The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_{64} or R_{64} analysis) or formed between one

WO 02/079441

PCT/US02/09820

nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

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

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

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.

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

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

20 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.

25 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.

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

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs

WO 02/079441

PCT/US02/09820

preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an 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 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; Imis, 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

WO 02/079441

PCT/US02/09820

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
5 Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource
10 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,
15 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.
20 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.
25 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
30 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,

WO 02/079441

PCT/US02/09820

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
5 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,
10 chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

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

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

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

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

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

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid

WO 02/079441

PCT/US02/09820

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

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

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

WO 02/079441

PCT/US02/09820

other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

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

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 cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders.

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

WO 02/079441

PCT/US02/09820

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
5 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
10 phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

15 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:2 is 50% identical, from residue F145 to residue D308, to human SrCyp (immunophilin/cyclophilin) protein (GenBank ID g1770526) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.6e-40, which indicates the probability of obtaining
20 the observed polypeptide sequence alignment by chance. SEQ ID NO:2 also contains a cyclophilin type peptidyl-prolyl cis-trans isomerase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:2 is a cyclophilin type peptidyl-prolyl cis-trans isomerase.

25 In an alternative example, SEQ ID NO:3 is 53% identical, from residue P223 to residue G333, to a novel collagen protein in chicken (GenBank ID g211610) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.9e-20, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ
30 ID NO:3 also contains a collagen triple helix repeat and an EGF-like domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS and MOTIFS analyses provide further corroborative evidence that SEQ ID NO:3 is a secreted protein (note that EGF-like domains are found in proteins known to be secreted).

WO 02/079441

PCT/US02/09820

In an alternative example, SEQ ID NO:17 is 64% identical, from residue M11 to residue Y60, to feline major allergen I (GenBank ID g163825) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.5e-11, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:17 also contains an uteroglobin family signature as determined using the ProfileScan algorithm which searches for structural and sequence motifs as defined in the Prosite database of protein families and domains. (See Table 3.) Data from BLIMPS and MOTIFS analyses provide further corroborative evidence that SEQ ID NO:17 is a secreted protein (note that "feline major allergen" (fel D) is a protein known to be structurally related to uteroglobulin, both being secreted as mature proteins, Morgenstern, J.P., et al., (1991) Proc. Natl. Acad. Sci. U.S.A. 88:9690-9694).

In an alternative example, SEQ ID NO:20 is 98% identical, from residue M1 to residue N541, to human novel Immunoglobulin domains containing protein (isoform 1) (GenBank ID g9930918) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 3.9e-294, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:20 also contains an Immunoglobulin 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 analyses provide further corroborative evidence that SEQ ID NO:20 is a secreted protein (note that "immunoglobulin domains" are characteristic of matrix proteins).

In an alternative example, SEQ ID NO:23 is 76% identical, from residue M1 to residue L154 and is 39% identical, from residue G47 to residue G188, to murine adipocyte-specific protein 3 (GenBank ID g15777917) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 3.1e-79 and 1.3e-23, which indicate the probabilities of obtaining the observed polypeptide sequence alignments by chance. SEQ ID NO:23 also has homology to a human protein containing two immunoglobulin (Ig) domains, which may be involved in protein-protein and protein-ligand interactions, as determined by BLAST analysis using the PROTEOME database. SEQ ID NO:23 also contains an immunoglobulin domain (e-value: 3.6e-3) as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains.

SEQ ID NO:1, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, and SEQ ID NO:25 were analyzed and annotated in a similar manner. The

WO 02/079441

PCT/US02/09820

algorithms and parameters for the analysis of SEQ ID NO:1-25 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
5 identification 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
10 example, in hybridization or amplification technologies that identify SEQ ID NO:26-50 or that
distinguish between SEQ ID NO:26-50 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
15 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
20 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
25 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
30 FLXXXXXX_gAAAAA_gBBBBB_1_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,

WO 02/079441

PCT/US02/09820

and *N* referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, gBBBBB).

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

Prefix	Type of analysis and/or examples of programs
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.

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 embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:26-50, which encodes SECP. The polynucleotide sequences of SEQ ID NO:26-50, as presented in the Sequence Listing, embrace the equivalent RNA sequences,

WO 02/079441

PCT/US02/09820

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
5 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:26-50 which has at least about 70%, or alternatively at least about 85%, or even at least about 95%
10 polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:26-50. 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
15 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
20 about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide sequence encoding SECP. For example, a polynucleotide comprising a sequence of SEQ ID NO:48 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:29; a polynucleotide comprising a sequence of SEQ ID NO:49 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:43; and a polynucleotide comprising a sequence of SEQ ID NO:50 is a splice variant
25 of a polynucleotide comprising a sequence of SEQ ID NO:47. 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
30 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

WO 02/079441

PCT/US02/09820

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 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:26-50 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.

WO 02/079441

PCT/US02/09820

(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, 5 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 from a circularized template. The template is derived from restriction fragments comprising 10 a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (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 15 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 20 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 25 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 30 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

WO 02/079441

PCT/US02/09820

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.

5 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.

10 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 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

15 sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve

20 the biological properties of SECP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively,

25 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

30 naturally occurring genes in a directed and controllable manner.

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

WO 02/079441

PCT/US02/09820

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, 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 occurring polypeptide.

10 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.)

15 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 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 20 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, 25 exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

30 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

WO 02/079441

PCT/US02/09820

Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding SECP. These include, but are not limited to, microorganisms such as bacteria transformed
5 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);
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
10 (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,
15 adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for
delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola,
M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA
90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994)
Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The
20 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
25 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.
30 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

WO 02/079441

PCT/US02/09820

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

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

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding 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
20 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 virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

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

For long term production of recombinant proteins in mammalian systems, stable expression of
30 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 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

WO 02/079441

PCT/US02/09820

being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

5 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 *ap^r* cells, respectively. (See, e.g., Wigler, M. et al. (1977) *Cell* 11:223-232; Lowy, I. et al. (1980) *Cell* 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to
10 methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci. USA* 77:3567-3570; Colbere-Garapin, F. et al. (1981) *J. Mol. Biol.* 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) *Proc.*
15 *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. (1995) *Methods Mol. Biol.* 55:121-131.)

20 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 marker gene can be placed in tandem with a sequence encoding SECP under the control of a single
25 promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding 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
30 amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of SECP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques

WO 02/079441

PCT/US02/09820

include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on SECP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g.,

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

A wide variety of labels and conjugation techniques are known by those skilled in the art and
10 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 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,
15 and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for case of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic
20 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 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
25 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 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
30 lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture

WO 02/079441

PCT/US02/09820

Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding 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 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). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the SECP encoding 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 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 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 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 binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the

WO 02/079441

PCT/US02/09820

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 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 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 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 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, 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

WO 02/079441

PCT/US02/09820

system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding 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).

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 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).

20 THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of SECP and secreted proteins. In addition, examples of tissues expressing SECP are normal and cancerous breast tissues, and normal and cancerous colon tissues, and also can be found in Table 6. Therefore, SECP appears to play a role in 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 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 cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary

WO 02/079441

PCT/US02/09820

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 lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a 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-Strausler-Scheinker

WO 02/079441

PCT/US02/09820

syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system 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, 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, 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 mucopolysaccharidosis, 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.

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.

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 cell proliferative, autoimmune/inflammatory, cardiovascular, 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

WO 02/079441

PCT/US02/09820

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 expression or activity of SECP including, but not limited to, those described above.

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

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

WO 02/079441

PCT/US02/09820

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.

5 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.)

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

15 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.)

20 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.)

25 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

WO 02/079441

PCT/US02/09820

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
5 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
10 SECP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a
ranging from about 10^8 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 dissociation of SECP, preferably in active form, from the antibody
15 (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
20 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
30 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

WO 02/079441

PCT/US02/09820

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
5 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.*
10 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
15 immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaise, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii)
20 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
25 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.

30 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 vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic

WO 02/079441

PCT/US02/09820

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. Ráscipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

5 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); 10 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 20 polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

25 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 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 30 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. 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

WO 02/079441

PCT/US02/09820

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
5 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 documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; Bouyhadi, 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
15 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 incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) *Nature* 18:389:239-242, both incorporated by reference herein.

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 respect to the expression of SECP. The use of herpes simplex virus (HSV)-based vectors may be
25 especially valuable for introducing SECP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Lin, 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
30 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

WO 02/079441

PCT/US02/09820

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
5 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,
10 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
15 (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
20 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
25 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,
30 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

WO 02/079441

PCT/US02/09820

by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example,
5 engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze 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,
10 corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared
15 by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding 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
20 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 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
25 within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a
30 compound which is effective in altering expression of a polynucleotide encoding SECP. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular

WO 02/079441

PCT/US02/09820

chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased 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

WO 02/079441

PCT/US02/09820

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.)

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

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient.

10 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 (Mack Publishing, Easton PA). Such compositions may consist of SECP, antibodies to SECP, and mimetics, agonists, antagonists, or inhibitors of SECP.

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

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the
20 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
25 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.

Specialized forms of compositions may be prepared for direct intracellular delivery of
30 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

WO 02/079441

PCT/US02/09820

transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, 5 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.

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 10 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 15 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 20 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 25 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 μg to 100,000 μg, 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. 30 Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

WO 02/079441

PCT/US02/09820

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 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 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 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 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 invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:26-50 or from

WO 02/079441

PCT/US02/09820

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 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 with expression of SECP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, 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 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 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

WO 02/079441

PCT/US02/09820

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, 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, 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, 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 cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss. The polynucleotide sequences encoding SECP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered SECP

WO 02/079441

PCT/US02/09820

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 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
5 from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding SECP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the
10 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 normal or standard profile for expression is established. This may be accomplished by combining
15 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. 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
20 presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several
25 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
30 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

WO 02/079441

PCT/US02/09820

enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding SECP, 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
5 quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences 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
10 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 (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
15 detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-
20 based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

SNPs may be used to study the genetic basis of human disease. For example, at least 16
25 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
30 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

WO 02/079441

PCT/US02/09820

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.)

5 Methods which may also be used to quantify the expression of SECP include radiolabeling or biotinylation 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
10 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
15 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
20 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
25 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 gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by
30 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. 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of

WO 02/079441

PCT/US02/09820

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.

5 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
10 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
15 signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the
20 rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released
25 February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated
30 biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are

WO 02/079441

PCT/US02/09820

indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl 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 proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for SECP to quantify the levels of SECP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendozze, L.G. et al. (1999) *Biotechniques* 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson,

WO 02/079441

PCT/US02/09820

N.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537, so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profiles. 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.

5 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
10 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
15 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.,
20 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,
25 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
30 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

WO 02/079441

PCT/US02/09820

constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355; Price, C.M. (1993) *Blood Rev.* 7:127-134; and Trask, B.J. (1991) *Trends Genet.* 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) *Proc. Natl. Acad. Sci. USA* 83:7353-7357.)

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, *supra*, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding 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.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) *Nature* 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, 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 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 washed. Bound SECP is then detected by methods well known in the art. Purified SECP can

WO 02/079441

PCT/US02/09820

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 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 properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, including U.S. Ser. No. 60/282,112, U.S. Ser. No. 60/280,527, U.S. Ser. No. 60/282,702, U.S. Ser. No. 60/283,855, U.S. Ser. No. 60/357,002, U.S. Ser. No. 60/339,236, and U.S. Ser. No. 60/343,718 are expressly incorporated by reference herein.

EXAMPLES

I. Construction of cDNA Libraries

Incye cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incye 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,

WO 02/079441

PCT/US02/09820

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 SUPERSKRIP 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., PBLUESKRIP plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or pINCY (Incyte 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

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 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 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

WO 02/079441

PCT/US02/09820

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the

5 MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the

10 ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

15 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 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

20 BLOCKS, PRINTS, 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

25 (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.

30 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

WO 02/079441

PCT/US02/09820

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
5 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
10 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
15 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
20 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:26-50. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

25 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., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) *J. Mol. Biol.* 268:78-94, and Burge, C. and S. Karlin
30 (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 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

WO 02/079441

PCT/US02/09820

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 cDNA sequences that had been annotated as secreted proteins. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpr public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

15 "Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpr public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended

WO 02/079441

PCT/US02/09820

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 Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of SECP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:26-50 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:26-50 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease

WO 02/079441

PCT/US02/09820

genes map within or in proximity to the intervals indicated above.

In this manner, SEQ ID NO:29 was mapped to chromosome 3 within the interval from 97.20 to 115.70 centiMorgans. SEQ ID NO:46 was mapped to chromosome 4 within the interval from 111.10 to 123.50 centiMorgans. SEQ ID NO:47 was mapped to chromosome 19 within the interval
5 from 61.40 to 65.80 centiMorgans and to chromosome 19 within the interval from 62.00 to 64.80 centiMorgans.

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
10 from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Anselmel (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 LIPSEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer
15 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:

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

20

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
25 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
30 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.

WO 02/079441

PCT/US02/09820

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

5 classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of

10 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

15 information are found in the LIFESEQ 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

20 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.

25 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+} , $(NH_4)_2SO_4$,

30 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

WO 02/079441

PCT/US02/09820

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 µl PICOGREEN
5 quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE
and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar,
Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II
(Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the
concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by
10 electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the
sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates,
digested with CviII cholera virus endonuclease (Molecular Biology Research, Madison WI), and
sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For
15 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
20 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
25 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
30 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.

WO 02/079441

PCT/US02/09820

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:26-50 using the LIF/ESQ 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 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:26-50 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

WO 02/079441

PCT/US02/09820

(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 10⁷ 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 (Schena (1999), *supra*). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., 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.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described

WO 02/079441

PCT/US02/09820

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 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 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 the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 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 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).

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.

WO 02/079441

PCT/US02/09820

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60° C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 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.

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 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 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

WO 02/079441

PCT/US02/09820

different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

5 The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and
10 measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used
15 for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

Expression

In this way it was shown that SEQ ID NO:46, is downregulated at least 2.2-fold in 4 out of 5 human breast cancer cell lines and in 5 out of 6 human lung adenocarcinoma cells and lung squamous cell carcinoma tumors. These experiments indicate that SEQ ID NO:21 and SEQ ID NO:46 have
20 utility in disease staging or in the use of SEQ ID NO:21 as a potential therapeutic target for cancer and tumors.

Propagation of Human Epithelial Cell Lines

HMEC is a human primary mammary epithelial cell strain derived from normal mammary tissue (Clonetics San Diego, CA). The following cell lines were obtained from ATCC (Manassus,
25 VA): Sk-BR3 is a breast adenocarcinoma cell line isolated from a malignant pleural effusion of a 43-year old female; MCF7 is a breast adenocarcinoma cell line derived from the pleural effusion of a 69-year old female; T47D is a breast carcinoma cell line derived from a pleural effusion from a 54-year old female with an infiltrating ductal carcinoma of the breast; BT20 is a breast carcinoma cell line derived *in vitro* from cells emigrating out of thin slices of a tumor mass isolated from a 74-year old
30 female; MDA-mb-435S and MDA-mb-231 are metastatic breast tumor cell line derived from the pleural effusion of a 51-year old female with metastatic breast carcinoma. All cell cultures were propagated in media according to the supplier's recommendations and grown to 70-80% confluence prior to RNA isolation.

WO 02/079441

PCT/US02/09820

The expression of cDNAs from the six tumor cell lines representing various stages of breast tumor progression (BT20, MCF7, MDA-mb-231, MDA-mb-435S, SKBr3, and T47D) were compared with that of the non-malignant mammary epithelial cell line, HMEC. Array elements that exhibited about at least a 2-fold change in expression and a signal intensity over 250 units, a signal-to-background ratio of at least 2.5, and an element spot size of at least 40% were identified as differentially expressed using the GEMTOOLS program (Incyte Genomics). SEQ ID NO:28 showed 2.75 fold lower expression vs. T47D cell, 3.7 fold lower expression vs. MD-am-B435S cells, 4.6 fold lower expression vs. BT20 cells, 4.7 fold lower expression vs. SKBr3 cells, 5.1 fold lower expression vs. MCF7 cells and 1.9 fold lower expression vs. MD-am-B231 cells. These experiments indicate that SEQ ID NO:28 was significantly under-expressed in five out of six breast tumor cell lines tested, further establishing the utility of SEQ ID NO:28 as a diagnostic marker or as a potential therapeutic target for breast cancer.

For example, SEQ ID NO:48 showed differential expression in inflammatory responses as determined by microarray analysis. The expression of SEQ ID NO:48 was increased by at least two fold in THP-1 human promonocyte line which had been stimulated for 48 hours with 100 nM PMA (phorbol 12-myristate 13-acetate) when compared to untreated THP-1 cells. THP-1 is promonocyte line derived from peripheral blood of a 1 year old male with acute monocytic leukemia. The cell line acquires monocytic characteristics upon stimulation with PMA. Monocytes play a critical role in the initiation and maintenance of inflammatory immune responses.

In another example, the expression of SEQ ID NO:48 was decreased by at least two fold in peripheral blood mononuclear cells (PBMCs, 12% B lymphocytes, 40% T lymphocytes, 20% NK cells, 25% monocytes, and 3% various cells that include dendritic and progenitor cells) treated with interleukin-10 (IL-10). IL-10 is a pleiotropic cytokine that can exert either immunostimulatory or immunosuppressive effects on a variety of cell types. As an immunostimulatory cytokine, IL-10 can act on B lymphocytes to enhance their viability, cell proliferation, Ig secretion, and class II MHC expression. Aside from B lymphocytes, IL-10 is also a growth co-stimulator for thymocytes and mast cells, as well as an enhancer of cytotoxic T-cell development. As an immunosuppressant cytokine, IL-10 is a potent inhibitor of monocyte/macrophage activation and its resultant cytotoxic effects. It can suppress the production of numerous cytokines including TNF- α , IL-1, IL-6, and IL-10, as well as the synthesis of superoxide anion, reactive oxygen intermediates, and reactive nitrogen intermediates by activated monocytes/macrophages.

In yet another example, the expression of SEQ ID NO:48 was decreased by at least two fold in human PBMCs with mixed lymphocyte reaction (MLR) relative to untreated PBMCs. When

WO 02/079441

PCT/US02/09820

lymphoid cells from genetically distinct animals of the same species are mixed together in tissue culture, mixed lymphocyte reaction (MLR) can occur. MLR is a recognized model of T cell activation and is relevant for investigating a number of syndromes where such activation takes place, including graft-verses-host disease and transplant rejection (host-verses-graft disease). In this experiment, human
5 PBMCs were collected from the blood of six healthy volunteer donors using standard gradient separation.

The three experiments described above show that SEQ ID NO:48 is useful in diagnostic assays for inflammatory responses.

In an alternative example, SEQ ID NO:48 showed differential expression in certain breast
10 carcinoma and mammary gland cell lines versus primary mammary epithelial cells as determined by microarray analysis. The breast carcinoma cell lines include BT20, a breast carcinoma cell line derived in vitro from cells emigrating out of thin slices of a tumor mass isolated from a 74-year-old female; BT474, a breast ductal carcinoma cell line isolated from a solid, invasive ductal carcinoma of the breast from a 60-year-old female; BT483, a breast ductal carcinoma cell line isolated from a
15 papillary invasive ductal tumor from a 23-year-old normal, menstruating, parous female; HS578T, a breast ductal carcinoma cell line isolated from a 74-year-old female with breast carcinoma; MCF7, a breast adenocarcinoma cell line derived from the pleural effusion of a 69-year-old female; and MDA-
mb-468, a breast adenocarcinoma cell line isolated from the pleural effusion of a 51-year-old female with metastatic adenocarcinoma of the breast. MCF10A, a breast mammary gland (luminal ductal
20 characteristics) cell line that was isolated from a 36 year old woman with fibrocystic breast disease was also compared. The primary mammary epithelial cell line HMEC was derived from normal human mammary tissue (Clonetics, San Diego, CA). The microarray experiments showed that the expression of SEQ ID NO:48 was increased by at least four fold in HS578T breast carcinoma line and decreased by at least two fold in MDA-
mb-468 breast carcinoma line and MCF10A breast mammary
25 gland cell line relative to cells from the primary mammary epithelial cell line, HMEC. Therefore, SEQ ID NO:48 is useful as diagnostic markers or as potential therapeutic targets for breast cancer and fibrocystic breast diseases.

In a further example, the expression of SEQ ID NO:48, as determined by microarray analysis, was increased by at least two fold in tumorous colon tissues relative to normal colon tissues. Both the
30 tumor and the normal colon tissues were isolated from a 58 year old female diagnosed with mucinous adenocarcinoma. Therefore, SEQ ID NO:48 is useful as a diagnostic marker or as a potential therapeutic target for colon cancer.

XII. Complementary Polynucleotides

WO 02/079441

PCT/US02/09820

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 5 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

10 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 15 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 20 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. 25 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 enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized 30 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-

WO 02/079441

PCT/US02/09820

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.

5 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. 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). 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

WO 02/079441

PCT/US02/09820

analysis or microarray techniques.

XV. Production of SECP Specific Antibodies

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 anti-peptide 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.

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

WO 02/079441

PCT/US02/09820

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 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).

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

Peptidyl prolyl *cis/trans* isomerase activity of SECP can be assayed by an enzyme assay described by Rahfeld, J.U., et al. (1994; FEBS Lett. 352: 180-184). The assay is performed at 10°C in 35 mM HEPES buffer, pH 7.8, containing chymotrypsin (0.5 mg/ml) and SECP at a variety of concentrations. Under these assay conditions, the substrate, Suc-Ala-Xaa-Pro-Phe-4-NA, is in equilibrium with respect to the prolyl bond, with 80-95% in *trans* and 5-20% in *cis* conformation. An aliquot (2 ul) of the substrate dissolved in dimethyl sulfoxide (10 mg/ml) is added to the reaction mixture described above. The *trans* to *cis* conversion is measured by the hydrolysis of the *cis* conformer by chymotrypsin, producing 4-nitroanilide which is detected spectrophotometrically by absorbance at 390 nm. 4-Nitroanilide appears in a time-dependent manner and is proportional to the amount of SECP in the assay.

Alternatively, 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. 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 acid-precipitable DNA.

Alternatively, an assay for SECP activity measures the stimulation or inhibition of

WO 02/079441

PCT/US02/09820

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 medium. Membrane currents are recorded from the myocyte. Increased or decreased current
5 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 is fractionated using methods known to those of skill in the art, for example, sucrose gradient
10 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 concentration of SECP in secretory organelles relative to SECP in total cell lysate is proportional to
15 the amount of SECP in transit through the secretory pathway.

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.

20 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
25 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
30 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.

WO 02/079441

PCT/US02/09820

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
5 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
10 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.

15 Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious
20 to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

WO 02/079441

PCT/US02/09820

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	CA2 Reagents
7757335	1	7757335CD1	26	7757335CB1	
1482559	2	1482559CD1	27	1482559CB1	
2234213	3	2234213CD1	28	2234213CB1	
1345785	4	1345785CD1	29	1345785CB1	90078841CA2
3807190	5	3807190CD1	30	3807190CB1	
4856078	6	4856078CD1	31	4856078CB1	4856078CA2, 900796645CA2
6106886	7	6106886CD1	32	6106886CB1	56050011CA2, 6106886CA2, 90079304CA2, 90079312CA2, 90079320CA2, 90079344CA2, 90079412CA2, 90079420CA2, 90079436CA2, 90079444CA2, 90079582CA2, 90079600CA2
5627037	8	5627037CD1	33	5627037CB1	
3688835	9	3688835CD1	34	3688835CB1	3688835CA2
2295419	10	2295419CD1	35	2295419CB1	5391296CA2
2437896	11	2437896CD1	36	2437896CB1	6519813CA2
2641482	12	2641482CD1	37	2641482CB1	
7594292	13	7594292CD1	38	7594292CB1	90067354CA2
7636238	14	7636238CD1	39	7636238CB1	9008978CA2, 9008964CA2
8095391	15	8095391CD1	40	8095391CB1	6053256CA2, 7644258CA2, 8095391CA2, 8324949CA2
532018	16	532018CD1	41	532018CB1	90090702CA2, 90090710CA2, 90090842CA2
7481063	17	7481063CD1	42	7481063CB1	
2044436	18	2044436CD1	43	2044436CB1	
4091564	19	4091564CD1	44	4091564CB1	4091540CA2, 4091564CA2, 4092602CA2
8039739	20	8039739CD1	45	8039739CB1	
126837	21	126837CD1	46	126837CB1	8764942CA2
5568527	22	5568527CD1	47	5568527CB1	1338190CA2
7503641	23	7503641CD1	48	7503641CB1	
7503458	24	7503458CD1	49	7503458CB1	
7500925	25	7500925CD1	50	7500925CB1	

WO 02/079441

PCT/US02/09820

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
1	757353CD1	g2138390	0	[Homo sapiens] acetyl-CoA carboxylase Abu-Elheiga, L., et al. (1997) J. Biol. Chem. 272:10669-10677
2	1482539CD1	g1770826	1.6E-40	[Homo sapiens] SRpP protein Boutquin, J.P. (1997) Nucleic Acids Res. 25:2055-2061
3	2234213CD1	g211610	1.9E-20	[Gallus gallus] novel collagen protein Marchant, J.K., et al. (1991) Proc. Natl. Acad. Sci. U.S.A. 88:1560-1564
4	1345785CD1	g15777917	0.0E+00	adipocyte-specific protein 3 [Mus musculus]
16	532018CD1	g15778556	0.0E+00	alpha-1-B glycoprotein precursor [Homo sapiens]
		g11877348	2.3E-90	[Rattus norvegicus] putative alpha 1B-glycoprotein Ishioh, N., et al., (1986) Proc Natl Acad Sci U.S.A. 83:2563-7
17	7481063CD1	g163825	1.5E-11	[Felis catus] major allergen 1 [Mus musculus]
18	2044436CD1	g13488611	3.0E-79	retinidin [Homo sapiens]
19	4091564CD1	g16974791	4.0E-05	dragline silk protein spidroin 2 [Nephila clavata]
20	8039739CD1	g0930918	3.9E-294	[Homo sapiens] hA36717.2.1 (novel immunoglobulin domains containing protein (isoform 1))
20	8039739CD1	g16033594	0.0E+00	SH2 domain-containing phosphatase anchor protein 2c [Homo sapiens]
20	8039739CD1	g18092655	0.0E+00	immunoglobulin superfamily receptor unassociated protein 3 [Homo sapiens]
23	7503641CD1	g13777917	4.0E-56	[Mus musculus] adipocyte-specific protein 3
		716615M GCC304	2.6E-80	[Homo sapiens] Protein containing two immunoglobulin (Ig) domains, which may be involved in protein-protein and protein-ligand interactions
24	7504456CD1	g13488611	1.9E-43	[Homo sapiens] retinidin

WO 02/079441

PCT/US02/09820

Table 2

Polypeptide SEQ ID NO:	Lucyite Polypeptide ID	GenBank ID NO: or PROTEOME ID NO:	Probability Score	Annotation
25	7500925CD1	377422 p083	8.4E-27	[Schizosaccharomyces pombe] Conserved protein containing a DUF28 domain
		6910 YGR021W	8.2E-20	[Saccharomyces cerevisiae] Likely involved in mitochondrial protein synthesis
		634830 orf6.2546	2.9E-17	[Candida albicans] Protein containing an uncharacterized domain, has low similarity to S. cerevisiae Ygr021, which is possibly involved in mitochondrial protein synthesis

Table 3

SEQ. Ineye ID No.	Ineye Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	775733CD1	2487	S35 S72 S78 S91 S105 S136 S148 S169 S174 S175 S183 S187 S192 S246 S278 S423 S469 S676 S705 S725 S876 S921 S998 S1018 S1060 S1226 S1242 S1278 S1437 S1443 S1477 S1613 S1639 S2099 S2276 S2297 S2467 S2479 T188 T206 T216 T300 T412 T668 T732 T909 T1031 T1086 T1301 T1371 T1491 T1505 T1773 T1825 T1894 T1961 T2024 T2034 T2054 T2085 T2113 T2350 T2457 T2443 Y1643 Y1774 Y2284	N118 N683 N1002 N1071 N1153 N1453 N1610 N1611 N1662 N2431	Signal Peptide: M1-T16	SPSCAN
					Signal Peptide: M1-W19	HMMER

WO 02/079441

PCT/US02/09820

Table 3

SEQ ID NO: 1	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1					Carbamoyl-phosphate synthase (CPSase); K262-S685	HMMER_FFAM
					Carboxyl transferase domain; G1794-L2363	HMMER_FFAM
					Biotin-requiring enzymes; T895-L961	HMMER_FFAM
					Transmembrane Segments: L4-I22 Q528-K556 Y822-H850 I1954-R1973	TMAP
					N-terminus non-cytosolic	
					Biotin-requiring enzymes BL00188; P900-1945	BLIMPS_BLOCKS
					Carbamoyl phosphate synthetase subdomain	BLIMPS_BLOCKS
					BL00866; F490-F535; A1909-Y2003	
					Biotin-requiring enzymes attachment site biotin.prf; L905-A954	PROFLIFESCAN
					CARBOXYLASE ACETYLCOA LIGASE BIOTIN ATP BINDING PD002092; H1177-D1632; V1629; G1793, S1070-F1144, W351-N367, L2384-W2404	BLAST_PRODROM
					CARBOXYLASE LIGASE BETA TRANSFERASE SUBUNIT FATTY ACID BIOSYNTHESIS PD001372; I1957-R2361, G1814-I1864, D241-L264	BLAST_PRODROM
					ACETYLCOA CARBOXYLASE INCLUDES: BIOTIN FATTY ACID BIOSYNTHESIS LIGASE PD011672; L2362-S2486 PD037401; M1-A155	BLAST_PRODROM

Table 3

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	cont				BIOTIN-REQUIRING ENZYMES ATTACHMENT SITE: DM01562 P11029 1510-2277; T1681-L2441 Q00955 1445-2194; N1682-D2422 P32874 1491-2243; T1681-Q2416 DM01850 P11029 619-1213; E762; F1144 Cell attachment sequence R777-D729	MOTIFS MOTIFS
					Carbamoyl phosphate synthase subdomain signature 1 F449-G463	MOTIFS
					Carbamoyl phosphate synthase subdomain signature 2 F574-L585	MOTIFS
2	1482539CD1	311	S18 S42 S53 S132 S185 T139 T170 T179 T227 T292	N81 N232 N253	Signal_peptide: M1-A40 Cyclophilin type peptidyl-prolyl cis-trans isomerase: F143-Y310 Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature: BL00170; G156-A182, Y193-N232, R240-1284 Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature & profile esa_ppiase.ppf: D165-D230	SPSCAN HMMER_PPAM BLIMPS_BLOCKS PROFILESCAN
					Cyclophilin peptidyl-prolyl cis-trans isomerase signature PF00153; Q356-D268, R269-L284, L162-L177; F198-Q210; G231-Q236	BLIMPS_PRINTS
					ISOMERASE ROTAMASE CYCLOPHILIN CISTRANS PEPTIDYLPROLYL PPIASE CYCLOSPORIN MULTIGENE FAMILY PROTEIN PF000341; V144-D308	BLAST_PRODOM

WO 02/079441

PCT/US02/09820

Table 3

SEQ ID NO: ID	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
2	cont				CYCLOPHILIN-TYPE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE DM00129 B4732858-229; D147-G307 P3041468-229; D147-G307 Q08752113-182; V144-G307 P520111-170; V144-G307	BLAST_DOMO
					SIMILAR TO CUTICULAR COLLAGEN PD067228; P248-G333	BLAST_PRODOM
					PRECURSOR SIGNAL COLLAGEN ALPHA 3(X) CHAIN EXTRACELLULAR MATRIX CONNECTIVE TISSUE PD038299; G247-R326	BLAST_PRODOM
					FIBRILLAR COLLAGEN CARBOXYL-TERMINAL DM000019; P39061455-643; S241-G333	BLAST_DOMO
					FIBRILLAR COLLAGEN CARBOXYL-TERMINAL DM000019; Q023881273-1547; L245-G333	BLAST_DOMO
					FIBRILLAR COLLAGEN CARBOXYL-TERMINAL DM000019; P29400705-886; P246-G333	BLAST_DOMO
					FIBRILLAR COLLAGEN CARBOXYL-TERMINAL DM000019; P08572178-947; P246-G333	BLAST_DOMO
					Aspartic acid and asparagine hydroxylation site C150-C161	MOTIFS
					FGF-like domain signature 2 C159-C174	MOTIFS

WO 02/079441

PCT/US02/09820

Table 3

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
3	2234213CD1	406	S62 S156 S316 S363 S364 S376 S383 T34 T54 T175 T184 T201 Y131	N142 N182	Signal Peptide: M1-K36 Signal cleavage: M1-T34 Collagen triple helix repeat (20 copies): G271-P329 EGF-like domain: C138-C174, C93-K126 Transmembrane domain: L16-W33, E319-N347 N-terminus is non-cytosolic Calcium-binding EGF-like domain proteins pattern proteins BL01187: C132-G143, C150-Y165 Type II EGF-like signature PR00010: D89-Q100, G155-Y165, A195-T201 COLLAGEN ALPHA PRECURSOR CHAIN REPEAT SIGNAL CONNECTIVE TISSUE EXTRACELLULAR MATRIX PD000007: P246-G333 PRECOLLAGEN P PRECURSOR SIGNAL PD072959: N242-S334 SIMILAR TO CUTTICULAR COLLAGEN PD067228: P248-G333 PRECURSOR SIGNAL COLLAGEN ALPHA 3IX CHAIN EXTRACELLULAR MATRIX CONNECTIVE TISSUE PD028299: G247-R326	HMMER SESCAN HMMER_PPAM HMMER_PPAM TMAP BLIMPS_BLOCKS BLIMPS_PRINTS BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM

WO 02/079441

PCT/US02/09820

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
3	cont				FIBRILLAR COLLAGEN CARBOXYL-TERMINAL DM00019/p39061455-643; S241-G333	BLAST_DOMO
					FIBRILLAR COLLAGEN CARBOXYL-TERMINAL DM00019/Q023881373-1547; L245-G333	BLAST_DOMO
					FIBRILLAR COLLAGEN CARBOXYL-TERMINAL DM00019/p29400705-886; P246-G333	BLAST_DOMO
					FIBRILLAR COLLAGEN CARBOXYL-TERMINAL DM00019/p08572178-947; P246-G333	BLAST_DOMO
					Aspartic acid and asparagine hydroxylation site C150-C161	MOTIFS
					EGF-like domain signature 2 C159-C174	MOTIFS
					Calcium-binding EGF-like domain pattern signature D134-C159	MOTIFS
4	11345785CD1	442	S56 S93 S113 S236 S377 T194 T290	N119 N306	Signal Peptide: M1-A30, M1-S26	HMMER
					Signal cleavage: M1-S24	SFSCAN
					Immunoglobulin domain: G179-L274, C47-L139	HMMER_Pfam
					Transmembrane domain: S334-A362	TMAP
					N-terminus is non-cytosolic	
5	3807190CD1	146			Signal Peptide: M1-H20, M1-P19, M1-K25, M1-Y24, M1-P26	HMMER
					Signal cleavage: M29-A89	SFSCAN
6	4856078CD1	109	T98		Signal Peptide: M1-G18	HMMER

WO 02/079441

PCT/US02/09820

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
6	cont.				Signal_cleavage: M1-T34	SFSCAN
7	6106886CD1	78	T29		Transmembrane domain: M1-G33 N-terminus is non-cytosolic Signal Peptide: M1-P27, M1-25 Signal_cleavage: M1-G24 Transmembrane domain: P4-H20N-terminus is non-cytosolic	TMAP HMIMER SFSCAN TMAP
8	5627027CD1	721	S57 S150 S157 S218 S294 S303 S491 S516 S538 S580 S602 S678 T36 T71 T90 T109 T128 T138 T147 T161 T181 T222 T265 T290 T326 T334 T496 T536 T667 T683	N88 N245 N435	Immunoglobulin domain: G270-A337, G84-S150, K5-A50, G185-G209, G370-A438	HMIMER_PPFAM
					Transmembrane domain: E455-F483 N-terminus is non-cytosolic GLYCOPROTEIN ANTIGEN PRE PD02327, L62-I73, T90-Q111 IRREGULAR CHIASM CROUCHEST PROTEIN PRECURSOR IRREC TRANSMEMBRANE IMMUNOGLOBULIN FOLD GLYCOPROTEIN SIGNAL_CELL ADHESION PD124347: 67-I177	TMAP BLIMPS_PRODROM BLAST_PRODROM
9	3688383CD1	144	S10 S79 S106 T36 T133	N82	Signal_cleavage: M1-G15	SFSCAN
10	2295419CD1	124	S102		signal_cleavage: M1-S17	SFSCAN

WO 02/079441

PCT/US02/09820

Table 3

SEQ ID NO: 10	Ineye Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
11	2437896CD1	142	S87 S99 T97		Signal Peptide: M1-R30 Transmembrane Domains: S84-F107 N-terminus is cytosolic signal_cleavage: M1-G21 Signal Peptide: M1-G21 Transmembrane Domains: L109-I132 N-terminus is cytosolic ATP/GTP-binding site motif A (P-loop) C80-S87	HMMER TMAP SFSCAN HMMER TMAP MOTIFS
12	2641482CD1	103	S83		signal_cleavage: M1-Y24 Signal Peptide: M1-Y24, M1-Q26	SFSCAN HMMER
13	7494292CD1	670	S47 S121 S129 S402 S417 S424 S441 S445 S491 S514 S584 T61 T106 T113 T156 T179 T222 T317 T327 T454 T494 T538 Y279	N57 N189	signal_cleavage: M1-A27	SFSCAN
14	7636288CD1	215	S127 S181 T99 T130		Signal Peptide: M1-Y21, M1-C25, M1-P29, M1-A27 Transmembrane Domains: F8-E36 N-terminus is non-cytosolic Recoverin family signature PR00450; L135-F109; P171-Q190 signal_cleavage: M1-S40 Signal Peptide: M1-L24, M1-C26, M1-A30	HMMER TMAP BLIMPS_PRINTS SFSCAN HMMER

WO 02/079441

PCT/US02/09820

Table 3

SEQ ID No.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
15	8095391	CD1_86	S51 S65 S78	N36	signal_cleavage: M1-A25 Signal Peptide: M1-A25	SPSCAN HAMMER
16	532018	CD1_495	S183 S237 S317 S365 S389 T186 T200 T257 T278 T445	N44 N179 N363 N371	signal_cleavage: M1-G16 Signal Peptide: M1-V18, M1-A21, M1-A22 Immunoglobulin domain: L42-S95, G225-Y281, G416-Y472, C318-Y376 Transmembrane Domains: L4-E20 N-terminus is non-cytosolic ALPHA1 B-GLYCOPROTEIN IMMUNOGLOBULIN FOLD GLYCOPROTEIN PLASMA PD138678: A22-A387 ALPHA1 B-GLYCOPROTEIN IMMUNOGLOBULIN FOLD GLYCOPROTEIN PLASMA KIAA0964 PD101369: F385-E489 IMMUNOGLOBULIN DM00001P04217574-459: H395-E481 DM00001P04217192-268: M213-W290 DM00001P04217110-82: L31-L104 DM00001P04217100-170: L12-L192	SPSCAN HAMMER HAMMER_PFAM TMAP BLAST_PRODOM BLAST_PRODOM
17	7481065	CD1_60	T10 T49		signal_cleavage: M1-C32 Signal Peptide: M1-F34, M1-C32, M1-LR29, M1-L134 Transmembrane Domains: T8-R29 N-terminus is non-cytosolic Uteroglobin signature PRO0486: K12-L26, S28-D43, D43-Y60 Uteroglobin family signature: L13-Y60	SPSCAN HAMMER TMAP BLIMPS_PRINTS PROFILSCAN

WO 02/079441

PCT/US02/09820

Table 3

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
17	20444436CD1	241	S62 S88 S142 S143		UTEROGLOBIN FAMILY DM02656(P04381)91; M11-Y60 Uteroglobin family signature 1_A30-L46 signal_cleavage: M1-G30, M7-A27, M7-G39, M1-G29, M7-A36 Transmembrane domain: M7-E26 N-terminus is non-cytosolic signal_cleavage: M1-G23	BLAST_DOMO MOTIFS SFSCAN TMAP SFSCAN
19	4091564CD1	174	S50 S62 S70 S142 T41		signal_cleavage: M1-L19	SFSCAN
20	8039739CD1	727	S81 S235 S261 S266 S399 S459 S492 S608 S617 S621 S696 S701 T170 T193 T328 T340 T368 T379 T472 T558 T691 Y349	N419 N431 N441 N512 N524 N534 N554 N561	Immunoglobulin domain: G295-A353, G197-Y255, C483-A539, G390-A446, G106-A158, G30-T77 Transmembrane domain: N561-Y589 N-terminus is non-cytosolic RECEPTOR FC IMMUNOGLOBULIN PD01270; P14_H53_165-H101_D107-N135 CELL SURFACE GLYCOPROTEIN GM2 PRECURSOR SIGNAL GPI ANCHOR MEMBRANE PD116497; P376-S559, Q177-N355, M1-P23, F313-N448	HMMER_FFAM TMAP BLIMPS_PRODOM BLAST_PRODOM

WO 02/079441

PCT/US02/09820

Table 3

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
20	count				RECEPTOR GAMMA AFFINITY IMMUNOGLOBULIN PRECURSOR PROTEIN TRANSMEMBRANE GLYCOPROTEIN SIGNAL PD002534; M1-K131	BLAST_PRODOM
					PLATELET ENDOTHELIAL CELL ADHESION PRECURSOR SIGNAL MOLECULE PECAM1 CD31 ANTIGEN PDI S0932; S362-D694	BLAST_PRODOM
					MYELIN-ASSOCIATED GLYCOPROTEIN DN00682P123198; 196-19-V178	BLAST_DOMO
					MYELIN-ASSOCIATED GLYCOPROTEIN DN00682P1422091-172; V13-V178	BLAST_DOMO
					MYELIN-ASSOCIATED GLYCOPROTEIN DN00682P14020428-216; M1-E180	BLAST_DOMO
					MYELIN-ASSOCIATED GLYCOPROTEIN DN00682P12764519-208; P14-Q179	BLAST_DOMO
					N 6 Adenine-specific DNA methylases signature L18- W24	MOTIFS
21	1265837CD1	132	S5 S38 S49 T31	N58	Transmembrane domain: T86-F111 N-terminus is non-cytosolic	TMAP
22	5566327CD1	287	S35 S120 S157 S284 T221	N135	HCP repeats signature: D19-K70 signal_cleavage: M1-A19	PROFILESCAN SFSCAN
23	7503641CD1	218	S56 S93 S113 S207	N119	Signal Peptide: M1-A26 Signal_cleavage: M1-S24	HMIMER
					Signal Peptide: M1-A19; M1-V20; M1-L22; M1-G25; M1-A30; M1-S24; M1-S27	SFSCAN
					Immunoglobulin domain: G47-L139 E-value: 3.6E-3	HMIMER_Pfam
24	7503458CD1	266	S87 S113 S167 S168		signal_cleavage: M1-G30	SFSCAN

WO 02/079441

PCT/US02/09820

Table 3

SBQ ID	Incyte Polypeptide No. ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
24	cont				Signal Peptide: M7-A27, M7-G29, M7-R32, M1-G30, M1-G29, M7-A36	HMMER
25	7500025CD1	246	S35 S120 S157 T221	N155	Signal cleavage: M1-A19 Signal Peptide: M1-A26 PROTEIN INTERGENIC REGION CONSERVED OF SECTION COAT MGG32 YMA7RP531A F17H15.14 PD004186; Y138-P236, PD004323; K62- K127 H10315; R1VC; 26-4; ASP5; DM02856 P2423720-238; E76-D223 G6414721-238; R77-D223 P441634721-238; R77-D223 P5321245-289; E76-G241	SPSCAN HMMER BLAST_PRODOM BLAST_DOMO

Table 4

Polymucleotide SEQ ID NO./ In-cyte ID/ Sequence Length	Sequence Fragments
26/7757335CBI/ 7633	1-512, 1-663, 1-836, 5-516, 6-509, 10-773, 29-702, 160-334, 160-391, 160-555, 160-655, 160-711, 160-760, 160-819, 162-713, 168-684, 182-335, 219-945, 227-891, 230-878, 238-956, 260-411, 265-823, 289-636, 318-843, 341-934, 344-805, 346-741, 355-1028, 376-943, 457-1037, 466-745, 477-1064, 483-1099, 493-1131, 501-1189, 505-1235, 512-1205, 525-1255, 539-853, 549-1121, 558-1138, 568-1215, 573-1236, 592-1096, 592-1241, 604-1195, 606-950, 606-1124, 611-1053, 622-1168, 626-1286, 633-1286, 636-832, 645-1138, 653-939, 657-1248, 675-1265, 705-899, 722-1286, 724-1287, 733-1094, 736-1286, 746-1094, 750-1094, 754-1085, 756-1286, 759-1201, 766-1286, 777-1286, 782-1286, 786-1286, 789-1286, 802-1286, 811-1286, 823-897, 823-930, 823-947, 823-1085, 823-1098, 823-1167, 823-1218, 823-1286, 856-1286, 837-1286, 837-1383, 844-1286, 857-1286, 858-1082, 875-1286, 882-1286, 892-1286, 901-1094, 901-1286, 900-1287, 908-1286, 910-1287, 912-1286, 916-1286, 928-1094, 953-2377, 954-1111, 954-1288.
26 cont.	954-1274, 954-1285, 954-1287, 957-1286, 963-1286, 964-1087, 968-1094, 969-1286, 1004-1286, 1008-1513, 1016-1150, 1017-1286, 1018-1100, 1030-1286, 1033-1286, 1053-1287, 1057-1286, 1076-1286, 1086-1286, 1086-1287, 1100-1287, 1123-680, 1127-1287, 1139-1286, 1147-1286, 1151-1287, 1164-1286, 1169-1286, 1169-1287, 1175-1287, 1181-1208, 1182-1287, 1183-1286, 1191-1286, 1193-1857, 1196-1287, 1204-1287, 1206-1286, 1214-1633, 1219-1287, 1225-1286, 1225-1287, 1227-1287, 1307-1978, 1457-2058, 1500-2092, 1550-2096, 1613-1633, 1655-2091, 1729-2361, 1865-2058, 1941-1982, 2198-2740, 2296-2786, 2415-2569, 2485-2788, 2513-2053, 2519-3045, 2520-3044, 2567-3652, 2588-2691, 2592-3183, 2756-3423, 2764-2850, 2932-2972, 2974-3456, 3228-3519, 3400-3591, 3766-3807, 3965-4105, 4665-5170, 5927-6285, 6473-6971, 7199-7369, 7232-7267.
27/1482339CBI /992	1-135, 63-135, 63-338, 63-518, 65-135, 65-583, 81-414, 100-367, 108-532, 109-523, 109-610, 111-545, 124-463, 204-791, 225-791, 231-528, 231-571, 231-585, 231-642, 231-644, 231-647, 239-759, 279-568, 319-759, 358-759, 484-630, 561-989, 687-992, 718-992.

Table 4

Polymethylene Seq ID NO./ Protein ID/ Sequence Length	Sequence Fragments
29 cont.	1326-1605, 1351-1610, 1333-1562, 1333-1600, 1344-1665, 1344-1674, 1346-1571, 1346-1587, 1347-1634, 1361-1648, 1384-1603, 1387-1614, 1387-1646, 1387-1922, 1394-1678, 1403-1617, 1406-1660, 1407-1683, 1410-1687, 1411-1683, 1422-1701, 1423-1763, 1434-1702, 1439-1634, 1439-1649, 1439-1697, 1439-1712, 1440-1708, 1442-1597, 1442-1688, 1444-1736, 1445-1683, 1445-1701, 1445-1708, 1445-1712, 1446-1623, 1446-1731, 1449-1699, 1449-1733, 1485-1789, 1488-1751, 1491-2066, 1498-1712, 1505-1800, 1507-1717, 1508-1747, 1508-1837, 1515-1750, 1515-1970, 1530-1776, 1531-1677, 1531-1827, 1546-2178, 1554-1756, 1554-1786, 1554-1926, 1612-1807, 1620-1838, 1630-1858, 1631-1899, 1652-2031, 1665-1899, 1669-1919, 1673-2147, 1680-1861, 1680-1876, 1680-2212, 1700-1932, 1706-1967, 1708-1994, 1719-1933, 1719-2276, 1719-2292, 1733-2014, 1737-2006, 1745-2004, 1755-1999, 1756-2377, 1761-2254, 1762-2056, 1765-2304, 1765-2042, 1777-2024, 1781-2076, 1782-2090, 1782-2104, 1790-2050, 1798-2063, 1801-2021, 1810-2092, 1810-2362, 1825-2080, 1835-2116, 1837-1857, 1837-1884, 1837-1886, 1837-1896, 1837-1900, 1837-1906, 1837-1912, 1837-1914, 1838-1908, 1840-1905, 1851-1912, 1852-2081, 1854-1912, 1854-2091, 1854-2108, 1863-2008, 1869-2399, 1869-2411, 1873-2138, 1874-1912, 1884-2196, 1885-2159, 1889-1912, 1896-2057, 1896-2114, 1896-2144, 1896-2167, 1896-2169, 1896-2424, 1902-2215, 1910-2180, 1931-2180, 1932-2208, 1934-2397, 1939-2021, 1940-2206, 1940-2323, 1946-2190, 1947-1971, 1947-1976, 1947-2007, 1947-2020, 1947-2022, 1951-2020, 1961-2435, 1962-2020, 1964-2195, 1967-2301, 1967-2423, 1981-2020, 1986-2020, 1987-2435, 1992-2242, 1993-2443, 1994-2026, 1997-2020, 2008-2279, 2011-2433, 2011-2435, 2019-2435, 2029-2286, 2029-2441, 2046-2306, 2046-2435, 2056-2299, 2058-2307, 2058-2436, 2061-2317, 2062-2337, 2062-2335, 2062-2435, 2074-2435, 2075-2435, 2077-2436, 2078-2435, 2079-2322, 2083-2436, 2083-2437, 2085-2443, 2091-2443, 2092-2430, 2100-2292, 2105-2362, 2105-2366, 2106-2365, 2111-2422, 2111-2438, 2119-2360, 2129-2435, 2131-2441, 2132-2440, 2141-2401, 2146-2443, 2149-2399, 2155-2376, 2157-2443, 2159-2435, 2159-2438, 2171-2420, 2171-2443, 2180-2435, 2187-2413, 2188-2394, 2189-2435, 2198-2443, 2208-2443, 2215-2435, 2228-2436, 2231-2443, 2235-2393, 2240-2443, 2251-2442, 2269-2443, 2273-2443, 2301-2443, 2302-2443
29 cont.	1782-2104, 1790-2050, 1798-2063, 1801-2021, 1810-2092, 1810-2362, 1825-2080, 1835-2116, 1837-1857, 1837-1884, 1837-1886, 1837-1896, 1837-1900, 1837-1906, 1837-1912, 1837-1914, 1838-1908, 1840-1905, 1851-1912, 1852-2081, 1854-1912, 1854-2091, 1854-2108, 1863-2008, 1869-2399, 1869-2411, 1873-2138, 1874-1912, 1884-2196, 1885-2159, 1889-1912, 1896-2057, 1896-2114, 1896-2144, 1896-2167, 1896-2169, 1896-2424, 1902-2215, 1910-2180, 1931-2180, 1932-2208, 1934-2397, 1939-2021, 1940-2206, 1940-2323, 1946-2190, 1947-1971, 1947-1976, 1947-2007, 1947-2020, 1947-2022, 1951-2020, 1961-2435, 1962-2020, 1964-2195, 1967-2301, 1967-2423, 1981-2020, 1986-2020, 1987-2435, 1992-2242, 1993-2443, 1994-2026, 1997-2020, 2008-2279, 2011-2433, 2011-2435, 2019-2435, 2029-2286, 2029-2441, 2046-2306, 2046-2435, 2056-2299, 2058-2307, 2058-2436, 2061-2317, 2062-2337, 2062-2335, 2062-2435, 2074-2435, 2075-2435, 2077-2436, 2078-2435, 2079-2322, 2083-2436, 2083-2437, 2085-2443, 2091-2443, 2092-2430, 2100-2292, 2105-2362, 2105-2366, 2106-2365, 2111-2422, 2111-2438, 2119-2360, 2129-2435, 2131-2441, 2132-2440, 2141-2401, 2146-2443, 2149-2399, 2155-2376, 2157-2443, 2159-2435, 2159-2438, 2171-2420, 2171-2443, 2180-2435, 2187-2413, 2188-2394, 2189-2435, 2198-2443, 2208-2443, 2215-2435, 2228-2436, 2231-2443, 2235-2393, 2240-2443, 2251-2442, 2269-2443, 2273-2443, 2301-2443, 2302-2443

WO 02/079441

PCT/US02/09820

Table 4

Polynucleotide SEQ ID NO./ Inverse ID/ Sequence Length	Sequence Fragments
3093807190CBI/ 957	1-226, 1-288, 1-379, 1-380, 1-518, 1-537, 1-555, 1-573, 1-577, 1-609, 1-621, 1-632, 1-691, 1-708, 1-709, 1-710, 1-776, 1-792, 2-694, 3-541, 12-697, 20-503, 164-423, 167-799, 195-769, 215-923, 266-949, 314-640, 330-585, 333-948, 330-888, 380-954, 408-957, 411-957, 439-893, 440-952, 457-957, 475-957, 480-510, 480-957, 303-957, 523-955, 534-655, 534-662, 537-662, 539-662, 551-957, 553-662, 615-655, 615-662, 726-952, 729-957, 823-950, 825-872, 825-934, 825-950, 825-951, 831-951, 832-862, 832-951, 834-860, 834-867, 834-872, 836-865, 836-951, 844-872, 847-951, 858-935, 885-950, 887-941, 888-951, 914-950, 914-951, 915-940, 915-950
31/4856078CBI/ 674	1-345, 28-559, 84-523, 305-559, 315-568, 315-674
32/61106886CBI/ 559	1-157, 1-264, 1-379, 66-559, 162-402
33/5627037CBI/ 2600	1-104, 1-469, 31-705, 37-650, 245-520, 300-963, 355-983, 456-1000, 604-882, 620-1000, 777-936, 850-1506, 939-1007, 941-1242, 942-1067, 944-1067, 945-1067, 1066-1280, 1066-1419, 1067-1242, 1180-1831, 1345-1541, 1460-2245, 1497-2080, 1618-2263, 1634-2263, 1676-2263, 1827-2149, 1827-2570, 1829-2538, 1964-2136, 1968-2600, 1982-2485, 1984-2198, 2396-2580
34/3688835CBI/ 788	1-278, 1-526, 165-455, 166-739, 166-788, 179-775, 185-775, 187-782, 576-754
35/2295419CBI/ 702	1-274, 1-597, 1-608, 8-248, 118-639, 188-640, 261-670, 272-411, 272-678, 272-702
36/2457896CBI/ 806	1-549, 1-640, 275-806, 379-651
37/2641482CBI/ 1112	1-214, 34-286, 105-342, 109-701, 128-399, 137-382, 161-412, 296-456, 379-1086, 456-526, 456-578, 456-646, 456-970, 456-1086, 477-578, 485-714, 486-569, 520-781, 566-1090, 677-801, 691-1105, 732-1024, 766-1033, 964-1112
38/494292CBI/ 2401	1-455, 2-455, 16-2401, 39-531, 72-591, 124-649, 124-676, 124-811, 151-647, 151-776, 324-1133, 430-1266, 513-821, 535-734, 507-1464, 666-1226, 720-585, 731-1071, 731-1457, 753-1474, 766-1725, 785-1562, 789-1320, 802-1652, 810-1543, 822-1337, 822-1342, 823-1749, 832-1449, 850-1513, 865-1603, 1790-2059, 2104-2396, 2108-2401

Table 4

Poly nucleotide SEQ ID NO/ Incyte ID/ Sequence Length	Sequence Fragments
397/63628/CB1/ 1575	1-512, 267-703, 341-990, 357-735, 379-858, 422-858, 428-735, 492-1061, 518-942, 769-1518, 769-1575
40/809539/CB1/ 998	1-625, 44-625, 56-330, 58-333, 92-650, 200-686, 254-549, 300-696, 341-578, 396-998
41/532018/CB1/ 1697	1-92, 1-310, 16-256, 16-333, 18-534, 21-351, 21-366, 21-461, 25-182, 25-557, 27-264, 27-286, 27-328, 27-339, 27-341, 27-485, 37-371, 37-560, 45-441, 127-440, 293-967, 293-976, 293-1043, 295-483, 298-917, 372-862, 390-712, 475-837, 475-902, 648-832, 706-1051, 752-1359, 986-1653, 993-1564, 1037-1638, 1092-1220, 1109-1248, 1155-1681, 1137-1393, 1139-1651, 1169-1467, 1198-1679, 1227-1598, 1274-1654, 1230-1654, 1231-1654, 1232-1654, 1233-1654, 1234-1654, 1237-1650, 1243-1654, 1244-1654, 1246-1654, 1249-1485, 1249-1543, 1254-1418, 1255-1540, 1256-1503, 1256-1646, 1259-1503, 1259-1654, 1260-1654, 1261-1491, 1261-1654, 1262-1464, 1263-1624, 1266-1654, 1267-1675, 1268-1505, 1269-1644, 1275-1416, 1275-1502, 1275-1654, 1275-1688, 1278-1538, 1280-1654, 1284-1654, 1296-1654, 1298-1551, 1299-1538, 1300-1554, 1300-1654, 1303-1691, 1303-1695, 1305-1654, 1306-1676, 1307-1371, 1310-1694, 1331-1571, 1338-1654, 1343-1505, 1352-1575, 1361-1691, 1369-1603, 1369-1609, 1369-1678, 1375-1654, 1376-1685, 1379-1646, 1380-1654, 1389-1688, 1397-1570, 1401-1685, 1408-1654, 1422-1685, 1443-1691, 1481-1643, 1481-1675, 1488-1683, 1532-1692, 1577-1697, 1581-1654, 1617-1683
42/7481063/CB1/ 592	1-91, 1-393, 31-183, 92-393, 153-506, 153-507, 153-508, 153-523, 153-529, 153-532, 153-566, 153-592
43/2044436/CB1/ 1031	1-642, 19-259, 24-277, 24-487, 69-270, 69-379, 69-380, 70-374, 70-379, 71-171, 71-176, 189-677, 256-388, 313-483, 365-487, 533-957, 556-1012, 556-1025, 562-596, 576-1021, 580-1020, 583-1005, 588-1031, 593-1013, 593-1031, 602-877, 611-1023, 618-1031, 642-1022, 658-1031, 662-1004, 680-1023, 745-1011, 767-1021, 802-1022, 861-1066, 940-1031
44/4091564/CB1/ 916	1-306, 1-566, 1-635, 217-907, 228-916
45/8039739/CB1/ 2577	1-1789, 1-1817, 32-1110, 827-918, 827-1037, 827-1197, 895-1151, 994-1797, 1075-1729, 1112-1364, 1112-1365, 1148-1884, 1258-1365, 1289-1328, 1298-1365, 1383-2115, 1391-1476, 1391-1644, 1447-1644, 1539-2026, 1577-2081, 1883-2303, 1889-2158, 1890-2299, 2027-2301, 2044-2577, 2165-2533

WO 02/079441

PCT/US02/09820

Table 4

Polynucleotide SEQ ID NO./ Inverte ID/ Sequence Length	Sequence Fragments
467263837(CBI/ 2999	1-577, 287-810, 317-548, 345-976, 573-826, 573-837, 575-835, 678-1226, 679-922, 680-900, 829-1081, 839-1093, 859-1342, 889-1168, 955-1212, 1212-1476, 1212-1697, 1283-1551, 1327-1588, 1333-1472, 1436-1706, 1466-1624, 1482-1751, 1482-2033, 1512-1714, 1536-1769, 1547-1752, 1610-1925, 1631-1882, 1637-1907, 1663-1946, 1785- 2064, 1792-2061, 1817-2078, 1836-2073, 1844-2065, 1844-2032, 1857-2099, 1901-2064, 1944-2171, 1964-2197, 1990-2235, 1990-2572, 1990-2580, 2009-2273, 2083-2309, 2112-2411, 2119-2346, 2119-2663, 2127-2460, 2161- 2775, 2168-2765, 2193-2418, 2243-2484, 2275-2548, 2275-2775, 2341-2615, 2350-2561, 2493-2711, 2493-2999, 2494-2921, 2575-2853, 2576-2811
47556827(CBI/ 1314	1-42, 1-236, 1-255, 1-346, 1-394, 1-423, 1-427, 1-442, 1-444, 1-497, 4-150, 6-160, 6-339, 13-162, 124-431, 217-509, 245-508, 259-441, 259-506, 329-437, 332-593, 345-609, 351-518, 384-624, 384-883, 384-990, 393-669, 400-879, 412-630, 412-739, 427-613, 450-762, 450-1044, 470-744, 478-564, 526-762, 564-681, 564-726, 564-1027, 569-783, 569-921, 571-716, 603-896, 609-905, 640-1294, 646-866, 648-1287, 651-1286, 657-1293, 658-1303, 664-1296, 670- 883, 687-1286, 747-1271, 761-1299, 764-1013, 771-1056, 775-983, 775-1296, 778-1314, 783-1076, 784-1023, 792- 1057, 822-1077, 847-998, 862-1108, 897-1314, 905-1118, 905-1296, 905-1314, 944-1259, 945-1130, 945-1131, 995- 1305, 999-1314, 1001-1305, 1023-1314, 1046-1314, 1073-1290, 1124-1314, 1157-1314
487503641(CBI/ 2619	1-2348, 1-49-507, 153-418, 166-486, 166-515, 169-506, 175-515, 178-506, 179-414, 184-497, 184-515, 187-270, 188- 479, 210-453, 248-475, 279-506, 604-1048, 653-1297, 656-948, 658-1309, 668-1363, 680-949, 727-1247, 751-1039, 751-1390, 760-1230, 780-990, 785-1247, 798-943, 802-1065, 802-1361, 802-1379, 802-1384, 806-1178, 806-1371, 816-1017, 832-1059, 836-1041, 836-1048, 843-1067, 843-1282, 843-1417, 843-1566, 850-1366, 854-1334, 855- 1154, 855-1323, 889-1454, 917-1112, 941-1205, 949-1394, 962-1233, 970-1616, 978-1248, 978-1678, 982-1355, 987-1289, 991-1578, 993-1229, 994-1575, 996-1289, 997-1586, 997-1659, 1029-1271, 1041-1539, 1041-1559, 1055- 1321, 1038-1784, 1079-1639, 1084-1366, 1085-1337, 1090-1335, 1090-1372, 1092-1352, 1094-1341, 1094-1354, 1100-1343, 1101-1331, 1102-1302, 1102-1340, 1102-1366, 1102-1507, 1102-1577, 1106-1330, 1106-1371, 1107- 1660, 1107-1690, 1110-1391, 1110-1684, 1127-1413, 1131-1330, 1132-1488, 1133-1485, 1133-1586, 1132-1408, 1156-1681, 1166-1377, 1167-1417, 1172-1467, 1174-1453, 1179-1337, 1179-1401

WO 02/079441

PCT/US02/09820

Table 4

Polynucleotide SEQ ID NO./ In-cyte ID/ Sequence Length	Sequence Fragments
49/7503458CB1/ 1370	2085-2340, 2093-2318, 2093-2355, 2094-2340, 2095-2386, 2096-2227, 2103-2366, 2113-2369, 2115-2300, 2122-2334, 2129-2373, 2133-2390, 2136-2331, 2136-2363, 2149-2341, 2154-2355, 2156-2383, 2179-2394, 2192-2232, 2197-2383, 2206-2423, 2209-2324, 2223-2361, 2232-2386, 2238-2354, 2239-2357, 2240-2334, 2247-2373, 2252-2406
50/7500925CB1/ 1201	1-1097, 1-1106, 19-463, 24-130, 361-480, 388-558, 406-561, 504-1097, 608-739, 631-740, 631-1087, 631-1100, 651-1096, 655-1095, 656-1056, 658-1080, 663-1108, 663-1116, 668-1088, 668-1113, 677-952, 686-1098, 693-1116, 704-799, 711-877, 730-997, 733-1116, 737-1114, 753-1098, 758-1111, 842-1096, 847-1091, 877-1097, 936-1081, 946-1107, 1015-1370
50/7500925CB1/ 1201	1-537, 11-467, 11-475, 22-190, 22-283, 22-295, 22-299, 22-400, 22-438, 22-468, 22-475, 22-478, 22-484, 23-386, 23-482, 23-484, 26-379, 27-276, 27-434, 38-468, 40-684, 41-1112, 69-355, 140-651, 164-471, 165-744, 257-549, 280-469, 299-481, 299-546, 372-633, 385-649, 389-693, 391-558, 424-694, 433-689, 453-670, 453-772, 453-786, 467-653, 490-626, 490-778, 510-782, 610-692, 611-756, 734-1020, 734-1065, 801-1111, 807-1111, 830-1111, 833-1201, 842-1113, 990-1111

WO 02/079441

PCT/US02/09820

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
26	7757335CB1	LIVRTUE01
27	1482539CB1	COLNNON05
28	2234213CB1	PANCTUT02
29	1345785CB1	LATRTUT02
30	3807190CB1	CONTTUT01
31	4856078CB1	BRSTTUT22
32	6106886CB1	MCLRNOC01
33	5627037CB1	BONEUNR01
34	3688835CB1	THYRNOT10
35	2295419CB1	KIDNNOT32
36	2437896CB1	BRAFDT02
37	2641482CB1	LUNGTUT08
38	7494292CB1	BRAIFER06
39	7636288CB1	SINTDIE01
40	8095391CB1	BRAINYO2
41	532018CB1	LIVRTUT04
43	2044436CB1	PTHYNOT03
44	4091564CB1	BSCNSZT01
45	8039739CB1	SPLNNOE01
46	1265837CB1	ENDANOT01
47	5568527CB1	BRAITUT01
48	7503641CB1	FIBRUNT02
49	7503458CB1	BRANDIT04
50	7500925CB1	MONOTXN05

WO 02/079441

PCT/US02/09820

Table 6

Library	Vector	Library Description
BONEUNR01	PCDNΔ2.1	This random primed library was constructed using pooled cDNA from two different donors. cDNA was generated using mRNA isolated from an untreated MG-63 cell line derived from an osteosarcoma tumor removed from a 14-year-old Caucasian male (donor A) and using mRNA isolated from satal bone tumor tissue removed from an 18-year-old Caucasian female (donor B) during an exploratory laparotomy and soft tissue excision. Pathology indicated giant cell tumor of the sacrum in donor B. Donor B's history included pelvic joint pain, constipation, urinary incontinence, unspecified abdominal/pelvic symptoms, and a pelvic soft tissue malignant neoplasm. Family history included prostate cancer in donor B.
BRAJLIT02	pNCY	Library was constructed using RNA isolated from frontal cortex tissue removed from a 77-year-old male. Patient history included carcinoma of the lung.
BRAJFER06	PCDNΔ2.1	This random primed 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. Serologies were negative.
BRAINYO2	pNCY	This large size-fractionated and normalized library was constructed using pooled cDNA generated using mRNA isolated from midbrain, inferior temporal cortex, medulla, and posterior parietal cortex tissues 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. 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. 0.28 million independent clones from this size-selected library were normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9238-9232 and Romaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRAJLIT01	PSPORT1	Library was constructed using RNA isolated from brain tumor tissue removed from a 50-year-old Caucasian female during a frontal lobectomy. Pathology indicated recurrent grade 3 oligoastrocytoma with focal necrosis and extensive calcification. Patient history included a speech disturbance and epilepsy. The patient's brain had also been irradiated with a total dose of 5,082 cGy (Fraction 8). Family history included a brain tumor.

WO 02/079441

PCT/US02/09820

Table 6

Library	Vector	Library Description
BRANDIT04	pINCY	Library was constructed using RNA isolated from pineal gland tissue removed from a 66-year-old Caucasian female who died from congestive heart failure. Neuropathology indicated mild to moderate Alzheimer disease, atherosclerosis, and multiple infarctions. Microscopically, there were diffuse and neuritic amyloid plaques throughout the cerebral cortex. There were neurofibrillary tangles in the temporal lobes particularly the entorhinal cortex. The frontal cortex contained scattered, ballooned neurons. The amygdala contained marked gliosis, neuritic plaques and intracellular neurofibrillary tangles. The hippocampus contained neuritic and diffuse plaques, and neurofibrillary tangles. The thalamus contained diffuse and focal neuritic amyloid plaques and scattered neurofibrillary tangles. There was area of cystic cavitation with surrounding gliosis in the left globus pallidus. The pallidum contained scattered intracellular neurofibrillary tangles. The caudate, putamen and nucleus accumbens contained diffuse plaques.
BRANDIT04 cont.		There was an area of cystic cavitation with lipid-laden macrophages in the right cerebellar hemisphere. Patient history included diabetes mellitus, rheumatoid arthritis, hyperthyroidism, amyloid heart disease, and dementia.
BRSTTUT22	pINCY	Library was constructed using RNA isolated from right breast tumor tissue removed from a 59-year-old Caucasian female during unilateral extended simple mastectomy. Pathology indicated invasive lobular carcinoma with extension into ducts forming an ill-defined mass situated in the right breast biopsy cavity site. The non-neoplastic breast parenchyma displays papillomatosis. Prior right breast biopsy indicated invasive grade 3, nuclear grade 3, in situ ductal carcinoma. Estrogen and progesterone immunostains were positive in the neoplastic cells. Patient history included cirrhosis of the liver, esophageal ulcer, hyperlipidemia, and neuropathy caused by an unspecified disease. Previous surgeries included segmental lung resection and a liver transplant. Patient medications included Prograf, prednisone, Imuran, Lozol, Zimata, Estraderm patches, and Provera.
BSCNSZ701	pINCY	Library was constructed using RNA isolated from diseased caudate nucleus tissue removed from the brain of a 49-year-old male. Patient history included schizophrenia.

WO 02/079441

PCT/US02/09820

Table 6

Library	Vector	Library Description
COLNON03	pNCY	This normalized colon and colon epithelial layer tissue library was constructed from 7 million independent clones from a colon and colon epithelial layer tissue library. Starting RNA was made from pooled cDNA from two different donors. cDNA was generated using mRNA isolated from colon tissue removed from a 16-year-old Black male (donor A) who died from a motor vehicle accident and from colon epithelium tissue removed from a 13-year-old Caucasian female (donor B) who died from a motor vehicle accident. For donor A, serology was positive for cytomegalovirus and remaining serologies were negative. Patient history included occasional alcohol (beer) use. The patient was treated with chest tube insertion post motor vehicle accident. The patient was not taking any medications. The library was normalized in two rounds using conditions adapted from Soares et al., <i>PNAS</i> (1994) 91:9228-9232 and Bonaldo et al., <i>Genome Research</i> (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
CONTTUT01	pNCY	Library was constructed using RNA isolated from tumorous soft tissue of the left lateral thigh removed from a 34-year-old Caucasian female during a soft tissue excision. Pathology indicated metastatic grade 2 myxoid liposarcoma which formed multiple, lobulated, circumscribed masses situated in the subcutaneous adipose tissue. Patient history included a malignant soft tissue neoplasm of the leg. Family history included benign hypertension, acute leukemia, benign hypertension, and type II diabetes.
ENDANOT01	PBLUESCRIPT	Library was constructed using RNA isolated from aortic endothelial cell tissue from an explanted heart removed from a male during a heart transplant.
FIBRUNT02	pNCY	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.
KIDNNOT32	pNCY	Library was constructed using RNA isolated from kidney tissue removed from a 49-year-old Caucasian male who died from an intracranial hemorrhage and cerebrovascular accident. Patient history included tobacco abuse.
LATRUT02	pNCY	Library was constructed using RNA isolated from a myxoma removed from the left atrium of a 43-year-old Caucasian male during amnioplasty. Pathology indicated atrial myxoma. Patient history included pulmonary insufficiency, acute myocardial infarction, atherosclerotic coronary artery disease, hyperlipidemia, and tobacco use. Family history included benign hypertension, acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.

WO 02/079441

PCT/US02/09820

Table 6

Library	Vector	Library Description
LIVRTUE01	PCDNA2.1	This 5' hissed random primed library was constructed using RNA isolated from liver tumor tissue removed from a 72-year-old Caucasian male during partial hepatectomy. Pathology indicated metastatic grade 2 (of 4) neuroendocrine carcinoma forming a mass. The patient presented with metastatic liver cancer. Patient history included benign hypertension, type I diabetes, prostate hyperplasia, prostate cancer, alcohol abuse in remission, and tobacco abuse in remission. Previous surgeries included destruction of a pancreatic lesion, closed prostate biopsy, transurethral prostatectomy, removal of bilateral testes and total splenectomy. Patient medications included Euflexin, Hytrin, Proscar, and insulin. Family history included atherosclerotic coronary artery disease and acute myocardial infarction in the mother; atherosclerotic coronary artery disease and type II diabetes in the father.
LIVRTUT04	pINCY	Library was constructed using RNA isolated from liver tumor tissue removed from a 50-year-old Caucasian male during a partial hepatectomy. Pathology indicated a grade 3-4 hepatoma, forming a mass. Patient history included benign hypertension and hepatitis B core antigen and hepatitis B surface antigen was present in the patient.
LUNGTUT08	pINCY	Library was constructed using RNA isolated from lung tumor tissue removed from a 63-year-old Caucasian male during a right upper lobectomy with fibrotic bronchospasm. Pathology indicated a grade 3 adenocarcinoma. Patient history included atherosclerotic coronary artery disease, an acute myocardial infarction, rectal cancer, an asymptomatic abdominal aortic aneurysm, tobacco abuse, and cardiac dysrhythmia. Family history included congestive heart failure, stomach cancer, and lung cancer, type II diabetes, atherosclerotic coronary artery disease, and an acute myocardial infarction.
MCLRNOC01	pINCY	This large size-fractionated library was constructed using RNA isolated from mononuclear cells obtained from the umbilical cord blood of multiple individuals of mixed age and sex. The cells were treated with G-CSF.
MONOTXN05	pINCY	This normalized treated monocyte cell tissue library was constructed from 1.03 million independent clones from a monocyte tissue library. Starting RNA was made from RNA isolated from treated monocytes from peripheral blood removed from a 42-year-old female. The cells were treated with interleukin-10 (IL-10) and lipopolysaccharide (LPS). The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9728-9732 and Bonaldo et al., Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used.

WO 02/079441

PCT/US02/09820

Table 6

Library	Vector	Library Description
PANGTUT02	pINCY	Library was constructed using RNA isolated from pancreatic tumor tissue removed from a 45-year-old Caucasian female during radical pancreaticoduodenectomy. Pathology indicated a grade 4 anaplastic carcinoma. Family history included benign hypertension, hyperlipidemia and atherosclerotic coronary artery disease.
PTHYNOT03	pINCY	Library was constructed using RNA isolated from the left parathyroid tissue of a 69-year-old Caucasian female during a partial parathyroidectomy. Pathology indicated hyperplasia. The patient presented with primary hyperparathyroidism.
SINTIDE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from small intestine tissue removed from a 49-year-old Caucasian female during gastroenterostomy, exploratory laparotomy, and vagotomy. The patient presented with acute stomach ulcer with obstruction, nausea and vomiting, and abnormal weight loss. Patient history included backache, abdominal hysterectomy, Patient medications included Premarin. Family history included benign hypertension, type II diabetes and congestive heart failure in the father.
SPLNNOE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from the spleen tissue of a 2-year-old Hispanic male, who died from cerebral anoxia. Past medical history and serologies were negative.
THYRNOT10	pINCY	Library was constructed using RNA isolated from the diseased left thyroid tissue removed from a 30-year-old Caucasian female during a unilateral thyroid lobectomy and parathyroid reimplantation. Pathology indicated lymphocytic thyroiditis.

WO 02/079441

PCT/US02/09820

Table 7

Program	Description	Reference	Parameter/Threshold
ABIFACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value=1.0E-8 or less; Full Length sequences: Probability value=1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.00E-6; Assembled ESTs: fasta Identity=length-200 bases or greater; fastx E values=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 PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Atwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value=1.0E-3 or less

WO 02/079441

PCT/US02/09820

Table 7

Program	Description	Reference	Parameter/Threshold
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; Sonnhammer, E.L.J. et al. (1988) <i>Nucleic Acids Res.</i> 26:520-522; Durbin, K. et al. (1998) <i>Our World View</i> , in a Nusshel, Cambridge Univ. Press, pp. 1-350.	PFAM, SMART or TIGRFAM hits: Probability values= 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.	Gribskov, M. et al. (1988) <i>CABIOS</i> 4:61-66; Gribskov, 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>CCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) <i>Genome Res.</i> 8:175-185; Ewing, B. and P. Green (1998) <i>Genome Res.</i> 8:186-194.	
Phrap	A Phis Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) <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
Coned	A graphical tool for viewing and editing Phrap assemblies.	Gordin, D. et al. (1998) <i>Genome Res.</i> 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nelson, 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 determine 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.	

Table 7

Program	Description	Reference	Parameter Threshold
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Somhlammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. On Intelligent Systems for Mol. Biol., Glasgow et al. eds., The Am. Assoc. for Artificial Intelligence (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 02/079441

PCT/US02/09820

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence selected from the group consisting of
5 SEQ ID NO:1-25,
 - 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:2-24,
 - c) a polypeptide comprising a naturally occurring amino acid sequence at least 96% identical
to the amino acid sequence consisting of SEQ ID NO:1,
 - 10 d) a polypeptide comprising a naturally occurring amino acid sequence at least 95% identical
to the amino acid sequence consisting of SEQ ID NO:25,
 - e) a biologically active fragment of a polypeptide having an amino acid sequence selected
from the group consisting of SEQ ID NO:1-25, and
 - 15 f) an immunogenic fragment of a polypeptide having an amino acid sequence selected from
the group consisting of SEQ ID NO:1-25.
2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the
group consisting of SEQ ID NO:1-25.
- 20 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
25 the group consisting of SEQ ID NO:26-50.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a
polynucleotide of claim 3.
- 30 7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

WO 02/079441

PCT/US02/09820

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
 - 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:1-25.
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:26-50,
 - 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:27-50,
 - c) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 96% identical to the polynucleotide sequence of SEQ ID NO:26,
 - d) a polynucleotide complementary to a polynucleotide of a),
 - e) a polynucleotide complementary to a polynucleotide of b),
 - e) a polynucleotide complementary to a polynucleotide of c), and
 - f) an RNA equivalent of a)-e).
13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.
14. A method of detecting a target polynucleotide in a sample, said target polynucleotide 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

WO 02/079441

PCT/US02/09820

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.

5

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:

- 10 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.

15 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:1-25.

20

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.

25 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.

30 21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.

22. A method for treating a disease or condition associated with decreased expression of

WO 02/079441

PCT/US02/09820

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 02/079441

PCT/US02/09820

28. A method of screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:
- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
 - b) detecting altered expression of the target polynucleotide, and
 - c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.
29. A method of assessing toxicity of a test compound, the method comprising:
- 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.
30. A diagnostic test for a condition or disease associated with the expression of SECP in a biological sample, the method comprising:
- a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and
 - b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.
31. The antibody of claim 11, wherein the antibody is:
- a) a chimeric antibody,
 - b) a single chain antibody,

WO 02/079441

PCT/US02/09820

- c) a Fab fragment,
 - d) a F(ab')₂ fragment, or
 - e) a humanized antibody.
- 5 32. A composition comprising an antibody of claim 11 and an acceptable excipient.
33. A method of diagnosing a condition or disease associated with the expression of SECP in a subject, comprising administering to said subject an effective amount of the composition of claim 32.
- 10 34. A composition of claim 32, wherein the antibody is labeled.
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.
- 15 36. A method of preparing a polyclonal 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 selected from the group consisting of SEQ ID NO:1-25, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
 - 20 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 selected from the group consisting of SEQ ID NO:1-25.
- 25 37. A polyclonal antibody produced by a method of claim 36.
38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.
39. A method of making a monoclonal antibody with the specificity of the antibody of claim 30 11, the method comprising:
- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-25, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

WO 02/079441

PCT/US02/09820

- b) isolating antibody producing cells from the animal,
c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
d) culturing the hybridoma cells, and
5 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:1-25.
40. A monoclonal antibody produced by a method of claim 39.
- 10 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 library.
- 15 43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.
44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-25 in a sample, the method comprising:
20 a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of
25 SEQ ID NO:1-25 in the sample.
45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-25 from a sample, the method comprising:
30 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 NO:1-25.

WO 02/079441

PCT/US02/09820

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, the method comprising:

- 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
- c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

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.

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.

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.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

WO 02/079441

PCT/US02/09820

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.
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.

WO 02/079441

PCT/US02/09820

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.
- 5 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.
- 10 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.
- 15 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.
- 20 80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.
81. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:26.
- 25 82. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:27.
83. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:28.
84. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:29.
- 30 85. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:30.
86. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:31.

WO 02/079441

PCT/US02/09820

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.
- 5 89. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:34.
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.
- 10 92. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:37.
93. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:38.
- 15 94. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:39.
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.
- 20 97. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:42.
98. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:43.
- 25 99. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:44.
100. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:45.
- 30 101. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:46.
102. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID

WO 02/079441

PCT/US02/09820

NO:47.

103. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:48.

5

104. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:49.

105. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
10 NO:50.

WO 02/079441

PCT/US02/09820

<110> INCYTE GENOMICS, INC.

BAUGHN, Mariah R.
BURFORD, Neil
DING, Li
DUGGAN, Brendan M.
ELLIOTT, Vicki S.
FORSYTHE, Ian J.
GANDHI, Ameena, R.
GIETZEN, Kimberly, J.
GRIFFIN, Jennifer A.
HE, Ann
HONCHELL, Cynthia D.
ISON, Craig H.
LAL, Preeti G.
LEE, Ernestine A.
LEE, Sally
LU, Dyung Aina M.
MASON, Patricia M.
SAMJANWALA, Madhu M.
SWARNKAR, Anita
RAMKUMAR, Jayalaxmi
TANG, Y. Tom
THANGAVELU, Kavitha
TRAN, Uyen K.
WALLA, Narinder K.
WARREN, Bridget A.
YAO, Monique G.
YUE, Henry
XU, Yuming

<120> SECRETED PROTEINS

<130> PI-0401 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/280,527; 60/282,112; 60/282,702; 60/283,855;
60/343,718; 60/339,236; 60/357,002<151> 2001-03-30; 2001-04-06; 2001-04-09; 2001-04-13;
2001-10-19; 2001-12-07; 2002-02-13

<160> 50

<170> PERL Program

<210> 1

<211> 2487

<212> FKT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7757335CD1

<400> 1

Met Val Leu Leu Leu Cys Leu Ser Cys Leu Ile Phe Ser Cys Leu

WO 02/079441

PCT/US02/09820

1	5	10	15
Thr Phe Ser Trp	Leu Lys Ile Trp Gly	Lys Met Thr Asp Ser Lys	
	20	25	30
Pro Ile Thr Lys	Ser Lys Ser Glu Ala	Asn Leu Ile Pro Ser Gln	
	35	40	45
Glu Pro Phe Pro	Ala Ser Asp Asn Ser	Gly Glu Thr Pro Gln Arg	
	50	55	60
Asn Gly Glu Gly	His Thr Leu Pro Lys	Thr Pro Ser Gln Ala Glu	
	65	70	75
Pro Ala Ser His	Lys Gly Pro Lys Asp	Ala Gly Arg Arg Arg Asn	
	80	85	90
Ser Leu Pro Pro	Ser His Gln Lys Pro	Pro Arg Asn Pro Leu Ser	
	95	100	105
Ser Ser Asp Ala	Ala Pro Ser Pro Glu	Leu Gln Ala Asn Gly Thr	
	110	115	120
Gly Thr Gln Gly	Leu Glu Ala Thr Asp	Thr Asn Gly Leu Ser Ser	
	125	130	135
Ser Ala Arg Pro	Gln Gly Gln Gln Ala	Gly Ser Pro Ser Lys Glu	
	140	145	150
Asp Lys Lys Gln	Ala Asn Ile Lys Arg	Gln Leu Met Thr Asn Phe	
	155	160	165
Ile Leu Gly Ser	Phe Asp Asp Tyr Ser	Ser Asp Glu Asp Ser Val	
	170	175	180
Ala Gly Ser Ser	Arg Glu Ser Thr Arg	Lys Gly Ser Arg Ala Ser	
	185	190	195
Leu Gly Ala Leu	Ser Leu Glu Ala Tyr	Leu Thr Thr Gly Glu Ala	
	200	205	210
Glu Thr Arg Val	Pro Thr Met Arg Pro	Ser Met Ser Gly Leu His	
	215	220	225
Leu Val Lys Arg	Gly Arg Glu His Lys	Lys Leu Asp Leu His Arg	
	230	235	240
Asp Phe Thr Val	Ala Ser Pro Ala Glu	Phe Val Thr Arg Phe Gly	
	245	250	255
Gly Asp Arg Val	Ile Glu Lys Val Leu	Ile Ala Asn Asn Gly Ile	
	260	265	270
Ala Ala Val Lys	Cys Met Arg Ser Ile	Arg Arg Trp Ala Tyr Glu	
	275	280	285
Met Phe Arg Asn	Glu Arg Ala Ile Arg	Phe Val Val Met Val Thr	
	290	295	300
Pro Glu Asp Leu	Lys Ala Asn Ala Glu	Tyr Ile Lys Met Ala Asp	
	305	310	315
His Tyr Val Pro	Val Pro Gly Gly Pro	Asn Asn Asn Asn Tyr Ala	
	320	325	330
Asn Val Glu Leu	Ile Val Asp Ile Ala	Lys Arg Ile Pro Val Gln	
	335	340	345
Ala Val Trp Ala	Gly Trp Gly His Ala	Ser Glu Asn Pro Lys Leu	
	350	355	360
Pro Glu Leu Leu	Cys Lys Asn Gly Val	Ala Phe Leu Gly Pro Pro	
	365	370	375
Ser Glu Ala Met	Trp Ala Leu Gly Asp	Lys Ile Ala Ser Thr Val	
	380	385	390
Val Ala Gln Thr	Leu Gln Val Pro Thr	Leu Pro Trp Ser Gly Ser	
	395	400	405
Gly Leu Thr Val	Glu Trp Thr Glu Asp	Asp Leu Gln Gln Gly Lys	
	410	415	420
Arg Ile Ser Val	Pro Glu Asp Val Tyr	Asp Lys Gly Cys Val Lys	

WO 02/079441

PCT/US02/09820

	425	430	435
Asp Val Asp Glu Gly	Leu Glu Ala Ala	Glu Arg Ile Gly Phe	Pro
440	445	450	
Leu Met Ile Lys Ala	Ser Glu Gly Gly	Gly Gly Lys Gly Ile	Arg
455	460	465	
Lys Ala Glu Ser Ala	Glu Asp Phe Pro	Ile Leu Phe Arg Gln	Val
470	475	480	
Gln Ser Glu Ile	Pro Gly Ser Pro	Ile Phe Leu Met Lys	Leu Ala
485	490	495	
Gln His Ala Arg His	Leu Glu Val Gln	Ile Leu Ala Asp Gln	Tyr
500	505	510	
Gly Asn Ala Val Ser	Leu Phe Gly Arg	Asp Cys Ser Ile Gln	Arg
515	520	525	
Arg His Gln Lys Ile	Val Glu Glu Ala	Pro Ala Thr Ile Ala	Pro
530	535	540	
Leu Ala Ile Phe Glu	Phe Met Glu Gln	Cys Ala Ile Arg Leu	Ala
545	550	555	
Lys Thr Val Gly Tyr	Val Ser Ala Gly	Thr Val Glu Tyr Leu	Tyr
560	565	570	
Ser Gln Asp Gly Ser	Phe His Phe Leu	Glu Leu Asn Pro Arg	Leu
575	580	585	
Gln Val Glu His Pro	Cys Thr Glu Met	Ile Ala Asp Val Asn	Leu
590	595	600	
Pro Ala Ala Gln Leu	Gln Ile Ala Met	Gly Val Pro Leu His	Arg
605	610	615	
Leu Lys Asp Ile Arg	Leu Leu Tyr Gly	Glu Ser Pro Trp Gly	Val
620	625	630	
Thr Pro Ile Ser Phe	Glu Thr Pro Ser	Asn Pro Pro Leu Ala	Arg
635	640	645	
Gly His Val Ile Ala	Ala Arg Ile Thr	Ser Glu Asn Pro Asp	Glu
650	655	660	
Gly Phe Lys Pro Ser	Ser Gly Thr Val	Gln Glu Leu Asn Phe	Arg
665	670	675	
Ser Ser Lys Asn Val	Trp Gly Tyr Phe	Ser Val Ala Ala Thr	Gly
680	685	690	
Gly Leu His Glu Phe	Ala Asp Ser Gln	Phe Gly His Cys Phe	Ser
695	700	705	
Trp Gly Glu Asn Arg	Glu Glu Ala Ile	Ser Asn Met Val Val	Ala
710	715	720	
Leu Lys Glu Leu Ser	Ile Arg Gly Asp	Phe Arg Thr Thr	Val Glu
725	730	735	
Tyr Leu Ile Asn Leu	Leu Glu Thr Glu	Ser Phe Gln Asn Asn	Asp
740	745	750	
Ile Asp Thr Gly Trp	Leu Asp Tyr Leu	Ile Ala Glu Lys Val	Gln
755	760	765	
Ala Glu Lys Pro Asp	Ile Met Leu Gly	Val Val Cys Gly Ala	Leu
770	775	780	
Asn Val Ala Asp Ala	Met Phe Arg Thr	Cys Met Thr Asp	Phe Leu
785	790	795	
His Ser Leu Glu Arg	Gly Gln Val Leu	Pro Ala Asp Ser	Leu Leu
800	805	810	
Asn Leu Val Asp Val	Glu Leu Ile Tyr	Gly Gly Val Lys Tyr	Ile
815	820	825	
Leu Lys Val Ala Arg	Gln Ser Leu Thr	Met Phe Val Leu Ile	Met
830	835	840	
Asn Gly Cys His Ile	Glu Ile Asp Ala	His Arg Leu Asn	Asp Gly

WO 02/079441

PCT/US02/09820

	845	850	855											
Gly	Leu	Leu	Leu	Ser	Tyr	Asn	Gly	Asn	Ser	Tyr	Thr	Thr	Tyr	Met
	860	865	870											
Lys	Glu	Glu	Val	Asp	Ser	Tyr	Arg	Ile	Thr	Ile	Gly	Asn	Lys	Thr
	875	880	885											
Cys	Val	Phe	Glu	Lys	Glu	Asn	Asp	Pro	Thr	Val	Leu	Arg	Ser	Pro
	890	895	900											
Ser	Ala	Gly	Lys	Leu	Thr	Gln	Tyr	Thr	Val	Glu	Asp	Gly	Gly	His
	905	910	915											
Val	Glu	Ala	Gly	Ser	Ser	Tyr	Ala	Glu	Met	Glu	Val	Met	Lys	Met
	920	925	930											
Ile	Met	Thr	Leu	Asn	Val	Gln	Glu	Arg	Gly	Arg	Val	Lys	Tyr	Ile
	935	940	945											
Lys	Arg	Pro	Gly	Ala	Val	Leu	Glu	Ala	Gly	Cys	Val	Val	Ala	Asp
	950	955	960											
Leu	Glu	Leu	Asp	Asp	Pro	Ser	Lys	Val	His	Pro	Ala	Glu	Pro	Phe
	965	970	975											
Thr	Gly	Glu	Leu	Pro	Ala	Gln	Gln	Thr	Leu	Pro	Ile	Leu	Gly	Glu
	980	985	990											
Lys	Leu	His	Gln	Val	Phe	His	Ser	Val	Leu	Glu	Asn	Leu	Thr	Asn
	995	1000	1005											
Val	Met	Ser	Gly	Phe	Cys	Leu	Pro	Glu	Pro	Val	Phe	Ser	Ile	Lys
	1010	1015	1020											
Leu	Lys	Glu	Trp	Val	Gln	Lys	Leu	Met	Met	Thr	Leu	Arg	His	Pro
	1025	1030	1035											
Ser	Leu	Pro	Leu	Leu	Glu	Leu	Gln	Glu	Ile	Met	Thr	Ser	Val	Ala
	1040	1045	1050											
Gly	Arg	Ile	Pro	Ala	Pro	Val	Glu	Lys	Ser	Val	Arg	Arg	Val	Met
	1055	1060	1065											
Ala	Gln	Tyr	Ala	Ser	Asn	Ile	Thr	Ser	Val	Leu	Cys	Gln	Phe	Pro
	1070	1075	1080											
Ser	Gln	Gln	Ile	Ala	Thr	Ile	Leu	Asp	Cys	His	Ala	Ala	Thr	Leu
	1085	1090	1095											
Gln	Arg	Lys	Ala	Asp	Arg	Glu	Val	Phe	Ile	Asn	Thr	Gln	Ser	
	1100	1105	1110											
Ile	Val	Gln	Leu	Val	Gln	Arg	Tyr	Arg	Ser	Gly	Ile	Arg	Gly	Tyr
	1115	1120	1125											
Met	Lys	Thr	Val	Val	Leu	Asp	Leu	Leu	Arg	Arg	Tyr	Leu	Arg	Val
	1130	1135	1140											
Glu	His	His	Phe	Ser	Lys	Ala	Arg	Asp	Ala	Asp	Ala	Asn	Thr	Ser
	1145	1150	1155											
Gly	Met	Val	Gly	Gly	Val	Arg	Ser	Leu	Ser	Phe	Thr	Ser	Val	Trp
	1160	1165	1170											
Cys	Phe	Ser	Pro	Pro	Ala	His	Tyr	Asp	Lys	Cys	Val	Ile	Asn	Leu
	1175	1180	1185											
Arg	Glu	Gln	Phe	Lys	Pro	Asp	Met	Ser	Gln	Val	Leu	Asp	Cys	Ile
	1190	1195	1200											
Phe	Ser	His	Ala	Gln	Val	Ala	Lys	Lys	Asn	Gln	Leu	Val	Ile	Met
	1205	1210	1215											
Leu	Ile	Asp	Glu	Leu	Cys	Gly	Pro	Asp	Pro	Ser	Leu	Ser	Asp	Glu
	1220	1225	1230											
Leu	Ile	Ser	Ile	Leu	Asn	Glu	Leu	Thr	Gln	Leu	Ser	Lys	Ser	Glu
	1235	1240	1245											
His	Cys	Lys	Val	Ala	Leu	Arg	Ala	Arg	Gln	Ile	Leu	Ile	Ala	Ser
	1250	1255	1260											
His	Leu	Pro	Ser	Tyr	Glu	Leu	Arg	His	Asn	Gln	Val	Glu	Ser	Ile

WO 02/079441

PCT/US02/09820

1265 1270 1275
 Phe Leu Ser Ala Ile Asp Met Tyr Gly His Gln Phe Cys Pro Glu
 1280 1285 1290
 Asn Leu Lys Lys Leu Ile Leu Ser Glu Thr Thr Ile Phe Asp Val
 1295 1300 1305
 Leu Pro Thr Phe Phe Tyr His Ala Asn Lys Val Val Cys Met Ala
 1310 1315 1320
 Ser Leu Glu Val Tyr Val Arg Arg Gly Tyr Ile Ala Tyr Glu Leu
 1325 1330 1335
 Asn Ser Leu Gln His Arg Gln Leu Pro Asp Gly Thr Cys Val Val
 1340 1345 1350
 Glu Phe Gln Phe Met Leu Pro Ser Ser His Pro Asn Arg Met Thr
 1355 1360 1365
 Val Pro Ile Ser Ile Thr Asn Pro Asp Leu Leu Arg His Ser Thr
 1370 1375 1380
 Glu Leu Phe Met Asp Ser Gly Phe Ser Pro Leu Cys Gln Arg Met
 1385 1390 1395
 Gly Ala Met Val Ala Phe Arg Arg Phe Glu Asp Phe Thr Arg Asn
 1400 1405 1410
 Phe Asp Glu Val Ile Ser Cys Phe Ala Asn Val Pro Lys Asp Thr
 1415 1420 1425
 Pro Leu Phe Ser Glu Ala Arg Thr Ser Leu Tyr Ser Glu Asp Asp
 1430 1435 1440
 Cys Lys Ser Leu Arg Glu Glu Pro Ile His Ile Leu Asn Val Ser
 1445 1450 1455
 Ile Gln Cys Ala Asp His Leu Glu Asp Glu Ala Leu Val Pro Ile
 1460 1465 1470
 Leu Arg Thr Phe Val Gln Ser Lys Lys Asn Ile Leu Val Asp Tyr
 1475 1480 1485
 Gly Leu Arg Arg Ile Thr Phe Leu Ile Ala Gln Glu Lys Glu Phe
 1490 1495 1500
 Pro Lys Phe Phe Thr Phe Arg Ala Arg Asp Glu Phe Ala Glu Asp
 1505 1510 1515
 Arg Ile Tyr Arg His Leu Glu Pro Ala Leu Ala Phe Gln Leu Glu
 1520 1525 1530
 Leu Asn Arg Met Arg Asn Phe Asp Leu Thr Ala Val Pro Cys Ala
 1535 1540 1545
 Asn His Lys Met His Leu Tyr Leu Gly Ala Ala Lys Val Lys Glu
 1550 1555 1560
 Gly Val Glu Val Thr Asp His Arg Phe Phe Ile Arg Ala Ile Ile
 1565 1570 1575
 Arg His Ser Asp Leu Ile Thr Lys Glu Ala Ser Phe Glu Tyr Leu
 1580 1585 1590
 Gln Asn Glu Gly Glu Arg Leu Leu Leu Glu Ala Met Asp Glu Leu
 1595 1600 1605
 Glu Val Ala Phe Asn Asn Thr Ser Val Arg Thr Asp Cys Asn His
 1610 1615 1620
 Ile Phe Leu Asn Phe Val Pro Thr Val Ile Met Asp Pro Phe Lys
 1625 1630 1635
 Ile Glu Glu Ser Val Arg Tyr Met Val Met Arg Tyr Gly Ser Arg
 1640 1645 1650
 Leu Trp Lys Leu Arg Val Leu Gln Ala Glu Val Lys Ile Asn Ile
 1655 1660 1665
 Arg Gln Thr Thr Thr Gly Ser Ala Val Pro Ile Arg Leu Phe Ile
 1670 1675 1680
 Thr Asn Glu Ser Gly Tyr Tyr Leu Asp Ile Ser Leu Tyr Lys Glu

WO 02/079441

PCT/US02/09820

	1685	1690	1695
Val Thr Asp Ser Arg	Ser Gly Asn Ile Met Phe His Ser Phe Gly		
	1700	1705	1710
Asn Lys Gln Gly Pro	Gln His Gly Met Leu Ile Asn Thr Pro Tyr		
	1715	1720	1725
Val Thr Lys Asp Leu	Leu Gln Ala Lys Arg Phe Gln Ala Gln Thr		
	1730	1735	1740
Leu Gly Thr Thr Tyr	Ile Tyr Asp Phe Pro Glu Met Phe Arg Gln		
	1745	1750	1755
Ala Leu Phe Lys Leu	Trp Gly Ser Pro Asp Lys Tyr Pro Lys Asp		
	1760	1765	1770
Ile Leu Thr Tyr Thr	Glu Leu Val Leu Asp Ser Gln Gly Gln Leu		
	1775	1780	1785
Val Glu Met Asn Arg	Leu Pro Gly Gly Asn Glu Val Gly Met Val		
	1790	1795	1800
Ala Phe Lys Met Arg	Phe Lys Thr Gln Glu Tyr Pro Glu Gly Arg		
	1805	1810	1815
Asp Val Ile Val Ile	Gly Asn Asp Ile Thr Phe Arg Ile Gly Ser		
	1820	1825	1830
Phe Gly Pro Gly Glu	Asp Leu Leu Tyr Leu Arg Ala Ser Glu Met		
	1835	1840	1845
Ala Arg Ala Glu Gly	Ile Pro Lys Ile Tyr Val Ala Ala Asn Ser		
	1850	1855	1860
Gly Ala Arg Ile Gly	Met Ala Glu Glu Ile Lys His Met Phe His		
	1865	1870	1875
Val Ala Trp Val Asp	Pro Glu Asp Pro His Lys Gly Phe Lys Tyr		
	1880	1885	1890
Leu Tyr Leu Thr Pro	Gln Asp Tyr Thr Arg Ile Ser Ser Leu Asn		
	1895	1900	1905
Ser Val His Cys Lys	His Ile Glu Glu Gly Gly Glu Ser Arg Tyr		
	1910	1915	1920
Met Ile Thr Asp Ile	Ile Gly Lys Asp Asp Gly Leu Gly Val Glu		
	1925	1930	1935
Asn Leu Arg Gly Ser	Gly Met Ile Ala Gly Glu Ser Ser Leu Ala		
	1940	1945	1950
Tyr Glu Glu Ile Val	Thr Ile Ser Leu Val Thr Cys Arg Ala Ile		
	1955	1960	1965
Gly Ile Gly Ala Tyr	Leu Val Arg Leu Gly Gln Arg Val Ile Gln		
	1970	1975	1980
Val Glu Asn Ser His	Ile Ile Leu Thr Gly Ala Ser Ala Leu Asn		
	1985	1990	1995
Lys Val Leu Gly Arg	Glu Val Tyr Thr Ser Asn Asn Gln Leu Gly		
	2000	2005	2010
Gly Val Gln Ile Met	His Tyr Asn Gly Val Ser His Ile Thr Val		
	2015	2020	2025
Pro Asp Asp Phe Glu	Gly Val Tyr Thr Ile Leu Glu Trp Leu Ser		
	2030	2035	2040
Tyr Met Pro Lys Asp	Asn His Ser Pro Val Pro Ile Ile Thr Pro		
	2045	2050	2055
Thr Asp Pro Ile Asp	Arg Glu Ile Glu Phe Leu Pro Ser Arg Ala		
	2060	2065	2070
Pro Tyr Asp Pro Arg	Trp Met Leu Ala Gly Arg Pro His Pro Thr		
	2075	2080	2085
Leu Lys Gly Thr Trp	Gln Ser Gly Phe Phe Asp His Gly Ser Phe		
	2090	2095	2100
Lys Glu Ile Met Ala	Pro Trp Ala Gln Thr Val Val Thr Gly Arg		

WO 02/079441

PCT/US02/09820

2105 2110 2115
 Ala Arg Leu Gly Gly Ile Pro Val Gly Val Ile Ala Val Glu Thr
 2120 2125 2130
 Arg Thr Val Glu Val Ala Val Pro Ala Asp Pro Ala Asn Leu Asp
 2135 2140 2145
 Ser Glu Ala Lys Ile Ile Gln Gln Ala Gly Gln Val Trp Phe Pro
 2150 2155 2160
 Asp Ser Ala Tyr Lys Thr Ala Gln Ala Val Lys Asp Phe Asn Arg
 2165 2170 2175
 Glu Lys Leu Pro Leu Met Ile Phe Ala Asn Trp Arg Gly Phe Ser
 2180 2185 2190
 Gly Gly Met Lys Asp Met Tyr Asp Gln Val Leu Lys Phe Gly Ala
 2195 2200 2205
 Tyr Ile Val Asp Gly Leu Arg Gln Tyr Lys Gln Pro Ile Leu Ile
 2210 2215 2220
 Tyr Ile Pro Pro Tyr Ala Glu Leu Arg Gly Gly Ser Trp Val Val
 2225 2230 2235
 Ile Asp Ala Thr Ile Asn Pro Leu Cys Ile Glu Met Tyr Ala Asp
 2240 2245 2250
 Lys Glu Ser Arg Gly Gly Val Leu Glu Pro Glu Gly Thr Val Glu
 2255 2260 2265
 Ile Lys Phe Arg Lys Lys Asp Leu Ile Lys Ser Met Arg Arg Ile
 2270 2275 2280
 Asp Pro Ala Tyr Lys Lys Leu Met Glu Gln Leu Gly Glu Pro Asp
 2285 2290 2295
 Leu Ser Asp Lys Asp Arg Lys Asp Leu Glu Gly Arg Leu Lys Ala
 2300 2305 2310
 Arg Glu Asp Leu Leu Leu Pro Ile Tyr His Gln Val Ala Val Gln
 2315 2320 2325
 Phe Ala Asp Phe His Asp Thr Pro Gly Arg Met Leu Glu Lys Gly
 2330 2335 2340
 Val Ile Ser Asp Ile Leu Glu Trp Lys Thr Ala Arg Thr Phe Leu
 2345 2350 2355
 Tyr Trp Arg Leu Arg Arg Leu Leu Leu Glu Asp Gln Val Lys Gln
 2360 2365 2370
 Glu Ile Leu Gln Ala Ser Gly Glu Leu Ser His Val His Ile Gln
 2375 2380 2385
 Ser Met Leu Arg Arg Trp Phe Val Glu Thr Glu Gly Ala Val Lys
 2390 2395 2400
 Ala Tyr Leu Trp Asp Asn Asn Gln Val Val Val Gln Trp Leu Glu
 2405 2410 2415
 Gln His Trp Gln Ala Gly Asp Gly Pro Arg Ser Thr Ile Arg Glu
 2420 2425 2430
 Asn Ile Thr Tyr Leu Lys His Asp Ser Val Leu Lys Thr Ile Arg
 2435 2440 2445
 Gly Leu Val Glu Glu Asn Pro Glu Val Ala Val Asp Cys Val Ile
 2450 2455 2460
 Tyr Leu Ser Gln His Ile Ser Pro Ala Glu Arg Ala Gln Val Val
 2465 2470 2475
 His Leu Leu Ser Thr Met Asp Ser Pro Ala Ser Thr
 2480 2485

<210> 2
 <211> 311
 <212> PRT
 <213> Homo sapiens

WO 02/079441

PCT/US02/09820

<220>
 <221> misc_feature
 <223> Incyte ID No: 1482539CD1

<400> 2
 Met Ala Arg Pro Gln Pro Cys Gly Pro Pro His Ala Arg Cys Gly
 1 5 10 15
 Ser Pro Ser Leu Pro Glu Arg Pro Leu Gln Val Lys Val Val Gly
 20 25 30
 Leu Phe Ser Cys Pro Asn Phe Gln Ile Ala Lys Ser Ala Ala Glu
 35 40 45
 Asn Leu Lys Asn Asn His Pro Ser Lys Phe Glu Asp Pro Ile Leu
 50 55 60
 Val Pro Leu Gln Glu Phe Ala Trp His Gln Tyr Leu Gln Glu Lys
 65 70 75
 Lys Arg Glu Leu Lys Asn Glu Thr Trp Glu Tyr Ser Ser Ser Val
 80 85 90
 Ile Ser Phe Val Asn Gly Gln Phe Leu Gly Asp Ala Leu Asp Leu
 95 100 105
 Gln Lys Trp Ala His Glu Val Trp Asp Ile Val Asp Ile Lys Pro
 110 115 120
 Ser Ala Leu Tyr Asp Ala Leu Thr Glu Asp Phe Ser Ala Lys Phe
 125 130 135
 Leu Arg Asp Thr Lys His Asp Phe Val Phe Leu Asp Ile Cys Ile
 140 145 150
 Asp Ser Ser Pro Ile Gly Arg Leu Ile Phe Glu Leu Tyr Cys Asp
 155 160 165
 Val Cys Pro Lys Thr Cys Lys Asn Phe Gln Val Leu Cys Thr Gly
 170 175 180
 Lys Ala Gly Phe Ser Gln Arg Gly Ile Arg Leu His Tyr Lys Asn
 185 190 195
 Ser Ile Phe His Arg Ile Val Gln Asn Gly Trp Ile Gln Gly Gly
 200 205 210
 Asp Ile Val Tyr Gly Lys Gly Asp Asn Gly Glu Ser Ile Tyr Gly
 215 220 225
 Pro Thr Phe Glu Asp Glu Asn Phe Ser Val Pro His Asn Lys Arg
 230 235 240
 Gly Val Leu Gly Met Ala Asn Lys Gly Arg His Ser Asn Gly Ser
 245 250 255
 Gln Phe Tyr Ile Thr Leu Gln Ala Thr Pro Tyr Leu Asp Arg Lys
 260 265 270
 Phe Val Ala Phe Gly Gln Leu Ile Glu Gly Thr Glu Val Leu Lys
 275 280 285
 Gln Leu Glu Leu Val Pro Thr Gln Asn Glu Arg Pro Ile His Met
 290 295 300
 Cys Arg Ile Thr Asp Ser Gly Asp Pro Tyr Ala
 305 310

<210> 3
 <211> 406
 <212> FRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2234213CD1

WO 02/079441

PCT/US02/09820

```

<400> 3
Met Val Pro Pro Pro Ser Arg Gly Gly Ala Ala Arg Gly Gln
1 5 10 15
Leu Gly Arg Ser Leu Gly Pro Leu Leu Leu Leu Ala Leu Gly
20 25 30
His Thr Trp Thr Tyr Arg Glu Glu Pro Glu Asp Gly Asp Arg Glu
35 40 45
Ile Cys Ser Glu Ser Lys Ile Ala Thr Thr Lys Tyr Pro Cys Leu
50 55 60
Lys Ser Ser Gly Glu Leu Thr Thr Cys Tyr Arg Lys Lys Cys Cys
65 70 75
Lys Gly Tyr Lys Phe Val Leu Gly Gln Cys Ile Pro Glu Asp Tyr
80 85 90
Asp Val Cys Ala Glu Ala Pro Cys Glu Gln Gln Cys Thr Asp Asn
95 100 105
Phe Gly Arg Val Leu Cys Thr Cys Tyr Pro Gly Tyr Arg Tyr Asp
110 115 120
Arg Glu Arg His Arg Lys Arg Glu Lys Pro Tyr Cys Leu Asp Ile
125 130 135
Asp Glu Cys Ala Ser Ser Asn Gly Thr Leu Cys Ala His Ile Cys
140 145 150
Ile Asn Thr Leu Gly Ser Tyr Arg Cys Glu Cys Arg Glu Gly Tyr
155 160 165
Ile Arg Glu Asp Asp Gly Lys Thr Cys Thr Arg Gly Asp Lys Tyr
170 175 180
Pro Asn Asp Thr Gly His Glu Lys Ser Glu Asn Met Val Lys Ala
185 190 195
Gly Thr Cys Cys Ala Thr Cys Lys Glu Phe Tyr Gln Met Lys Gln
200 205 210
Thr Val Leu Gln Leu Lys Gln Lys Ile Ala Leu Leu Pro Asn Asn
215 220 225
Ala Ala Asp Leu Gly Lys Tyr Ile Thr Gly Asp Lys Val Leu Ala
230 235 240
Ser Asn Thr Tyr Leu Pro Gly Pro Pro Gly Leu Pro Gly Gly Gln
245 250 255
Gly Pro Pro Gly Ser Pro Gly Pro Lys Gly Ser Pro Gly Phe Pro
260 265 270
Gly Met Pro Gly Pro Pro Gly Gln Pro Gly Pro Arg Gly Ser Met
275 280 285
Gly Pro Met Gly Pro Ser Pro Asp Leu Ser His Ile Lys Gln Gly
290 295 300
Arg Arg Gly Pro Val Gly Pro Pro Gly Ala Pro Gly Arg Asp Gly
305 310 315
Ser Lys Gly Glu Arg Gly Ala Pro Gly Pro Arg Gly Ser Pro Gly
320 325 330
Pro Pro Gly Ser Phe Asp Phe Leu Leu Leu Met Leu Ala Asp Ile
335 340 345
Arg Asn Asp Ile Thr Glu Leu Gln Glu Lys Val Phe Gly His Arg
350 355 360
Thr His Ser Ser Ala Glu Glu Phe Pro Leu Pro Gln Glu Phe Pro
365 370 375
Ser Tyr Pro Glu Ala Met Asp Leu Gly Ser Gly Asp Asp His Pro
380 385 390
Arg Arg Thr Glu Thr Arg Asp Leu Arg Ala Pro Arg Asp Phe Tyr
395 400 405
Pro

```

WO 02/079441

PCT/US02/09820

<210> 4
<211> 442
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1345785CD1

```

<400> 4
Met Ala Leu Pro Ser Arg Ile Leu Leu Trp Lys Leu Val Leu Leu
 1          5          10          15
Gln Ser Ser Ala Val Leu Leu His Ser Gly Ser Ser Val Pro Ala
 20          25          30
Ala Ala Gly Ser Ser Val Val Ser Glu Ser Ala Val Ser Trp Glu
 35          40          45
Ala Gly Ala Arg Ala Val Leu Arg Cys Gln Ser Pro Arg Met Val
 50          55          60
Trp Thr Gln Asp Arg Leu His Asp Arg Gln Arg Val Leu His Trp
 65          70          75
Asp Leu Arg Gly Pro Gly Gly Gly Pro Ala Arg Arg Leu Leu Asp
 80          85          90
Leu Tyr Ser Ala Gly Glu Gln Arg Val Tyr Glu Ala Arg Asp Arg
 95          100         105
Gly Arg Leu Glu Leu Ser Ala Ser Ala Phe Asp Asp Gly Asn Phe
 110         115         120
Ser Leu Leu Ile Arg Ala Val Glu Glu Thr Asp Ala Gly Leu Tyr
 125         130         135
Thr Cys Asn Leu His His His Tyr Cys His Leu Tyr Glu Ser Leu
 140         145         150
Ala Val Arg Leu Glu Val Thr Asp Gly Pro Pro Ala Thr Pro Ala
 155         160         165
Tyr Trp Asp Gly Glu Lys Glu Val Leu Ala Val Ala Arg Gly Ala
 170         175         180
Pro Ala Leu Leu Thr Cys Val Asn Arg Gly His Val Trp Thr Asp
 185         190         195
Arg His Val Glu Glu Ala Gln Gln Val Val His Trp Asp Arg Gln
 200         205         210
Pro Pro Gly Val Pro His Asp Arg Ala Asp Arg Leu Leu Asp Leu
 215         220         225
Tyr Ala Ser Gly Glu Arg Arg Ala Tyr Gly Pro Leu Phe Leu Arg
 230         235         240
Asp Arg Val Ala Val Gly Ala Asp Ala Phe Glu Arg Gly Asp Phe
 245         250         255
Ser Leu Arg Ile Glu Pro Leu Glu Val Ala Asp Glu Gly Thr Tyr
 260         265         270
Ser Cys His Leu His His His Tyr Cys Gly Leu His Glu Arg Arg
 275         280         285
Val Phe His Leu Thr Val Ala Glu Pro His Ala Glu Pro Pro Pro
 290         295         300
Arg Gly Ser Pro Gly Asn Gly Ser Ser His Ser Gly Ala Pro Gly
 305         310         315
Pro Asp Pro Thr Leu Ala Arg Gly His Asn Val Ile Asn Val Ile
 320         325         330

```

WO 02/079441

PCT/US02/09820

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

<210> 5
 <211> 146
 <212> FRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3807190CD1

<400> 5
 Met Pro Pro Ile Trp Ser Pro Cys Phe Leu Leu Cys Leu Ser Ser
 1 5 10 15
 Val His Phe Pro His Ser Ser Trp Tyr Lys Pro Val Lys Met Gln
 20 25 30
 Val Gly Ser Cys His Leu Trp Ala Lys Cys Tyr Ser Gly Phe Leu
 35 40 45
 Leu His Ser Val Asn Ala Asp Ile Leu Arg Arg Leu Ile Asn Lys
 50 55 60
 Tyr Pro Val Ile Ser Gly Leu Asn Leu Pro Leu Ser His Leu Ala
 65 70* 75
 Ser Thr Ile Leu Ala Ser Leu Ser His Ser Val Ala Gln Ala Gly
 80 85 90
 Val Gln Trp Arg Asp Leu Gly Ser Leu Gln Thr Leu Pro Pro Arg
 95 100 105
 Phe Lys Gln Ser Ser His Leu Ser Leu Pro Ser Ser Trp Tyr Tyr
 110 115 120
 Arg Cys Ala Pro Pro Cys Leu Ala Asp Phe Cys Ile Val Ser Arg
 125 130 135
 Asp Arg Ile Ser Pro Cys Trp Ser Asp Trp Phe
 140 145

<210> 6
 <211> 109
 <212> FRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4856078CD1

WO 02/079441

PCT/US02/09820

<400> 6
 Met Leu Leu Phe Leu Glu Leu Ser Leu Gly Leu Val Leu Arg Met
 1 5 10 15
 Val Leu Gly Gln Pro Ile Leu Phe Ser Ala Met Leu Gly Asn Thr
 20 25 30
 Ala Leu Gly Val Gln Lys Ser Ala Ala His Gly Phe Ser Gln Ala
 35 40 45
 Val Arg Met Gly Glu Ala Trp Ala Ala Gln Gln Arg Gly Arg Cys
 50 55 60
 Leu Glu Glu Arg Arg Gly Phe Cys Arg Gly Glu Cys Gly Gln Ala
 65 70 75
 Pro Gly Arg Pro Arg Leu Pro Pro Leu Cys Pro Cys Arg Val Leu
 80 85 90
 Ser Pro Glu Pro Pro Arg Gly Thr His Ser Glu Ala Arg Trp Gly
 95 100 105
 Glu Arg Ser Pro

<210> 7
 <211> 78
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 6106886CD1

<400> 7
 Met Ala Met Pro Ser His Ser Met Val Leu Leu Leu Leu Arg
 1 5 10 15
 Met Ile Leu Leu His Lys Gly Phe Arg Arg Cys Pro Arg Thr Arg
 20 25 30
 Arg Phe Pro Ser Ala Gln Leu His Glu Ala His Ile Ser Trp Asp
 35 40 45
 Gly Ile Arg Lys Met Ala Thr Phe Thr Ser Thr Tyr Val Leu Trp
 50 55 60
 His Gln Pro Leu Phe Ser Leu Pro Ser Pro Thr Ser His Pro Leu
 65 70 75
 Leu Leu Gly

<210> 8
 <211> 721
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5627037CD1

<400> 8
 Met Asp Gln Gly Arg Ala Gly Pro Gly His Gly Pro Gly Leu Lys
 1 5 10 15
 Ala Trp Pro Arg Tyr Arg Val Val Gly Ser Ala Asp Ala Gly Gln
 20 25 30
 Tyr Asn Leu Glu Ile Thr Asp Ala Glu Leu Ser Asp Asp Ala Ser

WO 02/079441

PCT/US02/09820

	35	40	45											
Tyr	Glu	Cys	Gln	Ala	Thr	Glu	Ala	Ala	Leu	Arg	Ser	Arg	Arg	Ala
	50								55					60
Lys	Leu	Thr	Val	Leu	Ile	Pro	Pro	Glu	Asp	Thr	Arg	Ile	Asp	Gly
	65								70					75
Gly	Pro	Val	Ile	Leu	Leu	Gln	Ala	Gly	Thr	Pro	His	Asn	Leu	Thr
	80								85					90
Cys	Arg	Ala	Phe	Asn	Ala	Lys	Pro	Ala	Ala	Thr	Ile	Ile	Trp	Phe
	95								100					105
Arg	Asp	Gly	Thr	Gln	Gln	Glu	Gly	Ala	Val	Ala	Ser	Thr	Glu	Leu
	110								115					120
Leu	Lys	Asp	Gly	Lys	Arg	Glu	Thr	Thr	Val	Ser	Gln	Leu	Leu	Ile
	125								130					135
Asn	Pro	Thr	Asp	Leu	Asp	Ile	Gly	Arg	Val	Phe	Thr	Cys	Arg	Ser
	140								145					150
Met	Asn	Glu	Ala	Ile	Pro	Ser	Gly	Lys	Glu	Thr	Ser	Ile	Glu	Leu
	155								160					165
Asp	Val	His	His	Pro	Pro	Thr	Val	Thr	Leu	Ser	Ile	Glu	Pro	Gln
	170								175					180
Thr	Val	Gln	Glu	Gly	Glu	Arg	Val	Val	Phe	Thr	Cys	Gln	Ala	Thr
	185								190					195
Ala	Asn	Pro	Glu	Ile	Leu	Gly	Tyr	Arg	Trp	Ala	Lys	Gly	Gly	Phe
	200								205					210
Leu	Ile	Glu	Asp	Ala	His	Glu	Ser	Arg	Tyr	Glu	Thr	Asn	Val	Asp
	215								220					225
Tyr	Ser	Phe	Phe	Thr	Glu	Pro	Val	Ser	Cys	Glu	Val	His	Asn	Lys
	230								235					240
Val	Gly	Ser	Thr	Asn	Val	Ser	Thr	Leu	Val	Asn	Val	His	Phe	Ala
	245								250					255
Pro	Arg	Ile	Val	Val	Asp	Pro	Lys	Pro	Thr	Thr	Thr	Asp	Ile	Gly
	260								265					270
Ser	Asp	Val	Thr	Leu	Thr	Cys	Val	Trp	Val	Gly	Asn	Pro	Pro	Leu
	275								280					285
Thr	Leu	Thr	Trp	Thr	Lys	Lys	Asp	Ser	Asn	Met	Gly	Pro	Arg	Pro
	290								295					300
Pro	Gly	Ser	Pro	Pro	Glu	Ala	Ala	Leu	Ser	Ala	Gln	Val	Leu	Ser
	305								310					315
Asn	Ser	Asn	Gln	Leu	Leu	Leu	Lys	Ser	Val	Thr	Gln	Ala	Asp	Ala
	320								325					330
Gly	Thr	Tyr	Thr	Cys	Arg	Ala	Ile	Val	Pro	Arg	Ile	Gly	Val	Ala
	335								340					345
Glu	Arg	Glu	Val	Pro	Leu	Tyr	Val	Asn	Gly	Pro	Pro	Ile	Ile	Ser
	350								355					360
Ser	Glu	Ala	Val	Gln	Tyr	Ala	Val	Arg	Gly	Asp	Gly	Gly	Lys	Val
	365								370					375
Glu	Cys	Phe	Ile	Gly	Ser	Thr	Pro	Pro	Pro	Asp	Arg	Ile	Ala	Trp
	380								385					390
Ala	Trp	Lys	Glu	Asn	Phe	Leu	Glu	Val	Gly	Thr	Leu	Glu	Arg	Tyr
	395								400					405
Thr	Val	Glu	Arg	Thr	Asn	Ser	Gly	Ser	Gly	Val	Leu	Ser	Thr	Leu
	410								415					420
Thr	Ile	Asn	Asn	Val	Met	Glu	Ala	Asp	Phe	Gln	Thr	His	Tyr	Asn
	425								430					435
Cys	Thr	Ala	Trp	Asn	Ser	Phe	Gly	Pro	Gly	Thr	Ala	Ile	Ile	Gln
	440								445					450
Leu	Glu	Glu	Arg	Glu	Val	Leu	Pro	Val	Gly	Ile	Ile	Ala	Gly	Ala

WO 02/079441

PCT/US02/09820

455 460 465
 Thr Ile Gly Ala Ser Ile Leu Leu Ile Phe Phe Phe Ile Ala Leu
 470 475 480
 Val Phe Phe Leu Tyr Arg Arg Arg Lys Gly Ser Arg Lys Asp Val
 485 490 495
 Thr Leu Arg Lys Leu Asp Ile Lys Val Glu Thr Val Asn Arg Glu
 500 505 510
 Pro Leu Thr Met His Ser Asp Arg Glu Asp Asp Thr Ala Ser Val
 515 520 525
 Ser Thr Ala Thr Arg Val Met Lys Ala Ile Tyr Ser Ser Phe Lys
 530 535 540
 Asp Asp Val Asp Leu Lys Gln Asp Leu Arg Cys Asp Thr Ile Asp
 545 550 555
 Thr Arg Glu Glu Tyr Glu Met Lys Asp Pro Thr Asn Gly Tyr Tyr
 560 565 570
 Asn Val Arg Ala His Glu Asp Arg Pro Ser Ser Arg Ala Val Leu
 575 580 585
 Tyr Ala Asp Tyr Arg Ala Pro Gly Pro Ala Arg Phe Asp Gly Arg
 590 595 600
 Pro Ser Ser Arg Leu Ser His Ser Ser Gly Tyr Ala Gln Leu Asn
 605 610 615
 Thr Tyr Ser Arg Gly Pro Ala Ser Asp Tyr Gly Pro Glu Pro Thr
 620 625 630
 Pro Pro Gly Pro Ala Ala Pro Ala Gly Thr Asp Thr Thr Ser Gln
 635 640 645
 Leu Ser Tyr Glu Asn Tyr Glu Lys Phe Asn Ser His Pro Phe Pro
 650 655 660
 Gly Ala Ala Gly Tyr Pro Thr Tyr Arg Leu Gly Tyr Pro Gln Ala
 665 670 675
 Pro Pro Ser Gly Leu Glu Arg Thr Pro Tyr Glu Ala Tyr Asp Pro
 680 685 690
 Ile Gly Lys Tyr Ala Thr Ala Thr Arg Phe Ser Tyr Thr Ser Gln
 695 700 705
 His Ser Asp Tyr Gly Gln Arg Phe Gln Gln Arg Met Gln Thr His
 710 715 720
 Val

<210> 9
 <211> 144
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3688835CD1

<400> 9
 Met Leu Cys Leu Pro Lys Leu Leu Cys Ser Ala Arg Leu Trp Gly
 1 5 10 15
 Arg Pro Cys Leu Leu Leu Glu Arg Pro Leu Gly Pro Ala Val Ser
 20 25 30
 Ala Thr Ser Pro Pro Thr Leu Arg Met Thr Val His Leu Leu Ser
 35 40 45
 Trp Thr Cys Ser His Leu Pro Ala Asp Leu Ser Tyr Pro His Thr
 50 55 60

WO 02/079441

PCT/US02/09820

Pro Gln Gly Ser Gly Ser Pro Pro Thr Pro Ala Pro Leu Ser Gly
65 70 75
His Ser Phe Ser Arg Arg Asn Cys Ser Ser Ser Cys Phe Arg Asn
80 85 90
Pro Phe His Asp Asn Asp Ile Gly Pro Glu Ile Phe Pro Ile Ser
95 100 105
Ser Gln Lys Glu Phe Pro Ile Phe Pro Cys Lys Thr Asp Phe Ile
110 115 120
Ile Ile Cys Asp Arg Gly Ala Thr Ile Ala Pro Arg Thr Gln Pro
125 130 135
Glu Thr His Ser Leu Ala Val Gln Glu
140

<210> 10
<211> 124
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2295419CD1

<400> 10
Met Gly Pro Gln Pro Ala Ala Gly Leu Thr Leu Pro Leu Pro Ser
1 5 10 15
Leu Ser Phe Phe Leu Ser Arg Gln Trp Trp Arg Gly Val Arg Arg
20 25 30
Asp Arg Glu Thr Glu Val Arg Gln Gly Leu Val Pro Ser Ala Ala
35 40 45
Leu Gly Ser Trp Pro Pro Ser Gly Pro Ser Val Phe Pro Pro Ala
50 55 60
Leu Cys Ser Ala Asn Pro His Pro Pro Thr Gly Ile Gly Leu Gly
65 70 75
Leu Leu His Pro Ser Ala Cys Leu Ser Ile Phe Ile Phe Ser Phe
80 85 90
Leu Leu Val Leu Leu Cys Cys Val Phe Tyr Phe Ser Leu Lys Lys
95 100 105
Phe Phe Cys Gly Leu Gly Val Val Ala His Pro Cys Asn Pro Ser
110 115 120
Thr Leu Trp Gly

<210> 11
<211> 142
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2437896CD1

<400> 11
Met Asp Ala Gly Ser Ala Val Val Ser Leu Pro Thr Ala Phe Cys
1 5 10 15
Pro Tyr Leu Cys Trp Gly His Trp Asn Pro Asp Glu Arg Leu Glu
20 25 30

WO 02/079441

PCT/US02/09820

Pro Ser Thr Cys Leu Ala Val Cys Pro Glu Glu Met Gly Leu Gln
 35 40 45
 Gly Arg Val His Gly Lys Ser Gly Leu Gln Glu Asn Pro Cys Val
 50 55 60
 Ala Leu Trp Ser Arg His Phe Ala Ser Leu Ser Arg Asp Leu Ser
 65 70 75
 His Leu Gly Lys Gly Lys Met Ser Gly Gly Lys Ser Pro Lys His
 80 85 90
 Leu Gln Ala Ser Arg Val Thr Phe Ser Asp Pro Asp Leu Gln Gly
 95 100 105
 Glu Gln Ile Leu Tyr Tyr Asp Pro Val Tyr Thr Tyr Ile Trp Phe
 110 115 120
 Val Phe Cys Cys Met Leu Phe Val Lys Pro Val Ile Lys Phe Glu
 125 130 135
 Val Trp Gln Val Gly Phe Val
 140

<210> 12
 <211> 103
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2641482CD1

<400> 12
 Met Arg Arg Ser His Phe Pro Arg His Leu Ala Ala Val Ala Phe
 1 5 10 15
 Leu His Leu Pro Gln Asp Ile Ser Val Gln Gln Thr Leu Asp Leu
 20 25 30
 Val Pro Val Leu Glu Gly Leu Ala Val Ser Thr His Pro Ala Leu
 35 40 45
 Gln Leu Arg Arg Leu Ala Cys Gly Gly Trp Gln Thr Arg Pro Gln
 50 55 60
 Ala Ala Ala Pro Pro Leu Trp Asp Phe Ser Gln Gly Arg Arg Ala
 65 70 75
 Asp Gly Val Ser Val Thr Ala Ser Tyr Arg Ala Leu Gly Gly Gly
 80 85 90
 Pro Cys Ser Arg Ser Arg Pro Leu Pro Arg Pro His Thr
 95 100

<210> 13
 <211> 670
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7494292CD1

<400> 13
 Met Arg Gly Pro Ile Val Leu His Ile Cys Leu Ala Phe Cys Ser
 1 5 10 15
 Leu Leu Leu Phe Ser Val Ala Thr Gln Cys Leu Ala Phe Pro Lys
 20 25 30

WO 02/079441

PCT/US02/09820

Ile Glu Arg Arg Arg Glu Ile Ala His Val His Ala Glu Lys Gly
 35 40 45
 Gln Ser Asp Lys Met Asn Thr Asp Asp Leu Glu Asn Ser Ser Val
 50 55 60
 Thr Ser Lys Gln Thr Pro Gln Leu Val Val Ser Glu Asp Pro Met
 65 70 75
 Met Met Ser Ala Val Pro Ser Ala Thr Ser Leu Asn Lys Ala Phe
 80 85 90
 Ser Ile Asn Lys Glu Ile Gln Pro Gly Gln Ala Gly Leu Met Gln
 95 100 105
 Thr Glu Arg Pro Gly Val Ser Thr Pro Thr Glu Ser Gly Val Pro
 110 115 120
 Ser Ala Glu Glu Val Phe Gly Ser Ser Gln Pro Glu Arg Ile Ser
 125 130 135
 Pro Glu Ser Gly Leu Ala Lys Ala Met Leu Thr Ile Ala Ile Thr
 140 145 150
 Ala Thr Pro Ser Leu Thr Val Asp Glu Lys Glu Glu Leu Leu Thr
 155 160 165
 Ser Thr Asn Phe Gln Pro Ile Val Glu Glu Ile Thr Glu Thr Thr
 170 175 180
 Lys Gly Phe Leu Lys Tyr Met Asp Asn Gln Ser Phe Ala Thr Glu
 185 190 195
 Ser Gln Glu Gly Val Gly Leu Gly His Ser Pro Ser Ser Tyr Val
 200 205 210
 Asn Thr Lys Glu Met Leu Thr Thr Asn Pro Lys Thr Glu Lys Phe
 215 220 225
 Glu Ala Asp Thr Asp His Arg Thr Thr Ser Phe Pro Gly Ala Glu
 230 235 240
 Ser Thr Ala Gly Ser Glu Pro Gly Ser Leu Thr Pro Asp Lys Glu
 245 250 255
 Lys Pro Ser Gln Met Thr Ala Asp Asn Thr Gln Ala Ala Ala Thr
 260 265 270
 Lys Gln Pro Leu Glu Thr Ser Glu Tyr Thr Leu Ser Val Glu Pro
 275 280 285
 Glu Thr Asp Ser Leu Leu Gly Ala Pro Glu Val Thr Val Ser Val
 290 295 300
 Ser Thr Ala Val Pro Ala Ala Ser Ala Leu Ser Asp Glu Trp Asp
 305 310 315
 Asp Thr Lys Leu Glu Ser Val Ser Arg Ile Arg Thr Pro Lys Leu
 320 325 330
 Gly Asp Asn Glu Glu Thr Gln Val Arg Thr Glu Met Ser Gln Thr
 335 340 345
 Ala Gln Val Ser His Glu Gly Met Glu Gly Gly Gln Pro Trp Thr
 350 355 360
 Glu Ala Ala Gln Val Ala Leu Gly Leu Pro Glu Gly Glu Thr His
 365 370 375
 Thr Gly Thr Ala Leu Leu Ile Ala His Gly Asn Glu Arg Ser Pro
 380 385 390
 Ala Phe Thr Asp Gln Ser Ser Phe Thr Pro Thr Ser Leu Met Glu
 395 400 405
 Asp Met Lys Val Ser Ile Val Asn Leu Leu Gln Ser Thr Gly Asp
 410 415 420
 Phe Thr Glu Ser Thr Lys Glu Asn Asp Ala Leu Phe Phe Leu Glu
 425 430 435
 Thr Thr Val Ser Val Ser Val Tyr Glu Ser Glu Ala Asp Gln Leu
 440 445 450

WO 02/079441

PCT/US02/09820

Leu Gly Asn Thr Met Lys Asp Ile Ile Thr Gln Glu Met Thr Thr
 455 460 465
 Ala Val Gln Glu Pro Asp Ala Thr Leu Ser Met Val Thr Gln Glu
 470 475 480
 Gln Val Ala Thr Leu Glu Leu Ile Arg Asp Ser Gly Lys Thr Glu
 485 490 495
 Glu Glu Lys Glu Asp Pro Ser Pro Val Ser Asp Val Pro Gly Val
 500 505 510
 Thr Gln Leu Ser Arg Arg Trp Glu Pro Leu Ala Thr Thr Ile Ser
 515 520 525
 Thr Thr Val Val Pro Leu Ser Phe Glu Val Thr Pro Thr Val Glu
 530 535 540
 Glu Gln Met Asp Thr Val Thr Gly Pro Asn Glu Glu Phe Thr Pro
 545 550 555
 Val Leu Gly Ser Pro Val Thr Pro Pro Gly Ile Met Val Gly Glu
 560 565 570
 Pro Ser Ile Ser Pro Ala Leu Pro Ala Leu Glu Ala Ser Ser Glu
 575 580 585
 Arg Arg Thr Val Val Pro Ser Ile Thr Arg Val Asn Thr Ala Ala
 590 595 600
 Ser Tyr Gly Leu Asp Gln Leu Glu Ser Glu Glu Thr Gly Phe His
 605 610 615
 His Val Ala Gln Ala Arg Leu Lys Leu Leu Gly Leu Arg Ser Leu
 620 625 630
 Pro Ala Ser Ala Ser Gln Ser Val Gly Ile Thr Ser Val Asn Ser
 635 640 645
 Cys Thr Gln Pro Arg Lys Tyr Leu Asn Ser Cys Leu Lys Trp Lys
 650 655 660
 Leu Asn Pro Lys His Phe Ala Thr Gly Val
 665 670

<210> 14

<211> 215

<212> FFT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7636288CD1

<400> 14

Met Gly Ala Gln Pro Cys Trp Pro Arg Val Leu Leu Gly Ala Leu
 1 5 10 15
 Gly Cys Ser Pro Leu Trp Pro Arg Leu Ser Cys Val Cys Trp Ala
 20 25 30
 Gly Pro Ala Cys Cys Gly Ala Ala Arg Ser Val Pro Ser Gln Leu
 35 40 45
 Pro Ala Leu Gly Ala Thr His Thr Thr Pro Trp Leu Lys Ile Cys
 50 55 60
 Arg Gly Leu Ala Ser Met Val Leu Arg Cys Thr Trp Gly Ser Gln
 65 70 75
 Glu Ala Asn Ser Leu His Arg Glu Ala Val Gly Thr Pro Leu Ser
 80 85 90
 Gly Leu Leu Pro Ser Asp Gly Thr Thr Gln Lys Ala Gly Asn Ser
 95 100 105
 Val Val Leu Gln His Arg Asp Val Lys Asp Ser Gly Pro His Ala

WO 02/079441

PCT/US02/09820

110 115 120
 Ala Pro Asp Arg Leu Leu Ser Ser Arg Thr Ile Arg Arg Gly Arg
 125 130 135
 Gly Val Arg Gly Gly Gly Trp Gly Gln Glu Thr Asp Trp Arg Gln
 140 145 150
 Gln Gly Gln Glu Val Ala Gly Arg Gly Gln Leu Pro Trp Pro Trp
 155 160 165
 Leu Gln Val Glu Glu Arg Ser Leu Ser Arg Pro His Ala Ala Ala
 170 175 180
 Ser Leu Arg Gln Val Pro Pro Lys Gly Ser Leu Ser Trp Pro Ala
 185 190 195
 Gln Ser Gly Ser Ser Pro Trp Ser Ile Pro Gly Thr Trp Ala Leu
 200 205 210
 Glu Ile Arg Cys Gln
 215

<210> 15
 <211> 86
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 8095391CD1

<400> 15
 Met Lys Ala Ile Cys Ala Arg Arg Val Ala Ala Ser Gln Ile Ala
 1 5 10 15
 Ala Ala Ser Gln Ile Ala Ala Ala Ser Ala Gly Pro Arg Trp Arg
 20 25 30
 Cys Arg Lys Gln Pro Asn Leu Ser Asp Cys Gln Ser Val Ser Pro
 35 40 45
 Gly Ser Phe Ile Ile Ser Arg Lys Ile Leu Leu Gly Ser Ile Ser
 50 55 60
 Ile Pro Ser His Lys Val Leu Cys Phe Gln Gly Ala Asp Pro Thr
 65 70 75
 Gly Gly Ser Phe Arg Lys Leu Ile Arg Cys Ser
 80 85

<210> 16
 <211> 495
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 532018CD1

<400> 16
 Met Ser Met Leu Val Val Phe Leu Leu Leu Trp Gly Val Thr Trp
 1 5 10 15
 Gly Pro Val Thr Glu Ala Ala Ile Phe Tyr Glu Thr Gln Pro Ser
 20 25 30
 Leu Trp Ala Glu Ser Glu Ser Leu Leu Lys Pro Leu Ala Asn Val
 35 40 45
 Thr Leu Thr Cys Gln Ala Arg Leu Glu Thr Pro Asp Phe Gln Leu

WO 02/079441

PCT/US02/09820

	50	55	60
Phe Lys Asn Gly Val	Ala Gln Glu Pro	Val His Leu Asp Ser	Pro
	65	70	75
Ala Ile Lys His Gln	Phe Leu Leu Thr	Gly Asp Thr Gln Gly	Arg
	80	85	90
Tyr Arg Cys Arg Ser	Gly Leu Ser Thr	Gly Trp Thr Gln Leu	Ser
	95	100	105
Lys Leu Leu Glu Leu	Thr Gly Pro Lys	Ser Leu Pro Ala Pro	Trp
	110	115	120
Leu Ser Met Ala Pro	Val Ser Trp Ile	Thr Pro Gly Leu Lys	Thr
	125	130	135
Thr Ala Val Cys Arg	Gly Val Leu Arg	Gly Val Thr Phe Leu	Leu
	140	145	150
Arg Arg Glu Gly Asp	His Glu Phe Leu	Glu Val Pro Glu Ala	Gln
	155	160	165
Glu Asp Val Glu Ala	Thr Phe Pro Val	His Gln Pro Gly Asn	Tyr
	170	175	180
Ser Cys Ser Tyr Arg	Thr Asp Gly Glu	Gly Ala Leu Ser Glu	Pro
	185	190	195
Ser Ala Thr Val Thr	Ile Glu Glu Leu	Ala Ala Pro Pro Pro	Pro
	200	205	210
Val Leu Met His His	Gly Glu Ser Ser	Gln Val Leu His Pro	Gly
	215	220	225
Asn Lys Val Thr Leu	Thr Cys Val Ala	Pro Leu Ser Gly Val	Asp
	230	235	240
Phe Gln Leu Arg Arg	Gly Glu Lys Glu	Leu Leu Val Pro Arg	Ser
	245	250	255
Ser Thr Ser Pro Asp	Arg Ile Phe Phe	His Leu Asn Ala Val	Ala
	260	265	270
Leu Gly Asp Gly Gly	His Tyr Thr Cys	Arg Tyr Arg Leu His	Asp
	275	280	285
Asn Gln Asn Gly Trp	Ser Gly Asp Ser	Ala Pro Val Glu Leu	Ile
	290	295	300
Leu Ser Asp Glu Thr	Leu Pro Ala Pro	Glu Phe Ser Pro Glu	Pro
	305	310	315
Glu Ser Gly Arg Ala	Leu Arg Leu Arg	Cys Leu Ala Pro Leu	Glu
	320	325	330
Gly Ala Arg Phe Ala	Leu Val Arg Glu	Asp Arg Gly Gly Arg	Arg
	335	340	345
Val His Arg Phe Gln	Ser Pro Ala Gly	Thr Glu Ala Leu Phe	Glu
	350	355	360
Leu His Asn Ile Ser	Val Ala Asp Ser	Ala Asn Tyr Ser Cys	Val
	365	370	375
Tyr Val Asp Leu Lys	Pro Pro Phe Gly	Gly Ser Ala Pro Ser	Glu
	380	385	390
Arg Leu Glu Leu His	Val Asp Gly Pro	Pro Pro Arg Pro Gln	Leu
	395	400	405
Arg Ala Thr Trp Ser	Gly Ala Val Leu	Ala Gly Arg Asp Ala	Val
	410	415	420
Leu Arg Cys Glu Gly	Pro Ile Pro Asp	Val Thr Phe Glu Leu	Leu
	425	430	435
Arg Glu Gly Glu Thr	Lys Ala Val Lys	Thr Val Arg Thr Pro	Gly
	440	445	450
Ala Ala Ala Asn Leu	Glu Leu Ile Phe	Val Gly Pro Gln His	Ala
	455	460	465
Gly Asn Tyr Arg Cys	Arg Tyr Arg Ser	Trp Val Pro His Thr	Phe

WO 02/079441

PCT/US02/09820

	470		475		480
Glu Ser Glu Leu Ser	Asp Pro Val Glu	Leu Leu Val Ala	Glu Ser		
	485		490		495

<210> 17
 <211> 60
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7481063CD1

<400> 17
 Met Gly Pro Asp Glu Leu Ala Thr Gly Thr Met Lys Leu Ala Pro
 1 5 10 15
 Thr Phe Gly Leu Leu Phe Thr Ala Val Leu Leu Ile Ser Arg Ala
 20 25 30
 Asn Cys Asp Ile Cys Pro Val Val Thr Lys Asp Val Asp Leu Phe
 35 40 45
 Leu Val Gly Thr Pro Asp Glu Tyr Val Asp His Val Ala Gln Tyr
 50 55 60

<210> 18
 <211> 241
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2044436CD1

<400> 18
 Met Asp Cys Arg Val His Met Arg Pro Ile Gly Leu Thr Trp Val
 1 5 10 15
 Leu Gln Leu Thr Leu Ala Trp Ile Leu Leu Glu Ala Cys Gly Gly
 20 25 30
 Ser Arg Pro Leu Gln Ala Arg Ser Gln Gln His His Gly Leu Ala
 35 40 45
 Ala Asp Leu Gly Lys Gly Lys Leu His Leu Ala Gly Pro Cys Cys
 50 55 60
 Pro Ser Glu Met Asp Thr Thr Glu Thr Ser Gly Pro Gly Asn His
 65 70 75
 Pro Glu Arg Cys Gly Val Pro Ser Pro Glu Cys Glu Ser Phe Leu
 80 85 90
 Glu His Leu Gln Arg Ala Leu Arg Ser Arg Phe Arg Leu Arg Leu
 95 100 105
 Leu Gly Val Arg Gln Ala Gln Pro Leu Cys Glu Glu Leu Cys Gln
 110 115 120
 Ala Trp Phe Ala Asn Cys Glu Asp Asp Ile Thr Cys Gly Ala Pro
 125 130 135
 Thr Trp Gly Ser Pro Thr Ser Ser Arg Lys Lys Lys Gly Gly Leu
 140 145 150
 Val Ser Pro Ala Cys Ala Tyr Leu Trp Thr Asp Leu Arg Arg Arg

WO 02/079441

PCT/US02/09820

<223> Incyte ID No: 8039739CD1

<400> 20

```

Met Leu Leu Trp Leu Leu Leu Ile Leu Arg Val Ala Pro Lys
 1          5          10          15
Ala Val Leu Leu Leu Asn Pro Pro Trp Ser Thr Ala Phe Lys Gly
 20          25          30
Glu Lys Val Ala Leu Ile Cys Ser Ser Ile Ser His Ser Leu Ala
 35          40          45
Gln Gly Asp Thr Tyr Trp Tyr His Asp Glu Lys Leu Leu Lys Ile
 50          55          60
Lys His Asp Lys Ile Gln Ile Thr Glu Pro Gly Asn Tyr Gln Cys
 65          70          75
Lys Thr Arg Gly Ser Ser Leu Ser Asp Ala Val His Val Glu Phe
 80          85          90
Ser Pro Asp Trp Leu Ile Leu Gln Ala Leu His Pro Val Phe Glu
 95          100          105
Gly Asp Asn Val Ile Leu Arg Cys Gln Gly Lys Asp Asn Lys Asn
 110          115          120
Thr His Gln Lys Val Tyr Tyr Lys Asp Gly Lys Gln Leu Pro Asn
 125          130          135
Ser Tyr Asn Leu Glu Lys Ile Thr Val Asn Ser Val Ser Arg Asp
 140          145          150
Asn Ser Lys Tyr His Cys Thr Ala Tyr Arg Lys Phe Tyr Ile Leu
 155          160          165
Asp Ile Glu Val Thr Ser Lys Pro Leu Asn Ile Gln Val Gln Glu
 170          175          180
Leu Phe Leu His Pro Val Leu Arg Ala Ser Ser Ser Thr Pro Ile
 185          190          195
Glu Gly Ser Pro Met Thr Leu Thr Cys Glu Thr Gln Leu Ser Pro
 200          205          210
Gln Arg Pro Asp Val Gln Leu Gln Phe Ser Leu Phe Arg Asp Ser
 215          220          225
Gln Thr Leu Gly Leu Gly Trp Ser Arg Ser Pro Arg Leu Gln Ile
 230          235          240
Pro Ala Met Trp Thr Glu Asp Ser Gly Ser Tyr Trp Cys Glu Val
 245          250          255
Glu Thr Val Thr His Ser Ile Lys Lys Arg Ser Leu Arg Ser Gln
 260          265          270
Ile Arg Val Gln Arg Val Pro Val Ser Asn Val Asn Leu Glu Ile
 275          280          285
Arg Pro Thr Gly Gly Gln Leu Ile Glu Gly Glu Asn Met Val Leu
 290          295          300
Ile Cys Ser Val Ala Gln Gly Ser Gly Thr Val Thr Phe Ser Trp
 305          310          315
His Lys Glu Gly Arg Val Arg Ser Leu Gly Arg Lys Thr Gln Arg
 320          325          330
Ser Leu Leu Ala Glu Leu His Val Leu Thr Val Lys Glu Ser Asp
 335          340          345
Ala Gly Arg Tyr Tyr Cys Ala Ala Asp Asn Val His Ser Pro Ile
 350          355          360
Leu Ser Thr Trp Ile Arg Val Thr Val Arg Ile Pro Val Ser His
 365          370          375
Pro Val Leu Thr Phe Arg Ala Pro Arg Ala His Thr Val Val Gly
 380          385          390
Asp Leu Leu Glu Leu His Cys Glu Ser Leu Arg Gly Ser Pro Pro

```

WO 02/079441

PCT/US02/09820

```

395          400          405
Ile Leu Tyr Arg Phe Tyr His Glu Asp Val Thr Leu Gly Asn Ser
410          415          420
Ser Ala Pro Ser Gly Gly Gly Ala Ser Phe Asn Leu Ser Leu Thr
425          430          435
Ala Glu His Ser Gly Asn Tyr Ser Cys Asp Ala Asp Asn Gly Leu
440          445          450
Gly Ala Gln His Ser His Gly Val Ser Leu Arg Val Thr Val Pro
455          460          465
Val Ser Arg Pro Val Leu Thr Leu Arg Ala Pro Gly Ala Gln Ala
470          475          480
Val Val Gly Asp Leu Leu Glu Leu His Cys Glu Ser Leu Arg Gly
485          490          495
Ser Phe Pro Ile Leu Tyr Trp Phe Tyr His Glu Asp Asp Thr Leu
500          505          510
Gly Asn Ile Ser Ala His Ser Gly Gly Gly Ala Ser Phe Asn Leu
515          520          525
Ser Leu Thr Thr Glu His Ser Gly Asn Tyr Ser Cys Glu Ala Asp
530          535          540
Asn Gly Leu Gly Ala Gln His Ser Lys Val Val Thr Leu Asn Val
545          550          555
Thr Gly Thr Ser Arg Asn Arg Thr Gly Leu Thr Ala Ala Gly Ile
560          565          570
Thr Gly Leu Val Leu Ser Ile Leu Val Leu Ala Ala Ala Ala Ala
575          580          585
Leu Leu His Tyr Ala Arg Ala Arg Arg Lys Pro Gly Gly Leu Ser
590          595          600
Ala Thr Gly Thr Ser Ser His Ser Pro Ser Glu Cys Gln Glu Pro
605          610          615
Ser Ser Ser Arg Pro Ser Arg Ile Asp Pro Gln Glu Pro Thr His
620          625          630
Ser Lys Pro Leu Ala Pro Met Glu Leu Glu Pro Met Tyr Ser Asn
635          640          645
Val Asn Pro Gly Asp Ser Asn Pro Ile Tyr Ser Gln Ile Trp Ser
650          655          660
Ile Gln His Thr Lys Glu Asn Ser Ala Asn Cys Pro Met Met His
665          670          675
Gln Glu His Glu Glu Leu Thr Val Leu Tyr Ser Glu Leu Lys Lys
680          685          690
Thr His Pro Asp Asp Ser Ala Gly Glu Ala Ser Ser Arg Gly Arg
695          700          705
Ala His Glu Glu Asp Asp Glu Glu Asn Tyr Glu Asn Val Pro Arg
710          715          720
Val Leu Leu Ala Ser Asp His
725

```

<210> 21

<211> 132

<212> PPT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1265837CD1

<400> 21

WO 02/079441

PCT/US02/09820

Met Cys Ser Ala Gly Glu Leu Leu Arg Gly Gly Asp Gly Gly Glu
 1 5 10 15
 Arg Asp Glu Asp Gly Asp Ala Leu Ala Glu Arg Glu Ala Ala Gly
 20 25 30
 Thr Glu Trp Asp Pro Gly Ala Ser Pro Arg Arg Arg Gly Gln Arg
 35 40 45
 Pro Lys Glu Ser Glu Gln Asp Val Glu Asp Ser Gln Asn His Thr
 50 55 60
 Gly Glu Pro Val Gly Asp Asp Tyr Lys Lys Met Gly Thr Leu Phe
 65 70 75
 Gly Glu Leu Asn Lys Asn Leu Ile Asn Met Gly Phe Thr Arg Met
 80 85 90
 Tyr Phe Gly Glu Arg Ile Val Glu Pro Val Ile Val Ile Phe Phe
 95 100 105
 Trp Val Met Leu Trp Phe Leu Gly Leu Gln Ala Leu Gly Leu Val
 110 115 120
 Ala Val Leu Cys Leu Val Ile Ile Tyr Val Gln Gln
 125 130

<210> 22

<211> 297

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5568527CD1

<400> 22

Met Ser Ala Trp Ala Ala Ala Ser Leu Ser Arg Ala Ala Ala Arg
 1 5 10 15
 Cys Leu Leu Ala Arg Gly Pro Gly Val Arg Ala Ala Pro Pro Arg
 20 25 30
 Asp Pro Arg Pro Ser His Pro Glu Pro Arg Gly Cys Gly Ala Ala
 35 40 45
 Pro Gly Arg Thr Leu His Phe Thr Ala Ala Val Pro Ala Gly His
 50 55 60
 Asn Lys Trp Ser Lys Val Arg His Ile Lys Gly Pro Lys Asp Val
 65 70 75
 Glu Arg Ser Arg Ile Phe Ser Lys Leu Cys Leu Asn Ile Arg Leu
 80 85 90
 Ala Val Lys Glu Gly Gly Pro Asn Pro Glu His Asn Ser Asn Leu
 95 100 105
 Ala Asn Ile Leu Glu Val Cys Arg Ser Lys His Met Pro Lys Ser
 110 115 120
 Thr Ile Glu Thr Ala Leu Lys Met Glu Lys Ser Lys Asp Thr Tyr
 125 130 135
 Leu Leu Tyr Glu Gly Arg Gly Pro Gly Gly Ser Ser Leu Leu Ile
 140 145 150
 Glu Ala Leu Ser Asn Ser Ser His Lys Cys Gln Ala Asp Ile Arg
 155 160 165
 His Ile Leu Asn Lys Asn Gly Gly Val Met Ala Val Gly Ala Arg
 170 175 180
 His Ser Phe Asp Lys Lys Gly Val Ile Val Val Glu Val Glu Asp
 185 190 195
 Arg Glu Lys Lys Ala Val Asn Leu Glu Arg Ala Leu Glu Met Ala

WO 02/079441

PCT/US02/09820

Ile	Glu	Ala	Gly	Ala	Glu	Asp	Val	Lys	Glu	Thr	Glu	Asp	Glu	Glu	200	205	210
															215	220	225
Glu	Arg	Asn	Val	Phe	Lys	Phe	Ile	Cys	Asp	Ala	Ser	Ser	Leu	His	230	235	240
Gln	Val	Arg	Lys	Lys	Leu	Asp	Ser	Leu	Gly	Leu	Cys	Ser	Val	Ser	245	250	255
Cys	Ala	Leu	Glu	Phe	Ile	Pro	Asn	Ser	Lys	Val	Gln	Leu	Ala	Glu	260	265	270
Pro	Asp	Leu	Glu	Gln	Ala	Ala	His	Leu	Ile	Gln	Ala	Leu	Ser	Asn	275	280	285
His	Glu	Asp	Val	Ile	His	Val	Tyr	Asp	Asn	Ile	Glu				290	295	

<210> 23
 <211> 218
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7503641CD1

<400> 23
 Met Ala Leu Pro Ser Arg Ile Leu Leu Trp Lys Leu Val Leu Leu
 1 5 10 15
 Gln Ser Ser Ala Val Leu Leu His Ser Gly Ser Ser Val Pro Ala
 20 25 30
 Ala Ala Gly Ser Ser Val Val Ser Glu Ser Ala Val Ser Trp Glu
 35 40 45
 Ala Gly Ala Arg Ala Val Leu Arg Cys Gln Ser Pro Arg Met Val
 50 55 60
 Trp Thr Gln Asp Arg Leu His Asp Arg Gln Arg Val Leu His Trp
 65 70 75
 Asp Leu Arg Gly Pro Gly Gly Gly Pro Ala Arg Arg Leu Leu Asp
 80 85 90
 Leu Tyr Ser Ala Gly Glu Gln Arg Val Tyr Glu Ala Arg Asp Arg
 95 100 105
 Gly Arg Leu Glu Leu Ser Ala Ser Ala Phe Asp Asp Gly Asn Phe
 110 115 120
 Ser Leu Leu Ile Arg Ala Val Glu Glu Thr Asp Ala Gly Leu Tyr
 125 130 135
 Thr Cys Asn Leu His His His Tyr Cys His Leu Tyr Glu Ser Leu
 140 145 150
 Ala Val Arg Leu Arg Glu Pro Arg Ala Arg Val Asp Arg Pro Ala
 155 160 165
 Arg Gly Gly Gly Ser Thr Gly Gly Ala Leu Gly Pro Ala Ala Ala
 170 175 180
 Arg Gly Pro Ala Arg Pro Arg Gly Pro Pro Ala Gly Pro Leu Arg
 185 190 195
 Val Gly Arg Ala Pro Arg Leu Arg Ala Pro Phe Ser Ala Arg Pro
 200 205 210
 Arg Gly Cys Gly Arg Gly Cys Leu
 215

<210> 24

WO 02/079441

PCT/US02/09820

<211> 266
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7503458CD1

<400> 24
 Met Asp Cys Arg Val His Met Arg Pro Ile Gly Leu Thr Trp Val
 1 5 10 15
 Leu Gln Leu Thr Leu Ala Trp Ile Leu Leu Glu Ala Cys Gly Gly
 20 25 30
 Ser Arg Pro Leu Gln Ala Arg Ser Gln Gln His His Gly Leu Ala
 35 40 45
 Ala Asp Leu Gly Lys Gly Lys Leu His Leu Ala Gly Asp Pro Gln
 50 55 60
 Ala Ser Val Pro Lys Pro Tyr Leu Asn Ile Gln Asp Pro Gly Ser
 65 70 75
 Gln Gly Ser Pro Leu Pro Gly Pro Cys Cys Pro Ser Glu Met Asp
 80 85 90
 Thr Thr Glu Thr Ser Gly Pro Gly Asn His Pro Glu Arg Cys Gly
 95 100 105
 Val Pro Ser Pro Glu Cys Glu Ser Phe Leu Glu His Leu Gln Arg
 110 115 120
 Ala Leu Arg Ser Arg Phe Arg Leu Arg Leu Leu Gly Val Arg Gln
 125 130 135
 Ala Gln Pro Leu Cys Glu Glu Leu Cys Gln Ala Trp Phe Ala Asn
 140 145 150
 Cys Glu Asp Asp Ile Thr Cys Gly Ala Pro Thr Trp Gly Ser Pro
 155 160 165
 Thr Ser Ser Arg Lys Lys Lys Gly Gly Leu Val Ser Pro Ala Cys
 170 175 180
 Ala Tyr Leu Trp Thr Asp Leu Arg Arg Arg Asp Gly Pro Leu Ser
 185 190 195
 Leu Gly Ser Gly Pro Arg Pro Thr Gly Gly Cys Ser Trp Ser Pro
 200 205 210
 Ser Leu Leu Gln His Leu His Leu Arg Gly Thr Ser Ser Gln Thr
 215 220 225
 Arg Thr Thr Gly Pro Gly Ser Ser Leu Pro Ala Phe Pro Gln Pro
 230 235 240
 Ser His Leu His Pro Gly Arg Cys Gly Gln Arg Glu Trp Gln Trp
 245 250 255
 Lys Arg Gln Arg Pro Leu Ala Asp Ala Trp Pro
 260 265

<210> 25
 <211> 246
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7500925CD1

<400> 25

WO 02/079441

PCT/US02/09820

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

<210> 26
 <211> 7633
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <223> Incyte ID No: 7757335CB1

<400> 26
 tgcagtaccy tccgaattcc cgggtogacg ggggcagtc ccagagtaag cagctagcag 60
 gcttagattc aggcctcag caaacaagga acctggaaaa tgtaaccctg aatgcacggt 120
 ggggaggaca tggcaagaga aaagcggcag gaataaagt attttctgaa tggcttctgt 180
 tctttgtota tcttctctga ttttctctg tctgacctt tcttggttaa aaatctgggg 240
 gaaaatgacg gactccaagc cgatcaccaa gagtaaatca gaagcaaacc teatcccgag 300
 ccagagcccc tttccagcct ctgataactc aggggagaca ccgcagagan atggggaggg 360
 ccacactctg cccaagacac ccagccaagg cgagccaagc tccaacaag gcccacaaga 420
 tgcggtctgg cggagaact cctaccacc ctcccaccag aagccccaa gaaaccacct 480
 ttctccagt gacgcagcac cctcccaga gottcaagcc aacgggactg gacacaagg 540
 tetggaggcc acagatacca atggcctgtc ctctcagcc aggccccag gccaagcagc 600
 tggctcccc tccaagaag acaagaagca ggcaaacatc aagagccagc tgatgacaa 660

WO 02/079441

PCT/US02/09820

```

cttcaatcctg ggcctottttg atgaactactc ctcogacgag gactctgttg ctggctcacc 720
tegtgagctc acccggaagg gcagccgggc cagcttgggg gccctgtccc tggaggctta 780
tcctgaccacn ggtgaagctg agacccegtt ccccaactatg aggcgagca tgtcgggact 840
ccacctggtg aagaggggac gggaacacaa gaagctggac ctgcaacagc actttaccgt 900
ggtctctccc gctgagtttg tcacacgctt tgggggggat cgggtcaatg agaaggtgct 960
tattgccaac aacgggattg ccgcccgtgaa gtgcattgct tccatccgca ggtgggecta 1020
tgaatgttc cgcaacgagc ggcgcaatcg gttctgtgtg atggtgacc ccgaggaact 1080
taaggccaac gcagagtaca tcaagatggc gnatcattat gtcccctcc caggagggcc 1140
caataacaac aactatgcca acgtggagct gattgtggac atgccaaga gaatcccctg 1200
gcaggcgggt tgggctggct ggggccatgc ttcagaaaac octaaacttc cggagctgct 1260
gtgcagaact ggggttgctt tcttaggccc tcccatggag gccatgtggg ccttaggaga 1320
taagatgccc tccaccgttg tgcocagac gctacaggtc ccaacctgct cctggagctg 1380
aaqcggccct acagtgagtg ggcagagaag tyatctgag cagggaanaa gactcagctg 1440
cccagaagat gtttatgaca aggttttctt gaaagacgta gatggggct tggaggcacc 1500
agaaagattt ggttttccat tgatgatcaa agcttctgaa ggtggcggag ggaagggaat 1560
ccggaaggct gaggatggcg aggacttccc gatccttttc agacaagtac agagtgagat 1620
cccaggctcg cccatcttcc tcatgaagct gcccagcac gccctcacc tggaaagtca 1680
gatcctcctg gaccagatg ggaatgctgt gtctctgttt ggtgcgact gctccatcca 1740
gcggcggcat cagaagatcg ttgaggagc accggccacc atcgcccgc tggcctattt 1800
cgagttcagt gacagtgct ccatccctct gggcaagacc gtggctatg tggatgcagg 1860
gaagtggaat tacctctata gtcaggatgg cagcttccac ttcttggagc tgaactctcg 1920
cttgagggtg gaaatccctt gcaacagaat gattgtctac gttaatctgc cggcccccac 1980
gtcacagctg gccatggggc tgcactgca cgggtgaaq gatatacggc ttctgtatgg 2040
agagtcacca tggggagtgat cctccatttc ttttgaacc cctcaaaac ctcocctcgc 2100
ccgagggcac gtcattgccc ccagaatcac cagcgaanaa ccagaagag gttttaagcc 2160
gagctccggg actgtccagg nactgaattt ccggagcagc aagaactgtt ggggttactt 2220
cagcgtggcc gctactggag gctgcaagc gtttgggat tcccaattg ggcactgctt 2280
ctcctgggga gagaaccggg aagagcccat ttcgaacatg gtggtgctt tgaaggaact 2340
gtccatccga ggtgacttta ggaactacgt ggaatacctc attaacctcc tggagaccga 2400
gagcttccag aacaacgaca tgcacacgg gtgttggac tacatcttg ctgagaaagt 2460
gcaggcccgag aaaccggata tcatgcttgg ggtgttatgc ggggcttga actgagccga 2520
tggagtgctc agaactgca tgcagatttt ctacactcc ctggaagggt gccaggtcct 2580
cccaggggat tcaactctga acctcgtaga tgtggaatta atttaaggag gttttaagta 2640
cattctcaag gtcgcccggc agtctctgac catgtctgct ctcatatga atggctgcca 2700
catcgagatt gatgcccacc ggtgaaatga tggggggctc ctgtctctct acaatgggaa 2760
cagctcaacc acctcaatga aggaagaggt tgaagtttac cgaattacca tggccaataa 2820
gagctgtgtg tttgagaagg agaacgatcc tacagtctct agatccccct cggctgggaa 2880
gctgacacag tacacagtg aggatggggg ccacgtttag gctgggagca gctacgtgta 2940
gatggcggtg atgaagatga tcatgaccct gaacgtttag gaagagggcc ggttgaagta 3000
catcaacgct ccaggtgccc tgcctgaaac agctctgctg gtcggccagg tggagctcga 3060
tgaacctctc aaagtccacc cggctgaaac gttcaacgga gaactccctg cccagcagac 3120
actgcccac ctcggagaga aaactgcaaa ggttttccac agcgtctctg aaaaacctac 3180
caactcctg agtggcttct gctgcccaga gccogttttt agcataaagc tgaaggagtg 3240
ggtgcaagag ctcatgatga cctccggca cccgtcacty ccgctgctgg agctgcagga 3300
gatcatgacc agcgtggcag gcgcacccc cggcccctgt gagaagtctg tcccgagggt 3360
gatggcccag tatgcccaga acatcacctc ggtgctgtgc cagttcccac gccagcagat 3420
agccaccatc ctggactgcc atgcagcaac cctgcagcgg aaggctgacc gagaggtctt 3480
cttcaacaac acccagagca tctgtcagtt ggtccagaga taccgcagcg ggaaccgctg 3540
ctatatgaaa accagtgtgt tggatctctt gagaagatc ttgctgtgtg agcaaccatt 3600
cagcaagaca agagatgctg atgccaacac cagtgggatg gtcgggggag tgaagagcct 3660
gagctttacc tctgtgtggt gtttttctcc ccagcccacc taagcaagat gttgtgataa 3720
cctcagggag cagtccaagc cagacatgct ccaggtgctg gactgcaatc tctcccaccg 3780
acaggtggcc aagaaagaac agctggtgat catgttgacc gatgagctgt gttggccaga 3840
cccttcccct tggagcagag tgatctccat cctcaacgag ctcaactcagc tgaacaaaag 3900
ggagcactgc aaagtggccc tcaagcccgc gaagatcctg attgctccc acctcccctc 3960
ctacagcctg cggcaaaccc aggtggagtc cattttcctg tctgcaattg acatgtacgy 4020

```

WO 02/079441

PCT/US02/09820

```

ccaccagttc tgcocagaga acctcaagaa attaataactt tcggaancaa ceatcttcga 4080
cgtccctgct actttctctc atcacgcaaa caaagtcgtg tgcctggcgt ccttgagggt 4140
ttacgtgggg aggggctaca tcgctcatga gttaaacagc ctgcagcaac ggcagctccc 4200
ggacggcaac tgcgtggtag aatccagtt catgctgccc tccctccccc caaacoggat 4260
gacgtgccc atcagatca ccaaccctga cctgctgagg cacagcacag agctcttcat 4320
ggacagcggc ttctccccc tgtgccagcg catgggagcc atggtagcct tcaggagatt 4380
cgaggacttc accaagaatt ttgatgaagt catctcttgc ttgcacaag tcoccaaga 4440
cacccccttc ttacagaggg ccgcgacct cctatactcc gaggatgact gcaagagcct 4500
cagagaagag ccaatccaca ttctgaagt gtccatccag tgtccagacc acctggagga 4560
tgaggcaact gtgcgattt tacggacatt cgtacagtc aagaaaaata tccitgtgga 4620
ttatggactc cgaagaatca cactcttgat tgcacaagag aaagaaattc ccaagltttt 4680
cacatccaga gcaagagatg agtttgcaga agatcgcatc taccgtcact tggaaactgc 4740
cctggccttc cagctggaac ttaaccggat gcttaacttc gatctgaccc cctggcctg 4800
tgcaaccac aagatgccc ttaccctggg tgcctgcaag gtagaggaag gctggaaggt 4860
gacggaccat aggtttctca tccgcgcaat catcaggcac tctgacctga tcaaaaagga 4920
agcctcttcc gaatacctgc agaaccaggg tggcggctg cctcctgggg ccatggacga 4980
gctggaggtg ccttcaata acaccagcgt ggcaccgagc tgcaaccaca tcttctccaa 5040
ctctgtgcc actgcatca tggaccocct caagatcgag gaggctcgtc gctacatggt 5100
tatcgctac ggcagccggc tgggaact cctgtgcta caggctgagg tcaagatcaa 5160
catccgcaag acccaaccgc gcagtgcgct tcccctcgc ctgttcatca ccaatgagtc 5220
gggtactac ctggacatca gctctcaaa agaagtgact gactccagat ctggaatat 5280
catgttccac tctctggcca acnagcaagg gccccagcac gggatgctga tcaatactcc 5340
ctacgtccc aagatctgc tccagccaa gcgatccag gccacagacc tgggaaccac 5400
ctacatctat gacttcccgc aatgttccag cagcctctc tttaaactgt ggggtccccc 5460
agacaagtat cccaaagaca tctgacata cactgaatta gtttggact ctcaggccca 5520
gctggggag atgaaccgac ttctgtgtgg aatgaggtg gctcgtgtg ccttcaaat 5580
gaggtttaag acccagagat acccgaagg ccgggatgtg atgctcctg gcaatgacat 5640
cacctttcgc attggtact ttggccctgt agagagacct ctgtacctg gggcatccga 5700
gatggccggc gcaagggcca tcccaaaat ttaactggca gccacagtg ggcctctat 5760
tgcatggca gaggagatca aacacatgtt caactgtgct tgggtggacc cagaagacc 5820
ccacaagga ttaaatacc tgaactgac tcccagacc tacaaccgaa tcaactcct 5880
gaactcctc cactgtaac acatcgagga agggagagag tccagatca tcatcagga 5940
tatcatcgtg aagatgatg gcttggcgt ggagactctg aggggctcag gcatgatgc 6000
tggggagtc tctctgctc acgaagagat cgtcaccatt agcttggta cctgcccagc 6060
cattgggatt gggcctact tggtaggct gggccagcga gtgctccag tggaaattc 6120
ccatcctc cccacaggg caagtctct caacaagctc ctgggaagag aggtctaac 6180
atcccaaac cagctgggtg cgttccgat catgcaaac aatggtgtc cccatccac 6240
cgtgccagat gactttgag gggttatac catcctggag tggctgtct atatgcaaa 6300
ggataatcac agccctgtcc ctatcctcc acccactgac cccatgaca gagaattga 6360
attctccca tccagagctc cctacagacc ccggtggatg cttgcaaggaa ggcctcacc 6420
aactctgaag ggaactggc agagcggat ctttgaccac ggcagttca aggaatcat 6480
ggacccttg gcgcagacc tggtagacag ccagcaagg cttgggggga tcccctggg 6540
agtgattgct gtagagaca ggcctgtgga ggtggcagtc cctgagacc ctgccaacct 6600
ggattctgag ccaagataa ttcagcaggc agacaggtg tggttcccag actcagccta 6660
caaaaccgcc caggccgta aggacttcaa cccgggaag ttgcccctga tgaactttgc 6720
caactggagg ggttctccc gtggcatgaa agacatgat gaccaggtgc tgaagtttgg 6780
agctacatc gtggacggcc tttagacaata caaacagccc atcctgatc atatcccgc 6840
ctatgggag ctcggggag gctcctgggt ggtcatgat gccaccatca acccctgtg 6900
catagaaatg tatgagaca aagagagcag ggttggctt ctggaaccag aggggacgt 6960
ggagattaag ttccgaaga aagatctgat aaagtccatg agaaggtac atccagctta 7020
caagaagctc atgaaacag taggggaacc tgatctccc gacaaggacc gaaagacct 7080
ggaggccggc ctaaggctc gcgaggaact gctgctccc atctaccacc aggtggcgt 7140
cgagttgccc gacttccatg acacaccggc ccggtctgct gagaaggcgc tcatctgca 7200
catctggag tggaaagcc caccacactt cctgtattgg cgtctgccc gctcctcct 7260
ggagggcagg gtcagcagg agatcctgca ggcagcggg gaggtagtc actgctat 7320
ccagtccatg ctgctcgtc ggttcgtgga gaccgaggg gctgtaagg cctactgtg 7380

```

WO 02/079441

PCT/US02/09820

```

ggacaacaac caggtggttg tgcagtggtt ggaacagcac tggcaggcag gggatggccc 7440
ggcctccaac atcogtgaga acatcaogta cctgaagcac gactctgtcc tcaagacoot 7500
cogaggcctg gttgaagaaa accocgaggt ggcogtgagc tgtgtgatat acctgagcca 7560
gcaotcaagc coagctgagc gggocgaggt cgttcaectg ctgtctacca tggacagccc 7620
ggcctccaac tga 7633

```

```

<210> 27
<211> 992
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1482539CB1

```

```

<400> 27
atggcaaggc cgcagcctg cgggcccccg cacgctaggt gcggtctgcc gtcgctgccc 60
gagcggccgc tgcaggtgaa ggtggtgggg ctcttcagct gccccaactt tcagattgcy 120
aagagcgcgc ctgagaatct gaagaataat catccatcca aatttgaaga tccatatta 180
gttctctctc aagaatttgc atggcatcaa tatctacagg agaaaaaag ggaactcaaa 240
aatgaaacct ggaataatcc ttctctctgt atttcttttg ttaatggtca gtttctgggt 300
gatgcattgg atctgcagaa atgggcccac gaggtgtggg ataatgttga cattaaaccc 360
totgcaatlt atgacgcact cactgaggat ttttcogcta agttottaag agacaccaa 420
catgatttgg tgtttttgga caattgtatt gattctcttc caattggaag attgattttt 480
gagotatact gtgatgtgtg tcccaaaaac tgtaaaaatt ttcaggctct gtgcacagga 540
aaagcagggt tttctcaacg tggcataaga ctacattaca aaaattccat ttttcacga 600
atagtaacga atggctggat acaaggaggg gatatagtgt atggaaaagg agataatgga 660
gagtcgattt atgtccaac atttgaagat gaaaactttc cagttcctca taataaaga 720
ggagtacttg gaatggccaa caaagccctg cacagcaacg ggtcacaatt ctatacaca 780
ctgcaagcaa ctcttatctc agatagaaaa tttgtggctt ttgggcaact gattgaagga 840
acngaatgc ttaacaactc agaattagtt ccaacacaga atgaaagacc aatacatatg 900
tgtagaatla ctgacagtg agatccttat gcttgatttt catatcaata ttttctgtga 960
taatttatta ctttctatct gatccagctgt tt 992

```

```

<210> 28
<211> 3788
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2234213CB1

```

```

<400> 28
cagggttagt cctcaccggg ganagacgaa caetggcctg acctgggct ttagaaacc 60
atggacaaca ggaagggaagc ttattaaccc ctctcttccc cctcctctct ctcccggty 120
gagaatcaagc tctctctctc cccggtctgt agaccctggt gcaaaatgca aacaaccgcc 180
cgtcttcccc acctctcgcg ctctctcccag aggagggaga gggaaagggc gagggcgcag 240
ccccaaaccg agtccagcaa gtctgtcaat cggacgtgaa cgtccccagc agacaacttg 300
aacctcaggg cctctcgccc agccggacac cgcaggtacc ccttttaact gctctctctg 360
tgggttttga tccgataccc caactcctag cctcatittc ccccagccc cataatattc 420
cggacccctg ctccccgtgc aggtgggtgca aacgcccctg ctctccggcg ccccctgccc 480
ggaagagggg aggcggcgcg gcccgggag gagcaggacg cttygtccgg accggagctcg 540
gctctgggaa gaagccggga gcttccctga tgytyccgoc gctccgagc cggggagag 600
ctgcccaggg coagctggcg aggagcctgg gtcogctgct gctgtctctg gcttgggac 660
acacgtggac ctacagagag gagccggagg accggcagcg sjaaatctgc tcagagagca 720

```

WO 02/07941

PCT/US02/09820

```

aaatcggcag gactaaatc cgggtctgta agtcttcagg cgaagctaac acatgctaca 780
ggaaaaagtg ctgcaaaagg tataaatttg ttcttggaca atgcatccca gaagattacg 840
acgtttgtgc cgaggctccc tgtgaacagc agtgcacagg caactttggc cgaagtgtctg 900
gtacttttta tccgggatac cgalatgacc gggagagaca ccggaagcgg gagaagccat 960
acgtctggga tattgatgag tgtgcccaga gcaatgggac gctgtgtgco cacatctgca 1020
tcaatacctt gggcagctac cgtctggagt gccgggaggg ctacatccgg gaagatgatg 1080
ggaaagcattg taccagggga gacaatatc ccaatgacac tggccatgag aagtctgaga 1140
acatggtgaa agccggaaact tgcctgtgcca catgcaagga tttctaccag atgaagcaga 1200
ccgtgtctga gctgaagcna aagatgtctc tgcctcccaa caatgcagct gaactgggca 1260
sgtatacacc tggtagcaag tgcctggcct caaacaccta coltccagga cctctctggc 1320
tgcctggggg ccaggggcoot ccogyctcac caggacccaaa gsgaaagccca ggcctccccc 1380
gtatgcccagg cctctctggg cagcccgccc cacggggctcc aatggggacc atggggccat 1440
ctctgatctg gtcccacatt aagcaaggcc ggaggggccc tgtgggtcca ccaggggcac 1500
caggaaagaga tggttctaag ggggagagag gacgcctcgg gccacagagg tctccaggac 1560
ccctcggctc tttgacttcc ctgctactta tgcctggctga cctccgcaat gacatcaactg 1620
agctgcagga aaagggtgtc gggcaaccgga ctcaactctc agcagagagg ttccctttac 1680
ctcaggaaatt tcccagctac ccagaagcca tggacctggg ctctggagat gaacctccaa 1740
gaagaactga gacaagagac ttgagagccc ccagagactt ctacccatag cacatcccaa 1800
ccocgtcagc ccaaaaggag agaaagatca actcaactgc agttaaacca totaaagaga 1860
agaaagacca ctgagagacc agaaaacata calttttctc ttctctctc ctgacgtctc 1920
tccactccct ttctccaaa tacgatgcta ttttcaagat cccctctag gccctgcagc 1980
atgaggggagt gaatgatgta ttcaactgct tctcaactag agtccattgg ggtggtttgc 2040
attgtaactt ttcttttaca tctatltttt ccaggaactt lggattlaag tactctcaca 2100
gtgtcttaaa tcaataatcc ttgaagttaa atttggcaga gtatcaaaag ggggaaaaatg 2160
acaaagctag ctctaagaaa atgtgaggct acttctaaga tgtgtgttca caatagacca 2220
taactctctc agtacaataa ttgggctctc tcagttaaaa aggggtgggg aggcacaaagc 2280
tgcctgattg ctttggggga gaatttttcc ctgtgtcttc tagtagactt taaatattgt 2340
atccctttgt caaaccttgt ttcccacatt caattaaaga gaggagagaa ttgaatggcg 2400
ttttagaagag atagaaaaga atcacagtca tatatttact ttatataga ttgccacatt 2460
ctaaaaatca aatcagctgc taaaglttcc atgcccctct tatctgtaag tatctattt 2520
agggaaagag attaaactct cttttcaaaa aaacaaagt aaatgcctgg atccacatta 2580
aaacaatggg ctctcgtttg ctataatatt ttaaagctgt ttaatcaaca gttggagtctg 2640
cctatanaat atagattatt tgttcaataa actggctcag cttagagaga ggtgcagaat 2700
tctctgttct gagcaggtgc ccagaaggtta ccattaggtg ccattgacca ggcctgaacca 2760
atatacagtg ggcctgaagt ctgcaaggag gtgtgctgct tgggctgacc tcaactatgc 2820
cacagcagc ggtaggtaaa ttttttctcc ttgggtatta caagtttttg tctggagcca 2880
accaagcttg ccacccacat atlgagagta atacactatt gaagttatc ttggtgggg 2940
agaaaaaaat atagtgtttt tctctgtttg caaaaactc ctctctattc tcaatttttc 3000
tcaattttct tcaatttagt ccaagtccca gttcttttag gccttctctt tgatttatt 3060
tccctctcat gtagaagca gttcagaaaa aggtctatat ctccaactcc tagtgagta 3120
gagtgcttcc tcaagaccoc ttctgggtggc aaagggaaag atgttctctg caaggtttgc 3180
tgtggattca gaagccagag gacgaagaga ccagaagat gatctgtccc tttgtaactg 3240
tgttgggggc cctctgtttt ccaatgagca gctlataggt tactcaactg ccactttctc 3300
actggacaca caaagtggct ctttatctac ctttgaggga gatlttcaet ctctgcacaa 3360
tgaactgtct cacactcata tttagctcat ttggaatttc ccatcctgcc atgtctcttc 3420
ccattttttt ttggcttttt tgcctccacc ttttagccca catcatltaa ctccaactac 3480
gtgaagcctt gottaaagaa aatccctctt ggccgggtgt gtagccacc gccctcaate 3540
ccagoccttt gggaggctga gggggggaga tcaacaagtc aggagatcga gaccagcctg 3600
accaacatgg tgaacccctg tctctactaa naatacaaaa attagctggg cgtgttggca 3660
cacaccctga atcccagcta ctcaaggagc tgaggcagga gaattacttt aacctgggg 3720
gggagcctag attgcgctac tgcactccag octagggcaac agagggagac tctgtctcat 3780
tcaaaaaa 3788

```

<210> 29
 <211> 2443
 <212> DNA

WO 02/079441

PCT/US02/09820

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1345785CB1

<400> 29

```

ataccgactc ggatcaacttg tacggacgca gtgtgctgga aagcaggtgc cagcgggggc 60
ggtcccccaga ccgtagacag gccccagcct ccgtgatgtc acgcccgggt ctaaggaggg 120
ggagggggta ggcctgtkttt tctggasaga gaactagagc cgaatgggac aaagcctggg 180
actggggggg ggcctatggcg ctgccaatcc gaatcctgct ttgaaactt gtgctctctc 240
agagctctgc tttctcctg cactcagggg octcggtaac cgcctctgct ggcagctccg 300
tggtgtccga gtccggggtg agctgggagg cgggcgcccg ggcgggtctg cgtctccaga 360
gcccggcat ggtgtggacc caggaccggc tgcaagaccg ccagcgcgtg ctccactggg 420
acctggcgtg ccccgggggt gggcccggcg ggcgctgct ggaactgtac tggcggggcg 480
agcagcgtgt gtacagaggg cgggaccggc ggcgctgga gctctcggcc tcggccttgc 540
acgacggcaa ctctctctg ctcatcggcg cgttgaggga gacggaacgc gggctgtaca 600
ctcgaacct gcaacctcac tactgccacc tctacgagag cctggccgtc cgcctggagg 660
tcaccgacgg ccccgggccc accccggcct actgggacgg cgaagaaggag gtgctggcgg 720
tgccggcggg cgcaccggcg cttctgacct gctggaaccg cgggcaacgt tggaccgacc 780
ggcacgtgga gtaggtctaa caggtgtgtc actgggacgc cagcgcgccc ggggtccgcg 840
acgaccggcg ggaacggcctg ctggacctct acgctcggg cagcgcgccc gcctacggcg 900
ccctttttct ggcgacccgc gtggctgtgg gcgcgatgc ctttgagcgc ggtgacttct 960
cactcgttat cgaacggctg gaggctgccg acgagggcac ctactcctgc caactgcacc 1020
accattactg tggcctgcac gaacggcggc tcttccacct gacggtcggc gaaccccagc 1080
cggagcggcc ccccgggggc tctcgggcca acggtccag ccacagcggc gcccccagcc 1140
cagaccacca actggcggcg ggcacaacg tcatcaatgt catctcccc gagagccgag 1200
cccactctct ccagcagctg ggtcaagtgc tggccacgct gctgtcttct atctctctac 1260
tggtcaactg cctcctggcc gcccgcagge gcgcgggagg ctacyaatc tcggaccaga 1320
agtcgggaaa gtcaaggggg aaggatgta acttggcggg gttcgtctgt gctcagggg 1380
accagatgct ttacagaggt gaggacatcc agctagatta caaaaacac atcctgaagg 1440
agagggcggg gctggccacc agcccctctg ctgccaahta catcgacct gcaaaaggt 1500
tcoggaagga gaatgcmeta tagggaggcc ctgggtctct ggcctggcca gcagctgcac 1560
ctctctctgc tgtctctctc gggcatctc ctgatctctc ggggtcacc cccctccag 1620
cggctgttgc cgtttctctg gaatttgccc tggcgtatg cagagccgcg ctccacacc 1680
ctccccaggg gcttgggtgg cagcatagcc cccacccttg cggcctttgc tccaggggtg 1740
cccctgcccc ccttgccaca accaaatccc cactgatgcc catcatgcc tcagacctt 1800
ctgggtctcg cccgtggggg gctgaagac attcctggag gacactccca tcagaacctg 1860
gcagcccaca aactggggtc agcctcagg agcaggtccc actcctccg ggcctctctc 1920
gtccggggtc gggagatgtt cctggagggg gacactccca tcagaacctg gcagcttga 1980
agttygggtc agcctcggca ggaatcccaa tccctctagg gtgctgctg ccaccgaag 2040
ctccccacc tttaccacca tglgggactc caggcaacct ctgtctccc cagggcaactg 2100
ctgacttgaa tgcagccctt tgcctctctg tgttctcttg ggcaccctgg ggcctgacc 2160
cctgaccttt ctctgcccc aacctacct agccttctc tcagcaccct tgatagtcac 2220
tgggtccctt gtgactctg acctgacac cctccctctg gactctgctt gggctggagt 2280
ctagggctgg ggtacalctt ggtttctgta ctgctgagg acagggagg gagtgaagt 2340
ggtttgggtt ggcctgtgtt gccactctca gcaccocaca ttgcatctg ctggtggacc 2400
tgccacactc acaataaagt ccccatctga tttttaaaaa aaa 2443

```

<210> 30

<211> 957

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

WO 02/079441

PCT/US02/09820

<223> Incyte ID No: 3807190CB1

```

<400> 30
ttatecctga ttcotttctg tccctcaatg aagccatcag caaatgctct tgattctaac 60
ttcaaaaaga tatccaaagt ctgaccaaatt ctgtcaatct caattgctac caccctggtc 120
caggctagct ctcttgggtt atgttagatg ccgcctatct ggtctccotg tttctectt 180
tgctctctt cagtccaitt tcaacacagc agctggtata agcctgtgaa aatgcaagtt 240
ggatcctgc acttatgggc caagtgctac agtgggttec tgttaacct cgttaaagct 300
gatctctca gaagacttat taataaatac cctgttaaat ctggcctca cctccccctt 360
tctcaactgg ctcaacct actggcctct ttgtctact ctgttgccca ggctggagtg 420
cagtggcggc atcttggctc acttcaaat ctgcctccca ggttaagca atcctctcac 480
ctcagcctcc caagttagctg gtattacagc tgcgcacctc catgcttggc tgatctttgt 540
attgttagta gagacaggat tcaaccatgt tggtaagact ggttttgaa cctcgaacct 600
aagtctgcca tcaacctggc cctccaaaag tctggggatt acaggtcatg gccactgtgc 660
ccaggttcc ttgttcttta caccagaat aacctctac ctgctggctt tatttctgt 720
tcaactctcc tagaattgctc ttctctcaat gcagtagact caatgtttat gctcccacc 780
ctgcaacccg gccaaatcat atgttaaac ctacccccct gcccgggcgc ggtggctcat 840
gcctgtaatc ccagcaacttt gggaggctga ggcagacaga tcaagaggtc aggagatgga 900
gccatctctg gtaacatgg tgaacctctg tctctactaa aaatacaaaa aaaaaaa 957

```

```

<210> 31
<211> 674
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4856078CB1

```

```

<400> 31
tctggaagca ccatgttctt ttgagtcaga ggtcacattt tgcacattt gttttccgg 60
aattcgggta tatcctgaga ctagatgttc acgttoatgt gctgctggtg gggccctgaa 120
tgctctctgt ccttgagctg tctctgggoc tggttctcag gatggtctct ggcagccca 180
tctctcttag cgcctgctt ggaacacagc ctcttgggtt ccaaaaatct cctgctcatg 240
ggttcagcca gccagtgagg atggagagag cctgggtgctc tcaacagagg gccaggtgcc 300
tggagagag gagaggttct tgcaggggag agtgtggcca ggtctctggg agccaccagg 360
tgctctctct gtgcccctgc aggggtctca gccctgagcc acccaagagt acacacagtg 420
aagccaggtg ggttagagaga agtccctgag cctgggacat gctctgaca caagggccca 480
gccagacctt ccaatccctg sagccacagg stgggggggc gggagtgca gcttggcaga 540
cagacagagg tgtgtccccc tcccttcca agtgcagggg gaggggcctt agggacaggg 600
gaacagaagg gccagggctt cgcctctgct ggcctctctt cctctctctt ctagctctctg 660
ggttgaaagg gagg 674

```

```

<210> 32
<211> 559
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5106886CB1

```

```

<400> 32
atgcagaagt caagycacag ggatgtgaac atacgtgggt catgtttcat agtacctgga 60
ccaagagaag aggtctctct atgttgagaa aatggcaatg cctccacaca gcatggtact 120
tcttctctct ctgagaatga tactgttaca taaaggtttt cggcgggtgc ccagaacacg 180

```

WO 02/079441

PCT/US02/09820

```

gagatttcc tctgcccagc tgcattgaggc acacatatcc tgggatggca ttaggaagat 240
ggccaacttc acatcaactt atgtactctg gcaccaacct ctcttctccc taccagagcc 300
aacttcccac ccactctctat tgggatgata tctgctcttg ggttaagaaat gcatatccaa 360
atgctacctc agtcagggga ggttaatactt ggcctcaatt ctctgcccctt agaattctgc 420
agccaaattg ggttaagaggt caagaccactt cgggtgtgctt attcagagga cccatctcac 480
gtgcaaaagc acacatgggc ccaaaaataaa gggatggagt aatatttacc aaggaataaa 540
gcacagacaa caacaaagc

```

```

<210> 33
<211> 2600
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5627037CB1

```

```

<400> 33
cggcgccagc ggggggggac agcatgctga gcctctctgt ctggatctct actctctccg 60
atactttctc ccaaggagac cagaccctgt tcagccagga gccagctgac cagacggtgg 120
tggctgggaca gggggccgtg ctccctctgt tctgtctcaa ctactctgga attgtgcaat 180
ggaccaagga cgggctggcc ctgggcatgt gccagggctc aaagcctggc caaggtaccg 240
ggttgtgggc tccgcagacg ctgggcafta caacctggag atcacagatg ctgagctctc 300
tgacgagccg tttacagagt gccagccca cggagccgoc ctgctctctc ggggggcaaa 360
actcaccgtg ctcatccccc cagaggacac caggattgac ggagccctg tgattctact 420
gcaggcagcg acccccacaa acctcacatg ccgggctctc aatgcaagc ctgctgccc 480
catcaatctg ttcggggagc ggcgcagcga gggggcgctc gtggccagca cgaattgct 540
gaaggtctgg aagagggaga ccactctgag ccaactgctt attaacccca cgaacttgg 600
catagggcgt gcttcaactt gccaaagcat gaaogaagoc atcctctagt gcaagagagc 660
ttccatcgag ctggtgtgtc accaccctcc tacagtgacc ctgtccattg agccacagac 720
ggtgcaaggc ggtgagcgtg ttgtctttac ctgccaggcc acagccaacc ccgagatctt 780
gggtctcagc tgggcaaaag ggggtttctt gattgaagac gccccagaga gtctctatga 840
gacaaatctg gattattctt ttttcaagca gctgtgtct tgtgaggttc acaacaaagt 900
gggaagcacc aatgtcagca ctttagtaaa tctccacttt gctccccgga tctagttgga 960
ccccaaacc acaaccacag acattggctc tgatgtgacc cttaactgtg tctgggttgg 1020
gaatcccccc ctcaactctca cctggaccaa aaaggaacta aataggggac ccagcctcc 1080
tggctcccca ccggagcctg ctctctctgc ccaggctctg agtaacagca accagctgct 1140
gctgaagctg gtagctcagg cagacgctgt cactcaacc tggcgggca cgtgtctctg 1200
aatcggagtg gctgagcggg aggtgcctgt ctatgtgaac gggcccca tcactctcag 1260
tgagcagctg cagtatctgt tgaggggtga cgggtggca gttgaggttt ctattggggg 1320
cacccaacc ccagaccgca tagctgggc ctggaaggag aactctctgg aggtggggac 1380
ctggaagcgc tacaagtg agagaccaa ctccggcagt ggggtgctat ccagcctcac 1440
atcaacaat gtcattggag ccgacttca gactcactac aactgcaag cctggaaacg 1500
cttggggca ggcacagca tcaatcagct ggaagagcga gagggtgtac ctgtgggcat 1560
catagctggg gccaccatcg cggcagcat ctgtctctac ttcttctca tctgcttgg 1620
atctctctc taccgggccc gcaagggcag tgcacaagac gtagccctga ggaagctga 1680
tatcaaggtg gagcagctga accagagacc acttaagatg cactctgacc gggaggtga 1740
cacggccagc gctctccagc caaccgggt catgaagccc atctactgt cgtttaagga 1800
tgatgtggat ctgaagcagg acctgctgtc cgaaccatc gacaccggg aggagtatga 1860
gatgaaggac ccaaccaatg gctactcaaa cgtgctgtcc catgaagacc gccgctctc 1920
cagggcagtg ctctatctgt actaccgtgc cctgggctct gccctctg accgccc 1980
ctcatccctg ctctccactt ccagcggcta tggccagctc aacacctata gccggggccc 2040
tgctctctac tatggccctg agcccaacc cctgggctct gctgcccag ctggcactga 2100
ccaaccagc cagctgtctc acgagaacta tgagaagttc aactccatc ccttccctgg 2160
ggcagctggg taccaccctt accgactggg ctacccccag gccaccctt ctggcctgga 2220
gcggaccaca tatgagggct atgaccccat tggcaagtac gccacagcca ctgattctc 2280

```

WO 02/079441

PCT/US02/09820

```

ctaacctcc cagcaactgg actacggcca gogattccag cagcgcgatg agactcaagt 2340
gtaggggcca gagcctggct ggggcattct tgcggggcag aggagaagcc tttaacagct 2400
gttccctgat attcaggggc attgctcatt gctccctctt cggacaagcc ttcttctccc 2450
caccatggca ggtggggagc aggtctccca gaacacccc gtcagaggat ggtcctctgt 2520
gcatgcccc aactcctggg gctgacattc ctcttcttgg ctaggaagtg tcccttctga 2580
actggactct tggctgagca

```

<210> 34
 <211> 788
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3688835C21

```

<400> 34
ctaaatgccc aatctgacca ccattgctctg tctacctaa gctcctctgca gtgcccggct 60
gtggggcagg cctctcttgc ttctggagag gccactgggc cctgcagtgt ctgccaacct 120
tctctctacc cttagaatga ccgtccatct tctctctctg acgtgctccc accttctctg 180
tgaccttagc taccccaca ctcccacagg atcaggcagc ccccctaccc ctgctcctct 240
ctctgggcat tcaattctct ggagaaattg ctctctcagc tgettcagga accctttcca 300
tgacaacgac attggccctg agatcttccc catcagctct caaaaagagt ttctattttt 360
tcctctcaaa actgacttca taattatttg tgaccgtggt gccaccattg ccccacgac 420
tcaacctgaa acccaactctc tggctgtgca agaattgctt ctccaacca ttctccagcc 480
cggcagtact gagaacttcc ccattcccgt tcccagact ctgtcgtggt atcgtacaca 540
tgctattggt atgtgtctat gctctgcttt agatattca ttcataagt atttattaag 600
cactgattct atgcccaggcc ctacagatca cagattaaa agcacaatac gtgctttaag 660
gactcaaaa aataggggag aagacaaaac ttcatttaat gacacaagag tgtgatgtct 720
acaacagtat aggtatggac aagtaactgt ggtggcccaa ggaagggagt gatcacacga 780
gtgggtaa

```

<210> 35
 <211> 702
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2295419C21

```

<400> 35
agacagggga gtatctgggg aatggccgct agtgccctgg ggagccaggc gaggggtagg 60
cagcgggagg tccctttgga ggaagagctc agaggtctgg gatgggtccc caacctgctg 120
cagccctcac acttcccctc ccaagcctca gcttcttccc gagcaggcag tggtagaggg 180
gggtgcttag agatagggaa actgaggtca ggcaggggct ggtccatagt gcagctctgg 240
gctcttggcc tcttcttggc cctctgtctt tctctctctg cctctgctct gctaaccccg 300
atccccccac tggcattggc ctggcctctt tacacccctc agcctgtctc tccatcttta 360
tcttttctct ctattagttt ttgttctgtt gtgtttttta ttttttaata aaaaagtttt 420
tttgggactc ggggtggttg gctcaccctc gtaatcccag cactttgtgg ggotgagcca 480
aatggatcac ttgagcccag gagttcgaga ccagtctggc caacatggtg aaacctcgtc 540
tctaactaaa atacaaaaaa ttactgtggc ttgggtggtg atgctctgta ttccagctac 600
tcaggagcct gaggcggaga atgctttaa cctggggagg agaggttgcg gtgagccag 660
gacagcccgg gtagacagat aagactctgt ctcaaaaaaa aa

```

<210> 36

WO 02/079441

PCT/US02/09820

<211> 806
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2437896CB1

<400> 36
ggctggagct angaactata gaactatagc tggacagggg gggccttctg gtcctagggg 60
gcaagaatga tttctggaca ctgttgacat tctaaccaagt cttctgtctg tggttaaggat 120
agaagggcag gacttcacaa cccgtagcat ggatgctggc ccagctgttg tctccctgcc 180
cactgccttc tgccttact tggctgggg gactgggaa ccagatgaac gtcctggagcc 240
ttccacttgc ctggctgtgt gtcagagga aatgggttta cagggaagag tgcattggaaa 300
atcaggtctg caggaaaaac catggtggc actgtggagc agacaattg cctcactctc 360
tcgtgatctg agtcatttgg gaaaggggaa aatgagtggt ggtaaatcac ccaaacatct 420
gcaggccagc agagtaactt tttctgacc tgacttacag ggagaacaaa ttttatatta 480
tgatccggtt tatacgtata tatggtttgt tttttgtgt atgttgtttg taaagccagt 540
gattaagttt gaggtttggc agttaggatt tgtttaaatt ccagcagatg gaactagaaa 600
agaccocatg tttccactgc ctcaagacc tgcctggatg atatatatt atctttgtat 660
ttatgtactg gataataatt agtgtatata gttacatct atagcatata tatactgtat 720
cagtggttaa aaatttcagc ataacctacg tatgtatgta tgaatataaa taaaaataaa 780
cataaggttt atgcataaa aaaaaa 806

<210> 37
<211> 1112
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2641482CB1

<400> 37
cgcccaccac ctccctcccc ctccctgccc tectgtctcc caggcaagccc tcaactaccc 60
ccccaccact ctaggccaag catccacctc cctgtcttaa ccagccaccc ccatccgccc 120
ccatcacaga ctanccagga gcacttttagc ccagctgctt cctcccccgc ccagcagctg 180
tcctgccttc ctgggttctct ggaacagagg gaggccctaa cctgcccatt ccagctccca 240
gagcaatacc ctccctctgg ttccttcaaa cccaagcccc tcttctcctt gacccacccc 300
atctccaccc accagtcctt cccttctct cagggtctgg gactcgtggg tgcagctggc 360
acctagctcc tgcctyacc cccaagccc aggggtctcy tgggaagggc tacagaagca 420
agcagatgag ggcagccat tccccgggc acctgtccgc cgtggccttc ctgcatctcc 480
cgcaggacat cagtgtccag cagacacttg acctgtcccc cgtgctggaa gggctggctg 540
tcagcactca cctgcctctg cagctccccc ggttgcctg cggagggctg cagaccaggg 600
ctcaggcagc ggcctctccc ctctggact tcagccaagg caggagggct gatggagtca 660
gtgtgacagc ctacataccga gccttgggag gtggtccttg tagtagaagc ccccctctgc 720
ctgcgccaca cacttgaggt ccacctgccc ctgttgccc acacggtaca cattggtggt 780
accatcagcc cagctcagcc tggccacact ccggcctgtc tccacatccc agcccaggt 840
gtccaccaca cggcccggtt tcccttccc tccctgggaa agaggccccc ccagcaggt 900
aaggctgagc catgagccag ccacagccca gccctgacc ccgccctctg cctccactca 960
ccatctctgt agcccaccc ccagtccggg cctgcaccca ccttctctcc ctggaagatg 1020
cccttagtg gcatcctctg gaggccctgg cggggactca gtgtgacact gcaagaggtg 1080
tcacaaataa agctctcaga agaaaaaaaa aa 1112

<210> 38
<211> 2401

WO 02/079441

PCT/US02/09820

<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7494292C81

<400> 38

```

ttcggctcga gggagggggt cgcgcgcgct accgcgcgag ccgcgcgagg gccgcgcgct 60
gggatgccga ggcgcgcggt cgcgcgctgc tctgtctccc gcgcgcgctgct cagctgaaag 120
cgcacaggat tcaattactg gaattgtcaa ctctgcacgt gtacgtgcaa ttctctcttc 180
actatgagag gaccgatgtt attgcaactt tgtctggott tctgtagcct tctgtctttc 240
agcgttgoca cacaatgtct ggccttcccc aaaaatagaa ggaggaggga gatagcaat 300
gttcattgag aaaaaggcca gtccgataag atgaacaacc atgaactaga aaatagctct 360
gttaactcaa agcagactcc ccaactgggt gctctgaaag atccaatgat gatgtcagca 420
gtaccatcgg caacatcatt aataaagca tctctgatta acaaaagaat ccagcctgga 480
caagctgggc tcatgcaaac agaacgcctt ggtgtttcca caactactga gtaaggtgtc 540
ccctcagctg aagaagtatt tggttccagg cagccagaga gaatactccc tgaagtgga 600
cttgccaagg coactttaac cactgtatcc actgcgactc cttctctgac tgttgatgaa 660
aaggaggaac tcccttaaac cactaaactt cagcccattg tagaagatc cacagaacc 720
acaaaagggt tctgaaagta tatggataat caatcatttg caactgaag tcaggaaaga 780
gttggtttgg gacattcacc ttcatctcat gtgaatacta aggaatgct aaccaccaat 840
cacaagactg agaaatttga agcagacaca gaccacagga caactctttt tctgtgtgt 900
gagtccacag caggcagtga gccctgaaag ctaccccctg ataaggagaa gccctgcgag 960
atgacagctg ataacacca ggcctgctgc accaagcaac cactcgaac ttcgagtac 1020
accctgagtg ttgagccaga aactgatagt ctgctggagg ccccagaag cacagtgagt 1080
gtcagcaacg ctgttccagc tgcctctgcc ttaagtgat agtgggatga caccaaatta 1140
gagagtgtaa gccggatag gacccccag cttggagaca atgaagagac tcagtgatga 1200
acggagatgt ctcagacagc acaagtaag catgagggtg tgaaggagg ccagccttgg 1260
caagaggctg cacagtggtc tctggggctg cctgaagggg aaacacacac gggcacagcc 1320
ctgcaaatag cgcattggaa tgagagatca cctgcttcca ctgatcaag ttcctttacc 1380
cccccaagtc tgatggaaga catgaaagt ttccatttga acttctcca aagtacggga 1440
gacttcaagg aatccacca ggaaaaogat gccctgtttt tcttgaaac cactgtttct 1500
gtctctgat atgagcttga ggcagcccaa ctgttggaaa atcaaatgaa agacatcctc 1560
actcaagaga tgcacaacagc tgttcaagag ccagatgcca ctttatccat ggtgacacaa 1620
gagcaggtgt ctaccctcga gcttatcaga gacagtggca agactgagga agaaaggagg 1680
gaccoccttc ctgtgtctga cgttctcgtt gttactcagc tgcacaagag atgggagcct 1740
ctggccaacta caatttcaac taagctgttc cctttgtctt ttgaagttae tcccactgtg 1800
gaagaacaaa tggacacagt cacagggcca aatgaggagt tcacaccagt tctgggatct 1860
ccagtgacac cctctggaat aatggtgggg gaaccacgca tttccctgc acttctgtct 1920
ttggaagcat cctctgagag aagaactgtt gttccatota ttactcgtgt taatacaget 1980
gctcatabg goctggacca acttgaatct gaagagcgg ggtttcccaa tgttgcccag 2040
gctcgtctca aactcctggg ctttaagaat ctgcctgcct gggcctcca aagtgttggg 2100
attacaagtg tgaactcctg caccacagca agaaaatatt taacagctg cctgaaatgg 2160
aaattgaaac caaaactttt tgcactcgtt gtttgaagtg gactatcttg tctcaagcct 2220
ttttatttcc tcatgaaagt accttctaaa tattgttaag gctacaatgg aggaattttt 2280
aaattattat atttactatt ttacagtgag tggaaaattt gaaatttagt aaaaagtttt 2340
caagaaattt ttatagatta ataattggag ttaaaaagtc atggcgatga ggcatttaac 2400
a

```

<210> 39
<211> 1575
<212> DNA
<213> Homo sapiens

<220>

WO 02/079441

PCT/US02/09820

<221> misc_feature

<223> Incyte ID No: 7636288CB1

<400> 39

```

cgaaccacag cagtcocact gtggcgaagg agccgaagla cgcacgaaag cgcggcccgg 60
tcgcactcag ggcggccacg aggtgcgcgc tgaacctctc cttaacatcc ttggsgtccg 120
ggacgtgtgc ttccgtgaga acagtgaata gabacatttg tgccggcctc gcctgggcca 180
cacaggcccc agcagggtgg ctgagcgctc gtcagggtcg cctggatctg gaacttttcc 240
agatgaacca tgtgacttc ccagtgctgc aaatttcaga acgcgtatt tatgtgtatt 300
tggcacactg actgtracgt gctgttcac tgaattcca tagaaactgg caccctctg 360
ctctgtgcta atggggagccc aacctgctg gccacagatg ctgctggggc ccttgggctg 420
cagccctctc tggccgaggg totcctgtgt ttgotgggca ggcacagcct gotgtggggc 480
agcagagtct gttccctccc agcttcacgc cctggagcc acccacaca caccctggct 540
gaagatctgc agagggcttg cctccatggt tctcagatgc acatgggggt ctcaggaggg 600
aaactccctc cacaggggag ctgtgggaac cccctgtcca gggctgtgc cttctgacgg 660
cacaccacag aaagccggga acagtgtgct cctccacaat cgggacgtga aggcacgtgg 720
tcctcatgca gccccagaca gacttctgtc ctccagaaca ataccgagag ggcgaggggt 780
gaggggtggg ggttggggggc agggacaga ctggaggcag caaggccagg aagttgctgg 840
gagggggcag ctccctgggc catggtcca gttggggag agtccactgt cccggcccca 900
cgtctctgcc tctctctggc agtccctccc caaggggagc ctacgtggc ctgcaacaac 960
aggcagctca cctggagca tccctggcac ctgggctta gagatcaggt gccatgaac 1020
gttgcttccc gtttttaaaa atgcacaga cagcagagaa gagcatcagg agacttggac 1080
tcgagaatat caagtacagg atataaaacc atagtgtcta accccaaggt ggtcacagtg 1140
ctgaactatg acgagggaac aagcacctcc aaaaggatgg aagatgtaaa aaagcaaat 1200
caaacattac gatggccggg caggtgggc ttacacctgt aatccagca ctttggagg 1260
ccgaggtggg cagatcaact gaggtcagga gttcaagacc agcctggcca acttggtyaa 1320
acctatctc tattaaaaat gcaaaaagt agctgggcat ggtggtgggt gotttgaatc 1380
ccactactt gggaggctga ggcaggagaa ttgtbtgaacc tgggagcgg aggtttagt 1440
gagctgagct tggggcagata taectcaagc ctggggcgag acagcaggat gtttcaaaaa 1500
aaaaaaaaa aaaaacccc aaaaacaaaa ggggaggttg gaaaagggag gggggcaaaa 1560
agaaaggggg aaagg                                     1575

```

<210> 40

<211> 998

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 8095391CB1

<400> 40

```

gggcaggtat attattaata ctgaaaacac caagcataaa tatcctaagt taatacttgt 60
tgctgggtag cacagggatt atatcagact ttcacaatcc ctatgactaa tagtettatt 120
tgtagaangc aaactggcca gaaacacaca cccccaacac aaacatgggt aaggtttaa 180
cccttaatta tgcttccctag caaaatgggt gctaatgaaa gccataggt ctagaagagt 240
ggctgctcag cagattgcgg ctgcctcaca gattgctgca gccctgcoag gcccaagatg 300
gaggtgtcgg aagcagccaa atctttctga ttgcccagtct gtttctcoag gaagctcat 360
tatatcoga aagatcttac tggggagat atccatccca tctcataagg taacttgrtt 420
ccagggtgct gatcctactg tgggatcatt cagaagttta atcaggtgct cctaaaaatg 480
tttttggaac ttccaagttc ttaactagaa ttccaagcca taggagacag gtatctaaag 540
tgccaagctg atccatltg gcaaaaaaatt tctctctgca agggagctac acatgaaact 600
aaaaaaaaa catgatgggg ggtgggcttt atacaagtt gagtggggt acaatggtac 660
agaggaatg ggaacttgyca atatatattt tctaaagttg ctacgcaactl ggtctacta 720
gcacacatca cctggaatgg ttaccaggac ctcaaaaggac tgcccacgg gctaaacagc 780
tyatccgtcc totgaagoca gacagtctta tctcctttaa gpatgccact 840

```

WO 02/079441

PCT/US02/09820

```

gattgagggc cggaaagctga acagagtgag ctccatcctc aacgtagtc tttatgctcc 900
ttatgaaacta gacttgagtg tatccctggac gacatttgtg atagccocaga atgtotaagc 960
taanacaagc atgagagata aaccagagtg tagtctag 998

```

```

<210> 41
<211> 1697
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 532019CB1

```

```

<400> 41
cagacactca ctgcaccgga gtgagcggga ccatcatgtc catgctcgtg gtatttctct 60
tgtctgtggg tgtaacctgg ggcccagtga cagaagcage catattttat gagacycage 120
ccagctgttg ggcagagctc gaataactgc tgaaccctt ggccaatgtg acgtgacgt 180
gcccagcccg cctggagact ccagacttcc agctgttcaa gaatgggggtg gccacggagc 240
ctgtgcaact tgactcaact gccatcaagc accagtctct gctgacgggt gacaccagg 300
gcgctaccg ctgcccctcg ggcttgtcca caggatggac ccagctgagc aagctcctgg 360
agctgacaag gccaaaagtc ttgctgtctc cctggctctc gatggcgcca gtgtcctgga 420
tcaccccggg cctgaaaaaa acagcagttg gccgaggtgt gctgggggt gtgacttttc 480
tgtctagagc ggagggcgac catgagtttc tggaggtgcc tgaggccca gaggatgtgg 540
aggccaactt tccagtccat cagcctggca actacagctg cagctaccgg accgatgggg 600
aaggccctct ctctgagccc agcgtactg tgaccattga ggagctcgtc gcaccaccac 660
cgctgtgtct gatgaccat ggagagctct cccagctctc gcaccctggc aacaaggtga 720
ccctcaactg cgtggctccc ctgagtgagc tggacttcca gctaccggcg ggggagaaag 780
agctgctggt acccagggagc agcaccagcc cagatgcat ctttcttcc ctgaacggcg 840
tggccctggg ggatggaggt caotacacct gcgctacog gctgcatgac aaccaaaag 900
gctggtccgg ggacagcgcg ccgctcgagc tgattctgag cgtgagagcg ctgcccgcg 960
cggagttctc cccggagccg gagtccggca gggccttggc gctgcggtgc ctggcggccc 1020
tggagggccc gcgcttcgcc ctggtcgccg aggcacgggg cggcgccgc gtgcccgtt 1080
tccagagccc cgtctggacc gaaggctctc tggagctgca caacatttcc gtgctgact 1140
ccgcaacta cagctcgctc taogtggacc tgaagccgcc tttcggggcg tccgcccoca 1200
ggagcgtt ggagctgca gttggacggc cccctcccag ccctcagctc cggcgagct 1260
ggagtggggc ggtcctggcg ggcagagat cgtctctgcy ctgagaggga ccatcccgc 1320
cgtcaactt cgagctgctg cggagggcg agcagaagc cgtgaagcgc gtcggcacc 1380
ccggggccc ggcaacctc gactgatct tctggggccc ccagcacgccc ggcactaca 1440
ggtgccccta cgtctcctg gtgcccaca ctttgaatc ggagctcagc gacctgtgg 1500
agctctggt ggcaagaagc tgactcagc gggggcccag ggtgctgtt gtgtctcag 1560
aagtcccggg gatctggac tggctccctc cctctctgtt gcagcaaacg gccggggtct 1620
ctggggggct ggagaaagct cctctattcc tcccaggaat taataaatgt gaagagagct 1680
ctgtttaaaa aaaaaaa 1697

```

```

<210> 42
<211> 592
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7481063CB1

```

```

<400> 42
atggggcctg acgaacttgc cactggaacc atgaagctgg ctctacttct cgggctgctc 60
ttcactgccc tgctcctgat ctccagggca aattgtgaca totgccagct tgtgacaaag 120

```

WO 02/079441

PCT/US02/09820

```

gatgttgatc tttctcctgt gggaaaccct gatgaatatg ttgatcatgt ggcacaatac 180
tgaacatcct cattaatatt gtccaatgct agaaagctga agaactgtat caatggcaaa 240
ttggcagacg aggacaagag gcatgtgctc agtgggctgg tgagtgcagc tgtgtgtcc 300
tgcacottag tctgggtgga gggctctctg tttctagggc agtgggocgg agtgttatt 360
ctctcccac caacttacc ttatgatatg taattggagc tactggaaat agatttacaat 420
gtgaggtgtg tagcaacagt aaaatgtatt ttaataaag gttataagag goataaaaat 480
ataaattcct gccatagggt aaagaattgt ttaaatagg aaaaatgaa gttaaaagtt 540
caacaagatg gtagaasgtg taaaaataaa tcttgcaaaa aaaaaaaaaa aa 592

```

```

<210> 43
<211> 1031
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2044436CB1

```

```

<400> 43
ttcagcagcc agtttctctg actggatcct gagggtgtctt agcggcctct agacccaat 60
ccagccaaca gctcttatct cagggtggaca tggactgcag ggtccacatg cgaccocatg 120
gcttgacgtg ggtgctgcaa ctgacottgg catggatcct gctagaagcc tgtggaggga 180
gccgccoact ccaagccagg tcccagcaac accatgggct ggcagctgat ctgggcaaaag 240
gtaagctgca ctggccagga ccttgttctc cctcagagat ggacacaaca gagacatcgg 300
gccctggaaa cactccagaa cgtctgtggg tgcagagccc tgaatgogaa tcttctctgg 360
aacactcca acgtgccctt cgcagtgcct tccgctcggc gctattgggg gtacgcaggg 420
cacagccgct ctgcagagg ctctgccagg cctggttctc caactgcgaa gatgatatca 480
cctgcggggc ccgcaattgg ggtcccacca ctctctcaag gaaaaaaaaa ggggggctgg 540
tgagcccagc ttgcgcttac ctatggacag acottgcag acgggaagga ccttctctgc 600
tcggctctgg gccacgccc accgtgtgct gctcctggag cccgcaactg ctccaacatc 660
tccatctctg cggtaacctg tcccagacca ggaagcaggg gccgggaagc tccctccgg 720
cgttcccgca gccctcgtac ctccactcct gacgctcggg gcagcgggag tggcagtgga 780
agcggcagcg gccctcagcg gacgcgtggc cctgagttgg gggagcagcc ctccccag 840
cccccccc caggacaacc agaacccacc cctcgtctct ctccgcttcc tgaatagtt 900
ttgagatgic tgtccctctc cctggagct ccagagacc acccctctcc agttatccc 960
agaaatgacc caactctctc acttttccct ctcccctttg aataaagtgc ccagctagag 1020
caaaaaaaaa a 1031

```

```

<210> 44
<211> 916
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4091564CB1

```

```

<400> 44
ctgtgtggcc agtgtactcg gagtgatcct ggagtctgag aagtccctga goactgtcca 60
cagggcgaca tgcacacagg gtagctgtat caacagccaa gggggatgag atgggccagg 120
gctccaggaa cccccagat: cgggccttgg gagtgtctga ggcgcatgga ggagctggcg 180
gtccaggaga agctgagaaa ggaagactgg aaaggagtgc agtttattct acttcaagag 240
attctgcttc ccttgtgtcc atgagaacat cagtgtcttt cctcaggaaa ttcagcaaga 300
cttggagtba agctccgtca aagaggcccc ccagccagcc cggttctcgg acattgagtg 360
cagccggggc tgggtcaact ggcagcacc tgcacggagc ttctgggcca gctcgcaca 420
aggatgctgc tgaagccagga gccgctgtgc ccagcctctt ttctggacca ctgcgagcag 480

```

WO 02/079441

PCT/US02/09820

```

cactgcagcc caggggagct ggagtcacg ttgtagcagc cacaggccca gggagctgtg 540
cgaaaaaggt agcccaaaag cagacagatg ccagaccaga cagagccggg aaccggcca 600
gggtcccccg catgaccctc ggggcccagg ctaagtgag tgctgctcag atgctgactaa 660
gaggatgac ctctctcagc gggaaaagct ctaagagcc acctgcaggt ccgtgtgggg 720
gaagagccca tgctgtctcg ggaccatggc acaagtgagg tggtgcaat ggtctgtaag 780
gaaaatcact gacagcagct gcctctggga gccaatctgg aaggtctaga ggggataga 840
atctactccc atctctatgt gatgcttggg aaatggttta accctgatca tgtactatgt 900
ttataatgaa agagca 916

```

```

<210> 45
<211> 2577
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 8039739CB1

```

```

<400> 45
atgctctgtt ggtctgctgt gctgacctct aggggtggccc caaaaagctgt acttctctctc 60
aatctctccat ggtccacagc ctctcaaaagg gaaaaagtgg ctctcatatg cagcagcata 120
tcaacttccc tagcccaggg agacacatat tggatcaagc atgagaagtt gttgaaaata 180
aaaatgaca agatcccaat tacagagcct ggaaattacc aatgtaagac ccgaggatcc 240
tccctcagtg atcccgtgca tgttggattt tcaactgact ggtctgactc gaggcttta 300
catctctgct tgaaggaga caatgtcatt ctgagatgc aggggaaaga caacaaaaac 360
actatcaaaa aggtttacta caaggtatgga aaacagcttc ctaattgtta taatttagag 420
aagatccagc tgaattcagc ctccagggat aatagcaaat atcattgtac tggctatagg 480
aagttttaca taactgacat tgaagtaact tcaaaacccc taaatctca agttcaagag 540
ctgtttctac atcctgtgct gagagccagc tcttccacgc ccatagaggg gagtcccatg 600
accctgacct gtgagaccga gctctctcca cagagggcag atgtccagct gcaattctcc 660
ctcttcagag atagccagac cctcggattg gctgggagca ggtcccccag actccagatc 720
cctgcaatgt ggaactgaaga ctccaggtct tactggtgtg aggtggagac agtgaactac 780
agcatcaaaa aaagggagcct gagatctcag atacgtgtac agagagctcc tggctotaat 840
tgaatcttag agatccggcc caccggaggc cagctgattg aaggagaaaa tatgctctt 900
attgctcagc tagcccaggg ttccagggact gtcacattct cctggcaca agaaggaaga 960
gtaagagacc tgggtgaaaa gaccagcagt tccctgttgg cagagctgca tgttctcacc 1020
gtgaaggaga gtgntgcagg gagatactac tgtgcagctg ataacttca cagcccctc 1080
ctcagcagct ggtatcgagt caccgtgaga atccggatct ctaaccctgt cctcaacttc 1140
aggtctccca gggcccacac tgtgttgggg gacctgctgg agcttactg tgaagctctg 1200
agaggtctct cccagatctc gtaccgattt tatcatgagg atgtccctt gggaaacagc 1260
tcaagccctc ctggagaggg agctctcttc aacctctctc tgaactgaga acattctgga 1320
naactactctc gtgatgcaga caatggcctg ggggcccagc acagctatgg agtgagtctc 1380
aggtcaacag tcccggtgct tggccccttc ctcaacctca gggctcccgg gggcccagct 1440
tggttggggg acctgctgga gcttcaactg gagtccctga gaggctcctt cccgactctg 1500
tactgttttt atcagagga tgaccctctg gggaaactct cggcccactc tggaggaggg 1560
goatccttca acctctctct gactacagaa cattctggaa actactcatg tgaggctgac 1620
aatgctctgg gggcccagca cagtaaatg gtgacactca atgttacagg aactccagc 1680
aacagaacag gccctaacgc tgggggaaac aoggggctgg tgcctcagat cctcgtctct 1740
gctgctgtgt ctgctctgct gcattacgcc agggcccagga ggaacccagg aggaattctt 1800
gccactggaa catctagtca cagtctcagt gagtgcagg agcctctctc gtcccagcct 1860
tccagatagc accctcaaga gcccaactca tctaaaccac tagcccactt ggagctggag 1920
ccaatgtaca gcaatgtaaa tccctggagat agcaaccaga tttattccca gatctggagc 1980
atccagcata caaaagaaaa ctccagtaat tgtccaatga tgcactaaga gcatggagaa 2040
cttaagctcc tctattcaga actgaagaag acacaccagc acgactctgc aggggaggtc 2100
agoagcagag gcaagggccc tgaagaagat gatgaagaaa actatgagaa tgtaccactg 2160
gtattactgg cctcagacca ctagcccctt acccagagtg gccccacagga aacagcctgc 2220

```

WO 02/07941

PCT/US02/09820

```

accattttt tttctgttct ctccaaccac acatcatcca tctctccaga ctctgctcc 2280
tacgaggctg ggctgcaggga ttctgaaccc aagaaagAAC aaagtccaag attttccatg 2340
tctttgtaac acttagccct gtgcaaatca gagtatgtga gtggaagaag gggtaggtcc 2400
taactgtaca tctcggatata aacaaatgtg caaattctga ttgattgccc tgtaaaatga 2460
attattctca tgcagtgctc taactgactt ttatcttga attocaaact aaaaaccat 2520
agococagaaa tctaaaaaaa gtaattttag tgagcctatg aaataaaagc cattaga 2577

```

```

<210> 46
<211> 2939
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1265837CBI

```

```

<400> 46
ctaacggtct cttacgcaca cacaaatctc ctcatcata cacacagcct ctctcatgta 60
caatctccct catgcagtgat atgtccatag tttttcacac aacacccctc ttctctcagt 120
ctgtggccct ctctctcaat ctcaacacaca cacccccgcc ccccaacaca ccgttgggccc 180
gcgccctctg cttggccctc cccctgtact gctcgcctcc gacggccggg ggttggggga 240
gccaagtaca ggaggggccc ggcctggtact tcagccccc aggcgggggg ggcgcactcc 300
ggcggctcct gtcaggggccc ggtgctggctg gatcccaagg cagccttcgg gggggggccc 360
tgccctgtgc gtcggggccc ggtcctccgg cggctgtccc gggcggtagg agttggctgc 420
gggatgtgct cagccgggga gctgtctcgg ggcggcagc ggggggaaag cgaagagacc 480
gggagcggcc tggcggagcc ggaggggccc gggcagcagc gggatcccgg ggcgagcccg 540
cggggcggcc gacagggccc gaaggagacc gagcaggatg ttgaagactc acagaacacc 600
actggtgagc cggttggaga tgactacaag aaaaatggaa cactttttgg tgaactgaac 660
aaaaacctta tcaacatggg ctccacaagg atgtattttg gagaacgaat agtgaacca 720
gtaaatgaca tttctctttg ggttatctgt tggttccctg gccctcaagc ccttggaacta 780
gttctgttcc tttgccttgg tattatttat gtgcaacagt aaaaatggc cgaattgaat 840
tgtttgcaat ttggtagcca tatatgtaat tgaagaagtt atatatctca ctttttgaca 900
accyaaaaag tttgccttgg ttcaaatcat gtgctggctg ttttgtaagt aaatttatac 960
atgtagtcca cttaaaaact aactcttgat cataacgggg ttgatatatc attttgaata 1020
taacttagct tatcaaaaac tctgttttca ctactgtgat ctctgtctcc tttatacccc 1080
tctatcccca tgcacaactc taagtcaacc caccagaagc tgaacagggg aataaacagg 1140
acatggaatt caaatcaagc aatatagttc ttatnaagag ttcccaataa scatttcaga 1200
agaaaagata tgaacaacagc taagaatagc ttcaacttag aaaaactctc cccactcaca 1260
ctccccacca aataatctca tattatttgg gaatatcttg patttcaatt gtcccacccc 1320
agctaaactc aaggttaaty ataatagaca taactactcc taatattgtg atgaaaatca 1380
glaaagatag aacgttttct aazaggtcga aataatatac ttattattac tggggaaagc 1440
tcccaggtta aatataactt ttttaaaatg taacatttgg accctagacct actttaaact 1500
atcatttgaa gtttcagaca attttggctg taatctactt ttgtgagttt ttaaagtctc 1560
atagcctagt tgaactgacc ctatggtaat gccatatttt cttgtatcta acaagttgca 1620
tattttttcc tagagagaca ttttcagtgat attttttttt agaaatttat aattttatag 1680
ttcttcaata acacttattc tgagttttga aacaatgtat ttctatctct gacatggatt 1740
ttttcaaaaa aaatttgatc ttactgttgg ttttcaggaa aaaaatcaga tcaattttct 1800
ttgatatcta taccagaagc gtacaatatt aacagtataa aaccaaagtc ttaaatttgg 1860
aacttagcca attttgataa tcttttttca agcctaagtt cacatcagta atggcttagc 1920
catgttatta aggtgtctta attcagcatt ttcaggtttt atattgaaat agcatttctc 1980
ttaaattctc tttgaaatg gagaaatgct ttagtgtatc tttgggtttt gttagagaac 2040
ctaaaatctt tacacttcca tctcaaaagat tataaaggaa aggggggtag ttaagattta 2100
gaattcaagt taaatttcag aaattggggc agtcaggcat ttgtatcttt ggtaggccaa 2160
caagtaaaac atgtagagtg ctgtctatcc caactcataa agctttttacc caatcttatt 2220
tctaaacctc tgtgcatctc tagtgtcttc tcaattctgaa acagaaaata aggaaaacca 2280
tttaacttag ttttcaaaaa tcaagataatc ctaaaacaaa atggttagtca gggtaactaa 2340

```

WO 02/079441

PCT/US02/09820

```

aaagtattgc acatttatat aaatacagtc cttttaaatt ttgactttta aaaaacaaaa 2400
gaactttgtac gatattgtgt ttttatgtct tttycaaat ttttatagta gcotttatga 2460
actongtata agtgcaagtt gtttgaagag gtgtttttat tagtgacaaa tagaattgty 2520
aggttttcaa tagatgtcat gagattttgt atatctacat aaatatcag taattttttt 2580
totaatgota ctggaaatct taacttttct ttgcaacaca taaatgata gatgtacaaa 2640
ataacagctc tggttccacc agtacctaag ttgcaaaaca tttttaaagt aatttttaat 2700
actaacatct tagtatactg toagtactgt aactctgccc actgggttta ataggggtata 2760
tattaaatta tataaagaaa taagatattt tgcgtttatt clttotacat atattattgg 2820
tcagtcactc aaataatatt tggctttgat atgggaaaa acaactttg cctatgtaat 2880
ggaaataaaa tattttcttt taagaaatah attggaatgc agattatact aaaaggsaaa 2940
aaaaagcabc gagagaaaga ayaagaaaaa aagagtgtgc cgataacgag tgaanaaac 2999

```

```

<210> 47
<211> 1314
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5568527CB1

```

```

<400> 47
ccggcgttgg ccgctgctgc tagcagcttg aaccccaggg tggggaccga tgtcggcttg 60
ggctgctgcc agcctaagca gggccgctgc ccgatgcttg ctggcacgag gcccccgggt 120
cagggcagct cotccgctgc acccccggcc cccccccc gagccccggg gctgcgctgc 180
cgtcccgccg aggaagctgc actttaaccg ggctgtcccc gccggccaca acaagtgttc 240
caaatcaagg cacatcaagg gtccgaagga cgtcgaagag agtcgcaatc tctccaaact 300
ctgtttgaac atccgctcgg cagtgaagga agggagcccc aaccctgagc acaacagcaa 360
cttggccaat atcttagagg tgtgtcgcag caaacatat cccaagtcaa cgaattgagac 420
agcactgaaa atgggagaaat ccaaggcac ttatttctgt tatgagggtc gagccctggc 480
tggctcttct ctgctcctcg aggcattatc taacagttag cacagtgcc aagcagacat 540
tagacatctc ctgaataaga atggaggagt gatggctgta ggagctcgtc actcttttga 600
caaaaagggg tggattgtgg ttgaagtgga gacagagag aagaaggctg tgaacctaga 660
gggtgcccgt gagatggcaa tcgaagcagg agctgaggag gtaacaggaaa ctgaagatga 720
agaagaaaag aacgttttta aatttatttg tgatgectct tcactgcaac aagtgaggaa 780
gaagctggac tcccggggcc tgtgttctgt gtccgtgaca ctgagttcca tcccacactc 840
aaaagtgccg ctggctgagc ccgacctgga acaggccgca catctcatte aggtctctcg 900
caaccacgag gatgtgcttc acgtctatga taacattgaa taaccaggctc acatgtgccc 960
ccgggtctct tcttagaaat gtggcagccc attccagcac acaggcttct gcagcaatct 1020
ctyaggytaa agccggtggg aggtccagca ggcaggagag cccaaggaca ggaacttgga 1080
ccttgaagcc aaaggaatct caattgtggg gcttctctgt cagctctgct gctgtctcag 1140
agcaatctgg atgagtgccc cgacaacctc toggatgcag ggcaggaca cccagctggt 1200
cagactctga tgttgggtag ctggcctctg tggggattgt aagtgccctg aggcgctctg 1260
tactagaaac tgcctcttaa aataacggtg attattgggt gctgcaaaaa aaaa 1314

```

```

<210> 48
<211> 2619
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7503641CB1

```

```

<400> 48
atacagcttc ggtactcttg taaggacgaa gtgtgctgga aagcaggtgc cagagggggc 60

```

WO 02/079441

PCT/US02/09820

```

ggtccccaga cgtgagaag gccacggcct cgtgatgtc accgggggtg ctaaggaggg 120
ggagggggta gggctgtttt tctggagaga gacttagaga cagctgggac aaagcctggg 180
gctggggggg gcccatggcg ctgccatccc gaatcctgct ttggaaactt gtgcttctgc 240
agactctctg tgttctctct caactcagggt cctcggtagc cgcgctgctt ggcagctccg 300
tggltgcccc gtcgggggtg agctggaggc cgggcccggg ggcggtgctg cgtctgocaga 360
gcccgcccat ggtgtggacc caggaccggc tgcccgaccg ccagcggctg ctccactggg 420
acctgcccgg ccccgggggt ggcggccgac ggcgctgctt ggaactgtac tggcggggcg 480
agcagcggct gtaacgaggcg cgggaccgcg gccgctgaga gctctcggcc tggcctctcg 540
accagcccaa cttctcgtct ctaaccggcg cgttggaggc gacggacggg gggctgtaca 600
cctgcaacct gcaaccatac taactgccacc tctacgagag cctggccgctc cgcctgcttg 660
aacggcgggc acgtgtggac ccagccggac gtyggaggag ctcaacaggt ggtgcactgg 720
gaccggcagc cgcgggggtt ccggccggac cggcggggac gctctgtgga cctctacggg 780
tcgggggagc gccgggccta cgggcccctt tttctgcccg accgctgggc tggggcggcg 840
gatgctcttg agcgggtgga cttctcactg cgtatcgagc cgtctggagt cgcgcagag 900
ggcaactact cctgcaacct gcaaccacct taactgtggc tgcaagaaag ccgctctctc 960
caactgacgg tgcgcaacc ccacggggag cgcggggggg ggggctctcc gggcaacggc 1020
tccagccaca gggggcggcc agggccagac ccaacactgg cgcggggcca caactcactc 1080
aatgtcatcg tcccggagag ccgagccacc ttcttcagc agctgggcta cgtctgggcc 1140
acgtgctgct tcttcatcct gctactgttc actgtctctc tggcggcccg caggcggcgc 1200
gggggctacg aatactcggc ccagaagtgc ggaagtcaa agggaaagga tgttaacttg 1260
cggagttctg ctgtgctgca aggggaccag atgctttaca ggggtgagga catccagcta 1320
gattacaaaa acaacatcct gaaggagggc ggggagctgg cccacagccc cctgctctgc 1380
aagtacatcg acctagacaa aggtttccgg aaggagaact gcaaataggg aggcctggg 1440
ctctggctcg ggcagcagc tgcaacctctc ctgtctgtgc tctctggggc atctctctgat 1500
gctccggggc tcaaccctct tccagcggct ggtccgctt tcttgggatt tggcctgggc 1560
gtatgcaagc gccgctctca caccctctcc ccaggggctt ggtggcagca tagcccccac 1620
ccctggggcc tttgctcagc ggtggccctg ccaaccctg gcaacaacaa aatccactg 1680
atgcccacta tgccctcaga cctctctggc ctctgcccgc tgggggctcg aagcaatcc 1740
tggaggacac tcccatcaga acctgcccgc ccaaaaactg ggttcagct cagggcagga 1800
gtcccactcc tccagggctc tctgtctccg gggctggggg atgttctctg agggagcac 1860
tccatcaga acttggcagc cttgaagtgt ggtcagcct ccgagaggt cccactctc 1920
ctggggtgct gcctgcaacc aagagctccc caacctgtao caocatgtg gactccagc 1980
accaatgttt ctcccaggg acctgctgac ttgaatgcaa cccctgtctc ctctgtgttg 2040
ctttgggcca cctggggctg caccctctg ccttctctct ccccatctct acctagctt 2100
tgctctcagc cacttgata gtcactggc tccctgtgac tcttgacct gacaccctc 2160
ccttgactc tgctgggct ggaacttagc gctgggcta catttgctt cgtactggc 2220
tgaggacagg gggggagtg aagtgtgtt ggggtggct gtttgccac tctcagacc 2280
caacattgc atctctggtt ggaactgcaa ccatcacaat aaagtccca tctgatltt 2340
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2400
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa gggggcgcca 2460
cagagggcac gctcccggc ggggataatt gccgggagga gtcagtgag ggcgcaaat 2520
tgctctcaga aggggcacta acaaaactgg ggaaaaacac gaggctaatg tgcctctga 2580
gaaatgttt ctgcggcgca atcccacag atcagaaaa 2619

```

```

<210> 49
<211> 1370
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7503458CB1

```

```

<400> 49
ttcagcagcc agtttctctg actggatcct gagggtcttt agcggcctct agaccacaat 60
ccagccaaca gctcttatct cagggtggaca tggactgcag ggtccaatg ccagccatcg 120

```

WO 02/079441

PCT/US02/09820

```

gectgacgtg ggtgctgcaa ctgaccttgg catggatcct gctagaagcc tgtggggga 180
gocgccoact ccaagccagg tccagcaaac acatgggctt ggcaactgat ctgggcaag 240
gtaagctgca octggcaggg gatcccaggc cctcagttcc aaaaacttac ctgaatctcc 300
aagaccttgg ctcaagaggt toccctttgc caggaccttg ttgtccctca gagatggaca 360
caacagagac atcgggccct gnaaacctcc cagaacgtgt tggagtgcgc agccctgaat 420
gcgaatcctt octggaacac ctccaactgt cctttgcag tccgtccgc ctgggctat 480
tgggggtacg ccagccacag ccgctctcgc aggagctctg ccaggcctgg ttcgccact 540
gggaagatga tctaacctgc ggggcccoga cttgggctc ccccaacttc tcaagaaaa 600
aaaaaggggg gctgggtgag cccgcttggc ctacctatg gacagacctt cgcagcagg 660
acggaccttt gtcgctcggc tctggggcac gccctaccgg tggctgtctc tggagccctt 720
cactgcttca acatctccat ctccggggta cctcgtccca gaccaggagc acggggccgg 780
gaagctccct cccggcgttc ccgcagccct cgcacctcca tccctggacc tggggccagc 840
gggaagtggca gtgggaagcg cagcggcccc tagcggagcc gttggccctg gttggggag 900
cgaccttcc cccagccocg cccctcagga caccagaaac cccacccttc gtcctctcg 960
cctctgtaa tagtllttag algtctgtcc ctctccctg gagctccaga gaccacc 1020
tctccagtt atcccagaaa tgaccctaac ctctcaact tccctctccc ctttgaataa 1080
agtccccagc tagagcaaaa anananaaaa anananaaaa aaaaaaaa aaaaaaaa 1140
ananaaaa ananaaaa aaaaaaaa aaaaaaaa aaaaaaaa gggggccccc 1200
caaaaggttc ccttttggc ctagggttat tctcgtgag agactttatt atgtatgagc 1260
aaaaaaatt atttggggg ggggttttaa aaagggggg gggggaaaaa cccggggggg 1320
aaaaataat tagcggaaa acacaacaag tttttttttt ttttccccc 1370

```

```

<210> 50
<211> 1201
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7500925CB1

```

```

<400> 50
ggggtcocga actgcttatt cgggcagtgg aagagacggc cgggcgttgg ccgctgtgc 60
tagcagcttg aaccocaggg tcgggcccca gttcggcttg gctgctgcc agcctaagca 120
gggcccctgc ccgatgcttg ctggcacgag gccccggggt cagggccgct cctccgcgcg 180
accocccgcc ctcccacccc gagccccggg gctgctgtgc cgtctccggc aggacgctgc 240
actttaccgc ggtgttcccc gccgggcaca acaagtggtc caaagttagc cacatcaagg 300
gtccgaaagg cgtcgaagg agtgcacat tctccaaact ctgtttgaa atccgctgg 360
cagtgaaaga aggagggccc aacctgagc acaacagcaa cctggccaat atcttagagg 420
tgtctcgaag caaacatatt cccaagtcaa cgtattgagc agcactyaaa atggagaaat 480
ccnaggcaac ttatttgcctg tatgagggtc gaggccctgg tggctctctc ctgctcaatc 540
aggcattatc taacagtagc caaagtgtcc aagcagacat tagacatctc ctgaataaga 600
atggagagat gatgctgta gtagctctgc actcttttga caaaaagggg gtagattgtg 660
ttgaagtgga ggaacagagag aagaagctg tgaacctaga gcgtgccctg gagatggcaa 720
tcgaagcagg agctgagat gtcaaggaaa ctgaagatga agnagaaggg aactttttaa 780
aaaaatgtgg cagccatctc cagcaacacag gcttctgcag caatctctga gggtaaaagg 840
ggtgggaggg tcaagcagcc aggagcccca aggcagagac ttgcagacct gaagcaaaag 900
gaatctcaat tgtggggcct ccttgcagc tctgctgctg tctcagagcc atctggatga 960
gtgtcccagc accctctcgg atgcagggca ggaaccacca gctggtcaga ctctgatgtt 1020
gggtagctgg cctctgttgg gattgtaagt gccctgagc gctctgtact agaaactgct 1080
cttaataata acggtgatta ttggttctgc caaaaaaaa agaggagtta gaggagtg 1140
agggaggggt gttgtgggggt ggggaaggac tcccaggggg gggccccggc accccaattg 1200
g 1201

```

【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
10 October 2002 (10.10.2002)

PCT

(10) International Publication Number
WO 2002/079441 A3(51) International Patent Classification: G01N 33/53,
C12P 21/06, A61K 38/00, C07K 14/00

(21) International Application Number: PCT/US2002/009820

(22) International Filing Date: 29 March 2002 (29.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/280,527 30 March 2001 (30.03.2001) US
60/282,112 6 April 2001 (06.04.2001) US
60/282,702 9 April 2001 (09.04.2001) US
60/283,855 13 April 2001 (13.04.2001) US
60/243,718 19 October 2001 (19.10.2001) US
60/339,236 7 December 2001 (07.12.2001) US
60/557,002 13 February 2002 (13.02.2002) US(71) Applicant (for all designated States except US): INCYTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): BAUGHN, Mariah,
R. [US/US]; 14244 Santiago Road, San Leandro, CA
94577 (US). BURFORD, Neil [GB/US]; 105 Wildwood
Circle, Durham, CT 06422 (US). DING, Li [CN/US];
3353 Alma Street # 146, Palo Alto, CA 94306 (US).
DUGGAN, Brendan, M. [AU/US]; 243 Buena Vista
Avenue # 306, Sunnyvale, CA 94086 (US). ELLIOTT,
Vicki, S. [US/US]; 3770 Polton Place Way, San Jose,
CA 95121 (US). FORSYTHE, Ian, J. [US/US]; 308
Roble Avenue, Redwood City, CA 94061 (US). GANDHI,
Ameena, R. [US/US]; 705 5th Avenue, San Francisco,
CA 94118 (US). GIBTZEN, Kimberly, J. [US/US]; 691
Los Hencos Drive, San Jose, CA 95123 (US). GRIFFIN,
Jennifer, A. [US/US]; 33691 Mello Way, Fremont, CA
94555 (US). HE, Ann [CN/US]; 4601 Cahalina Drive,
San Jose, CA 95129 (US). HONCHELL, Cynthia, D.
[US/US]; 400 Laurel Street #203, San Carlos, CA 94070
(US). ISON, Craig, H. [US/US]; 1242 Weathersfield Way,
San Jose, CA 95118 (US). LAL, Preeti, G. [US/US]; P.O.
Box 5142, Santa Clara, CA 95056 (US). LEE, Ernestine,
A. [US/US]; 624 Kains Street, Albany, CA 94706 (US).
LEE, Sally [US/US]; 825 East Evelyn #425, Sunnyvale,
CA 94086 (US). LU, Dyang Aina, M. [US/US]; 233 CoyDrive, San Jose, CA 95123 (US). MASON, Patricia,
M. [US/US]; 360 Clarke Lane, Morgan Hill, CA 95037
(US). SANJANWALA, Madhu, M. [US/US]; 210 Silvia
Court, Los Altos, CA 94024 (US). SWARNAKAR,
Anita [CA/US]; 8 Locksley Avenue #5D, San Francisco,
CA 94122 (US). RAMKUMAR, Jayalaxmi [IN/US];
34359 Maybird Circle, Fremont, CA 94555 (US). TANG,
Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA
95118 (US). THANGAVELU, Kavitha [IN/US]; 1950
Montecito Avenue #23, Mountain View, CA 94043 (US).
TRAN, Uyen, K. [US/US]; 2638 Mabury Square, San
Jose, CA 95133 (US). WALLA, Narinder, K. [US/US];
890 Davis Street A # 205, San Leandro, CA 94577 (US).
WARREN, Bridget, A. [US/US]; 2250 Homestead Court
2, Los Altos, CA 94024 (US). YAO, Montique, G.
[US/US]; 1189 Woodgate Drive, Carmel, IN 46033 (US).
XU, Yuning [US/US]; 1739 Walnut Drive, Mountain
View, CA 94040 (US). YUE, Henry [US/US]; 826 Lois
Avenue, Sunnyvale, CA 94087 (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, GR, 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, TI, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VN, YU, ZA, ZM, ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:
11 March 2004For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: 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 2002/079441 A3

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US02/09820
A. CLASSIFICATION OF SUBJECT MATTER IPC(C) : G01N33/53; C12P21/06; A61K 38/00; C07K 14/00 US CL : 435/7.1, 69.1; 514/2; 530/350 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. : 435/7.1, 69.1; 514/2; 530/350 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) MEDLINE, CAPLUS, AGRICOLA, BIOSIS, EMBASE, WPI/DS, GENBANK, EMBL, SWISSPROT, EAST		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WIDMER et al. Identification of a second human acetyl-CoA carboxylase gene. Biochem J. 1996, Vol. 316, pages 915-922, particularly page 917. HACC-beta is 94% identical to SEQ ID NO:1.	1,2,10,17,18,20,56-80
<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 "B" earlier application or patent published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family		
Date of the actual completion of the international search 02 April 2003 (02.04.2003)	Date of mailing of the international search report 20 JUN 2003	
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 David J. Steadman <i>Jamie Ford</i> Telephone No. (703) 308-0196	

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/09820

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

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

1. Claim Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. Claim Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. Claim Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)This International Searching Authority found multiple inventions in this international application, as follows:
Please See Continuation Sheet

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1,2,9,10,17,18,20,56

Remark on Protest The additional search fees were accompanied by the applicant's protest.
 No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

PCT/US02/09820

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

Group I, claims 1, 2, 9, 10, 17, 18, 20, and 56-80, drawn to a polypeptide comprising SEQ ID NO:1-25, a method of making a polypeptide, and a method of using a polypeptide to screen a compound as an agonist.

Group II, claims 3-8, 12, 13, 46-55, and 81-105, drawn to a polynucleotide encoding a polypeptide comprising SEQ ID NO:1-25, a polynucleotide comprising SEQ ID NO:26-50, a host cell, a transgenic organism, an array, a method for detecting a target polynucleotide by hybridization or PCR, and a method for generating an expression profile.

Group III, claims 11 and 30-43, drawn to an antibody that binds the polypeptide of SEQ ID NO:1-25, a composition comprising an antibody, a method for making an antibody, a diagnostic test for a condition associated with the expression of SECP using an antibody, and a method of diagnosing a condition associated with expression of SECP.

Group IV, claim 14-16, drawn to a method of detecting a target polynucleotide in a sample having a sequence of SEQ ID NO:26-50.

Group V, claim 19, drawn to a method for treating a disease associated with decreased expression of human secreted proteins (SECP) by administering a composition comprising SEQ ID NO:1-25.

Group VI, claims 21 and 22, drawn to a composition comprising an agonist of SEQ ID NO:1-25 and a method for treating a disease associated with decreased expression of SECP by administering a composition comprising an agonist of SEQ ID NO:1-25.

Group VII, claim 23, drawn to a method of screening a compound as an antagonist of SEQ ID NO:1-25.

Group VIII, claims 24 and 25, drawn to a composition comprising an antagonist of SEQ ID NO:1-25 and a method for treating a disease associated with overexpression of SECP by administering a composition comprising an antagonist of SEQ ID NO:1-25.

Group IX, claims 26 and 27, drawn to a method for screening a compound that binds to or modulates SEQ ID NO:1-25.

Group X, claim 28, drawn to a method of screening a compound for effectiveness in altering expression of SEQ ID NO:26-50.

Group XI, claim 29, drawn to a method of assessing toxicity of a test compound using the polynucleotide of SEQ ID NO:26-50.

Group XII, claim 44 drawn to a method for detecting a polypeptide using an antibody that binds the polypeptide of SEQ ID NO:1-25.

Group XIII, claim 45, drawn to a method for purifying a polypeptide using an antibody that binds the polypeptide of SEQ ID NO:1-25.

If applicant should elect the invention of Group I, III, V-IX, XII, or XIII, applicant should additionally elect one of the following inventions to be searched:

Group A, drawn to the polypeptide of SEQ ID NO:1.
 Group B, drawn to the polypeptide of SEQ ID NO:2.
 Group C, drawn to the polypeptide of SEQ ID NO:3.
 Group D, drawn to the polypeptide of SEQ ID NO:4.
 Group E, drawn to the polypeptide of SEQ ID NO:5.
 Group F, drawn to the polypeptide of SEQ ID NO:6.
 Group G, drawn to the polypeptide of SEQ ID NO:7.
 Group H, drawn to the polypeptide of SEQ ID NO:8.
 Group I, drawn to the polypeptide of SEQ ID NO:9.
 Group J, drawn to the polypeptide of SEQ ID NO:10.
 Group K, drawn to the polypeptide of SEQ ID NO:11.
 Group L, drawn to the polypeptide of SEQ ID NO:12.
 Group M, drawn to the polypeptide of SEQ ID NO:13.
 Group N, drawn to the polypeptide of SEQ ID NO:14.
 Group O, drawn to the polypeptide of SEQ ID NO:15.
 Group P, drawn to the polypeptide of SEQ ID NO:16.
 Group Q, drawn to the polypeptide of SEQ ID NO:17.
 Group R, drawn to the polypeptide of SEQ ID NO:18.
 Group S, drawn to the polypeptide of SEQ ID NO:19.
 Group T, drawn to the polypeptide of SEQ ID NO:20.
 Group U, drawn to the polypeptide of SEQ ID NO:21.
 Group V, drawn to the polypeptide of SEQ ID NO:22.
 Group W, drawn to the polypeptide of SEQ ID NO:23.
 Group X, drawn to the polypeptide of SEQ ID NO:24.
 Group Y, drawn to the polypeptide of SEQ ID NO:25.

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

INTERNATIONAL SEARCH REPORT

PCT/US02/09820

In the absence of any response from the Applicant, this Authority will establish the International Search Report based on the main invention. The claims drawn to the main invention are as follows:

Claims: 1, 2, 9, 10, 17, 18, 20, 56 and SEQ ID NO:1.

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 P 1/16	A 6 1 P 1/16	4 B 0 6 5
A 6 1 P 3/10	A 6 1 P 3/10	4 C 0 8 4
A 6 1 P 5/16	A 6 1 P 5/16	4 C 0 8 5
A 6 1 P 7/00	A 6 1 P 7/00	4 H 0 4 5
A 6 1 P 7/06	A 6 1 P 7/06	
A 6 1 P 9/10	A 6 1 P 9/10	1 0 1
A 6 1 P 9/12	A 6 1 P 9/12	
A 6 1 P 11/00	A 6 1 P 11/00	
A 6 1 P 11/06	A 6 1 P 11/06	
A 6 1 P 13/12	A 6 1 P 13/12	
A 6 1 P 17/00	A 6 1 P 17/00	
A 6 1 P 17/06	A 6 1 P 17/06	
A 6 1 P 19/04	A 6 1 P 19/04	
A 6 1 P 19/06	A 6 1 P 19/06	
A 6 1 P 19/10	A 6 1 P 19/10	
A 6 1 P 21/04	A 6 1 P 21/04	
A 6 1 P 25/08	A 6 1 P 25/08	
A 6 1 P 25/14	A 6 1 P 25/14	
A 6 1 P 25/16	A 6 1 P 25/16	
A 6 1 P 25/28	A 6 1 P 25/28	
A 6 1 P 27/06	A 6 1 P 27/06	
A 6 1 P 27/12	A 6 1 P 27/12	
A 6 1 P 29/00	A 6 1 P 29/00	1 0 1
A 6 1 P 31/00	A 6 1 P 31/00	
A 6 1 P 31/18	A 6 1 P 31/18	
A 6 1 P 33/00	A 6 1 P 33/00	
A 6 1 P 35/00	A 6 1 P 35/00	
A 6 1 P 35/02	A 6 1 P 35/02	
A 6 1 P 37/00	A 6 1 P 37/00	
A 6 1 P 37/08	A 6 1 P 37/08	
C 0 7 K 14/47	C 0 7 K 14/47	
C 0 7 K 16/18	C 0 7 K 16/18	
C 0 7 K 16/46	C 0 7 K 16/46	
C 0 7 K 19/00	C 0 7 K 19/00	
C 1 2 M 1/00	C 1 2 M 1/00	A
C 1 2 N 1/15	C 1 2 N 1/15	
C 1 2 N 1/19	C 1 2 N 1/19	
C 1 2 N 1/21	C 1 2 N 1/21	
C 1 2 N 5/10	C 1 2 P 21/02	C
C 1 2 P 21/02	C 1 2 P 21/08	
C 1 2 P 21/08	C 1 2 Q 1/68	A
C 1 2 Q 1/68	G 0 1 N 33/15	Z
G 0 1 N 33/15	G 0 1 N 33/50	Z
G 0 1 N 33/50	G 0 1 N 33/53	D
G 0 1 N 33/53	G 0 1 N 33/53	M
G 0 1 N 33/566	G 0 1 N 33/566	
G 0 1 N 37/00	G 0 1 N 37/00	1 0 2

C 1 2 N 15/00 F
 C 1 2 N 5/00 A
 A 6 1 K 37/02

- (31)優先権主張番号 60/283,855
 (32)優先日 平成13年4月13日(2001.4.13)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/343,718
 (32)優先日 平成13年10月19日(2001.10.19)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/339,236
 (32)優先日 平成13年12月7日(2001.12.7)
 (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)

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

- (72)発明者 バーフォード、ニール
 アメリカ合衆国コネチカット州06422・ダラム・ワイルドウッドサークル 105
 (72)発明者 ディング、リー
 アメリカ合衆国カリフォルニア州94306・パロアルト・#146・アルマストリート 3353
 (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94086・サニーベイル・#306・ブエナビスタアベニュー 243
 (72)発明者 エリオット、ビッキー・エス
 アメリカ合衆国カリフォルニア州95121・サンノゼ・ポルトンブレイスウェイ 3770
 (72)発明者 フォーサイス、イアン・ジェイ
 アメリカ合衆国カリフォルニア州94061・レッドウッドシティ・ローブルアベニュー 308
 (72)発明者 ガンディー、アミーナ・アール
 アメリカ合衆国カリフォルニア州94118・サンフランシスコ・フィフスアベニュー 705
 (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95123・サンノゼ・ロスウエコドライブ 691
 (72)発明者 グリフィン、ジェニファー・エイ
 アメリカ合衆国カリフォルニア州94555・フレモント・メローウェイ 33691
 (72)発明者 ヒー、アン
 アメリカ合衆国カリフォルニア州95129・サンノゼ・カタリナドライブ 4601
 (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94070・サンカルロス・#203・ローレルストリート 400
 (72)発明者 アイソン、クレイグ・エイチ
 アメリカ合衆国カリフォルニア州95118・サンノゼ・ウェザーズフィールドウェイ 1242
 (72)発明者 ラル、プリーティ・ジー
 アメリカ合衆国カリフォルニア州95056・サンタクララ・ピーオーボックス 5142

- (72)発明者 リー、アーンステーション・エイ
アメリカ合衆国カリフォルニア州9 4 7 0 6・アルバニー・ケインズストリート 6 2 4
- (72)発明者 リー、サリー
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・# 4 2 5・イーストイープリン 8
2 5
- (72)発明者 リュ、デュング・アイナ・エム
アメリカ合衆国カリフォルニア州9 5 1 2 3・サンノゼ・コイドライブ 2 3 3
- (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 サンジャンワラ、マデュー・エム
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・シルビアコート 2 1 0
- (72)発明者 スウォーナカール、アニータ
アメリカ合衆国カリフォルニア州9 4 1 2 2・サンフランシスコ・# 5ディー・ロックスリーアベ
ニュー 8
- (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 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 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 0 4 3・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 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 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 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 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 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
- (72)発明者 ユエ、ヘンリー
アメリカ合衆国カリフォルニア州9 4 0 8 7・サニーベイル・ルイスアベニュー 8 2 6

F ターム(参考) 2G045 AA34 AA35 AA40 BB50 BB51 CB01 DA12 DA13 DA36 FA11
FA12 FB02 FB03 FB12 GC15
4B024 AA01 AA11 BA43 BA61 CA02 CA04 CA09 FA02 GA05 HA14
HA15
4B029 AA07 BB20 FA15
4B063 QA01 QA05 QA18 QA19 QQ43 QR08 QR55 QR62 QS25 QS34
4B064 AG01 AG27 CA19 CC24 DA05 DA06 DA08 DA13
4B065 AB01 BA02 CA24 CA25 CA44 CA46
4C084 AA01 AA02 AA17 BA01 BA02 CA18 CA53 DC50 MA05 NA14
ZA022 ZA062 ZA152 ZA162 ZA332 ZA362 ZA422 ZA452 ZA592 ZA662
ZA682 ZA752 ZA812 ZA892 ZA942 ZA962 ZA972 ZB072 ZB132 ZB152
ZB262 ZB272 ZB322 ZB372 ZC352 ZC542
4C085 AA13 AA14 AA16 BB11 BB41 BB43 CC22 CC23 EE01
4H045 AA10 AA11 AA20 AA30 BA10 CA40 DA76 EA21 EA22 EA23

EA24 EA25 EA27 EA28 EA29 EA50 FA74

专利名称(译)	分泌蛋白質		
公开(公告)号	JP2005506049A	公开(公告)日	2005-03-03
申请号	JP2002578444	申请日	2002-03-29
[标]申请(专利权)人(译)	洞察Genomics公司		
申请(专利权)人(译)	洞察基因组公司		
[标]发明人	ボーグンマライアアール バーフォードニール デイングリー ダガンプレンダンエム エリオットビッキーエス フォーサイスイアンジェイ ガンディーアミーナアール ギーツエンキンバリージェイ グリフィンジェニファーエイ ヒーアン ホンチエルシンシアディー アイソクレイグエイチ ラルプリーティジー リーアーンステイーンエイ リーサリー リュデユングアイナエム メイソンパトリシアエム サンジャンワラマデューエム スウォーナカールアニータ ランクマールジャヤラクシミ タングワイトム サンガベルカピサ トランユエンケイ チョーラナリンダーケイ ワレンブリジットエイ ヤオモニークジー スーユーミング ユエヘンリー		
发明人	ボーグン、マライア・アール バーフォード、ニール デイング、リー ダガン、ブレンダン・エム エリオット、ビッキー・エス フォーサイス、イアン・ジェイ ガンディー、アミーナ・アール ギーツエン、キンバリー・ジェイ グリフィン、ジェニファー・エイ ヒー、アン ホンチエル、シンシア・ディー アイソン、クレイグ・エイチ ラル、プリーティ・ジー リー、アーンステイーン・エイ リー、サリー リュ、デユング・アイナ・エム		

メイソン、パトリシア・エム
サンジャンワラ、マデュー・エム
スウォーナカール、アニータ
ランクマール、ジャヤラクシミ
タング、ワイ・トム
サンガベル、カピサ
トラン、ユエン・ケイ
チョーラ、ナリンダー・ケイ
ワレン、ブリジット・エイ
ヤオ、モニー・ク・ジー
スー、ユーミン
ユエ、ヘンリー

IPC分類号	G01N33/50 A01K67/00 A61K38/00 A61K39/395 A61K45/00 A61P1/04 A61P1/16 A61P3/10 A61P5/16 A61P7/00 A61P7/06 A61P9/10 A61P9/12 A61P11/00 A61P11/06 A61P13/12 A61P17/00 A61P17/06 A61P19/04 A61P19/06 A61P19/10 A61P21/04 A61P25/08 A61P25/14 A61P25/16 A61P25/28 A61P27/06 A61P27/12 A61P29/00 A61P31/00 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37/00 A61P37/08 C07H21/04 C07K14/47 C07K16/18 C07K16/46 C07K19/00 C12M1/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N9/00 C12N15/09 C12P21/02 C12P21/08 C12Q1/68 G01N33/15 G01N33/53 G01N33/566 G01N37/00
CPC分類号	A01K2217/05 A61K38/00 A61P1/04 A61P1/16 A61P3/10 A61P5/16 A61P7/00 A61P7/06 A61P9/10 A61P9/12 A61P11/00 A61P11/06 A61P13/12 A61P17/00 A61P17/06 A61P19/04 A61P19/06 A61P19/10 A61P21/04 A61P25/08 A61P25/14 A61P25/16 A61P25/28 A61P27/06 A61P27/12 A61P29/00 A61P31/00 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37/00 A61P37/08 C07K14/47
FI分類号	C12N15/00.ZNA.A A61K39/395.D A61K39/395.N A61K45/00 A61P1/04 A61P1/16 A61P3/10 A61P5/16 A61P7/00 A61P7/06 A61P9/10.101 A61P9/12 A61P11/00 A61P11/06 A61P13/12 A61P17/00 A61P17/06 A61P19/04 A61P19/06 A61P19/10 A61P21/04 A61P25/08 A61P25/14 A61P25/16 A61P25/28 A61P27/06 A61P27/12 A61P29/00.101 A61P31/00 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37/00 A61P37/08 C07K14/47 C07K16/18 C07K16/46 C07K19/00 C12M1/00.A C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 G01N37/00.102 C12N15/00.F C12N5/00.A A61K37/02
F-TERM分類号	2G045/AA34 2G045/AA35 2G045/AA40 2G045/BB50 2G045/BB51 2G045/CB01 2G045/DA12 2G045/DA13 2G045/DA36 2G045/FA11 2G045/FA12 2G045/FB02 2G045/FB03 2G045/FB12 2G045/GC15 4B024/AA01 4B024/AA11 4B024/BA43 4B024/BA61 4B024/CA02 4B024/CA04 4B024/CA09 4B024/FA02 4B024/GA05 4B024/HA14 4B024/HA15 4B029/AA07 4B029/BB20 4B029/FA15 4B063/QA01 4B063/QA05 4B063/QA18 4B063/QA19 4B063/QQ43 4B063/QR08 4B063/QR55 4B063/QR62 4B063/QS25 4B063/QS34 4B064/AG01 4B064/AG27 4B064/CA19 4B064/CC24 4B064/DA05 4B064/DA06 4B064/DA08 4B064/DA13 4B065/AB01 4B065/BA02 4B065/CA24 4B065/CA25 4B065/CA44 4B065/CA46 4C084/AA01 4C084/AA02 4C084/AA17 4C084/BA01 4C084/BA02 4C084/CA18 4C084/CA53 4C084/DC50 4C084/MA05 4C084/NA14 4C084/ZA022 4C084/ZA062 4C084/ZA152 4C084/ZA162 4C084/ZA332 4C084/ZA362 4C084/ZA422 4C084/ZA452 4C084/ZA592 4C084/ZA662 4C084/ZA682 4C084/ZA752 4C084/ZA812 4C084/ZA892 4C084/ZA942 4C084/ZA962 4C084/ZA972 4C084/ZB072 4C084/ZB132 4C084/ZB152 4C084/ZB262 4C084/ZB272 4C084/ZB322 4C084/ZB372 4C084/ZC352 4C084/ZC542 4C085/AA13 4C085/AA14 4C085/AA16 4C085/BB11 4C085/BB41 4C085/BB43 4C085/CC22 4C085/CC23 4C085/EE01 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/CA40 4H045/DA76 4H045/EA21 4H045/EA22 4H045/EA23 4H045/EA24 4H045/EA25 4H045/EA27 4H045/EA28 4H045/EA29 4H045/EA50 4H045/FA74
優先権	60/280527 2001-03-30 US 60/282112 2001-04-06 US 60/282702 2001-04-09 US 60/283855 2001-04-13 US 60/343718 2001-10-19 US 60/339236 2001-12-07 US 60/357002 2002-02-13 US

摘要(译)

本发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。 本发明还提供了用于诊断，治疗或预防与SECP异常表达有关的疾病的方法。

接頭コード	解析のタイプおよび/またはプログラムの例
GNN GFG ENST	例えば GENSCAN (Stanford University, CA, USA) または FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK)を用いた、ゲノム配列からのエキソン予測
GBI	手動で編集された、ゲノム配列群の解析
FL	ステイッチまたはストレッチゲノム配列(実施例5参照)
INCY	ゲノムへのEST配列のマッピングからの完全長転写物とエキソンとの予測。エキソン群と生じる転写物とを予測するために、ゲノム位置とEST構成とのデータが組み合わされる。