

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-500870

(P2004-500870A)

(43) 公表日 平成16年1月15日(2004.1.15)

(51) Int.Cl. ⁷	F I	テーマコード (参考)
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 45/00	4 B O 2 4
A 6 1 K 45/00	A 6 1 P 1/04	4 B O 6 3
A 6 1 P 1/04	A 6 1 P 1/16	4 B O 6 4
A 6 1 P 1/16	A 6 1 P 1/18	4 B O 6 5
審査請求 未請求 予備審査請求 有 (全 317 頁) 最終頁に続く		

(21) 出願番号	特願2002-504308 (P2002-504308)	(71) 出願人	301005050
(86) (22) 出願日	平成13年6月20日 (2001.6.20)		インサイト・ゲノミックス・インコーポレイテッド
(85) 翻訳文提出日	平成14年12月18日 (2002.12.18)		アメリカ合衆国カリフォルニア州94304・パロアルト・ポータードライブ 3160
(86) 国際出願番号	PCT/US2001/019862		
(87) 国際公開番号	W02001/098353	(74) 代理人	100089266
(87) 国際公開日	平成13年12月27日 (2001.12.27)		弁理士 大島 陽一
(31) 優先権主張番号	60/212, 890	(72) 発明者	ヒルマン、ジェニファー・エル
(32) 優先日	平成12年6月20日 (2000.6.20)		アメリカ合衆国カリフォルニア州94040・マウンテンビュー・#17・モンロードライブ 230
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/213, 466		
(32) 優先日	平成12年6月23日 (2000.6.23)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/214, 601		
(32) 優先日	平成12年6月27日 (2000.6.27)		
(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】

本発明はヒト分泌性タンパク質 (S E C P) および S E C P 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、S E C P 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 (d) を有する群から選択した実質上単離されたポリペプチド。

(a) 配列番号 1 乃至 44 (SEQ ID NO: 1-44) を有す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: 1-44 を有する群から選択したアミノ酸配列と少なくとも 90% が同一であるようなアミノ酸配列を有する天然のポリペプチド

(c) SEQ ID NO: 1-44 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

(d) SEQ ID NO: 1-44 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抗原性断片 10

【請求項 2】

SEQ ID NO: 1-44 を有する群から選択した請求項 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。

【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: 45-88 を有する群から選択した請求項 4 に記載の単離されたポリヌクレオチド。 20

【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結合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。

【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、 30

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

(b) そのように発現した前記ポリペプチドを受容する過程とからなり、前記組換えポリヌクレオチドが、請求項 1 に記載の前記ポリペプチドをコードするポリヌクレオチドに機能的に結合したプロモーター配列を有することを特徴とする方法。

【請求項 10】

請求項 1 に記載のポリペプチドと特異結合するような単離された抗体。

【請求項 11】

以下の (a) 乃至 (d) を有する群から選択した実質上単離されたポリヌクレオチド。 40

(a) 配列番号 45 乃至 88 (SEQ ID NO: 45-88) を有す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: 45-88 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% が同一であるような天然のポリヌクレオチド

(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

(e) (a) ~ (d) の RNA 等価物

【請求項 12】

請求項 11 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 13】

請求項 11 に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中から検出する方法であって、

(a) 前記サンプル中の前記標的ポリヌクレオチドに相補的な配列を有する少なくとも 20 の連続したヌクレオチドを含むプローブを用いて前記サンプルをハイブリダイズする過程と、

(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。

10

【請求項 14】

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 13 に記載の方法。

【請求項 15】

請求項 11 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) 前記標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含むことを特徴とする方法。

20

【請求項 16】

有効量の請求項 1 のポリペプチドと、薬剤として許容できる賦形剤とを有することを特徴とする成分。

【請求項 17】

前記ポリペプチドが、SEQ ID NO: 1-44 を有する群から選択したアミノ酸配列を含むことを特徴とする請求項 16 に記載の成分。

【請求項 18】

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 16 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。

30

【請求項 19】

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すステップと、

(b) 前記サンプルにおいてアゴニスト活性を検出するステップとを含むことを特徴とするスクリーニング方法。

【請求項 20】

請求項 19 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成分。

【請求項 21】

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 20 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。

40

【請求項 22】

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すステップと、

(b) 前記サンプルにおいてアンタゴニスト活性を検出するステップとを含むことを特徴とするスクリーニング方法。

【請求項 23】

請求項 22 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

50

る賦形剤とを含むことを特徴とする成分。

【請求項 2 4】

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

【請求項 2 5】

請求項 1 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 つの試験化合物に結合させる過程と、

(b) 請求項 1 に記載のポリペプチドの試験化合物との結合を検出し、それによって請求項 1 に記載のポリペプチドに特異結合する化合物を同定する過程とを含むことを特徴とする方法。 10

【請求項 2 6】

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

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

(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。 20

【請求項 2 7】

請求項 5 の配列を含む標的ポリヌクレオチドの発現を変化させるのに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ステップと、

(b) 前記標的ポリヌクレオチドの変異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。 30

【請求項 2 8】

試験化合物の毒性を算定する方法であって、

(a) 核酸を含む生物学的サンプルを前記試験化合物で処理する過程と、

(b) 処理した前記生体サンプルの核酸と、請求項 1 1 のポリヌクレオチドの少なくとも 2 0 の連続するヌクレオチドを含むプローブをハイブリダイズさせるステップであって、このハイブリダイゼーションゼーションが、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われ、前記標的ポリヌクレオチドが、請求項 1 1 のポリヌクレオチドのポリヌクレオチド配列またはその断片を含むポリヌクレオチドである、前記ステップと、

(c) ハイブリダイゼーション複合体の収量を定量するステップと、 40

(d) 前記処理された生物学的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、処理されていない生物学的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、前記処理された生物学的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 2 9】

生物学的サンプル中の S E C P 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) 前記抗体が前記ポリペプチドに結合し、抗体とポリペプチドとの複合体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 1 0 に記載の抗体と結合する過程と、 50

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

【請求項 30】

前記抗体が、

(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 10 に記載の抗体。

【請求項 31】

請求項 10 に記載の抗体と、許容できる賦形剤とを含む化合物。

【請求項 32】

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 31 に記載の化合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 33】

前記抗体が標識されることを特徴とする請求項 31 に記載の化合物。

【請求項 34】

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 33 に記載の化合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 35】

請求項 10 に記載の抗体の特異性を有するポリクローナル抗体を調製する方法であって、

(a) 抗体反応を誘発する条件下で、SEQ ID NO: 1-44 を有する群から選択したアミノ酸配列含むポリペプチドまたはその免疫抗原性断片を用いて動物を免疫化する過程と、

(b) 前記動物から抗体を単離する過程と、

(c) 前記単離された抗体をポリペプチドでスクリーニングし、それによって、SEQ ID NO: 1-44 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 36】

請求項 35 に記載の方法で産出した抗体。

【請求項 37】

請求項 36 に記載の抗体及び適切なキャリアを含む化合物。

【請求項 38】

請求項 10 に記載の抗体の特異性を有する抗体を用いてモノクローナル抗体を製造する方法であって、

(a) 抗体反応を誘発する条件下で、SEQ ID NO: 1-44 を有する群から選択したアミノ酸配列含むポリペプチドまたはその免疫抗原性断片を用いて動物を免疫化する過程と、

(b) 前記動物から抗体産出細胞を単離する過程と、

(c) 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、

(d) 前記ハイブリドーマ細胞を培養する過程と、

(e) SEQ ID NO: 1-44 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結合するような前記培養モノクローナル抗体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 39】

請求項 38 に記載の方法で産出したモノクローナル抗体。

【請求項 40】

請求項 39 に記載の抗体及び適切なキャリアを含む化合物。

【請求項 41】

10

20

30

40

50

F a b 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 10 に記載の抗体。

【請求項 42】

組換え免疫グロブリンライブラリのスクリーニングにより前記抗体を産出することを特徴とする請求項 10 に記載の抗体。

【請求項 43】

SEQ ID NO: 1-44 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を検出する方法であって、

(a) 前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項 10 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: 1-44 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がサンプル中に存在することを標示することを特徴とする方法。

10

【請求項 44】

SEQ ID NO: 1-44 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の特異結合を許容する条件下で、サンプルを用いて請求項 10 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: 1-44 を有する群から選択した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。

20

【請求項 45】

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 46】

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 47】

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 48】

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 49】

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。

30

【請求項 50】

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 51】

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 52】

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 53】

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 54】

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。

40

【請求項 55】

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 56】

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 57】

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 58】

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 59】

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 60】
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 61】
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 62】
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 63】
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 64】
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 65】
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 66】
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 67】
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 68】
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 69】
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 70】
SEQ ID NO: 26 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 71】
SEQ ID NO: 27 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 72】
SEQ ID NO: 28 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 73】
SEQ ID NO: 29 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 74】
SEQ ID NO: 30 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 75】
SEQ ID NO: 31 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 76】
SEQ ID NO: 32 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 77】
SEQ ID NO: 33 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 78】
SEQ ID NO: 34 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 79】
SEQ ID NO: 35 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 80】
SEQ ID NO: 36 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 81】
SEQ ID NO: 37 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 82】
SEQ ID NO: 38 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 83】
SEQ ID NO: 39 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 84】

10

20

30

40

50

SEQ ID NO: 40 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 85】

SEQ ID NO: 41 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 86】

SEQ ID NO: 42 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 87】

SEQ ID NO: 43 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 88】

SEQ ID NO: 44 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 89】

SEQ ID NO: 45 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 90】

SEQ ID NO: 46 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 91】

SEQ ID NO: 47 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 92】

SEQ ID NO: 48 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 93】

SEQ ID NO: 49 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 94】

SEQ ID NO: 50 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 95】

SEQ ID NO: 51 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 96】

SEQ ID NO: 52 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 97】

SEQ ID NO: 53 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 98】

SEQ ID NO: 54 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 99】

SEQ ID NO: 55 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 100】

SEQ ID NO: 56 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 101】

SEQ ID NO: 57 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 102】

SEQ ID NO: 58 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレ

10

20

30

40

50

オチド。

【請求項 1 0 3】

SEQ ID NO : 5 9 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 0 4】

SEQ ID NO : 6 0 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 0 5】

SEQ ID NO : 6 1 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

10

【請求項 1 0 6】

SEQ ID NO : 6 2 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 0 7】

SEQ ID NO : 6 3 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 0 8】

SEQ ID NO : 6 4 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 0 9】

SEQ ID NO : 6 5 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

20

【請求項 1 1 0】

SEQ ID NO : 6 6 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 1 1】

SEQ ID NO : 6 7 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 1 2】

SEQ ID NO : 6 8 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

30

【請求項 1 1 3】

SEQ ID NO : 6 9 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 1 4】

SEQ ID NO : 7 0 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 1 5】

SEQ ID NO : 7 1 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

40

【請求項 1 1 6】

SEQ ID NO : 7 2 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 1 7】

SEQ ID NO : 7 3 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 1 8】

SEQ ID NO : 7 4 のポリヌクレオチド配列を有する請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 1 9】

50

SEQ ID NO: 75 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 120】

SEQ ID NO: 76 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 121】

SEQ ID NO: 77 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 122】

SEQ ID NO: 78 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。 10

【請求項 123】

SEQ ID NO: 79 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 124】

SEQ ID NO: 80 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 125】

SEQ ID NO: 81 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。 20

【請求項 126】

SEQ ID NO: 82 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 127】

SEQ ID NO: 83 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 128】

SEQ ID NO: 84 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 129】

SEQ ID NO: 85 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。 30

【請求項 130】

SEQ ID NO: 86 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 131】

SEQ ID NO: 87 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。

【請求項 132】

SEQ ID NO: 88 のポリヌクレオチド配列を有する請求項 11 に記載のポリヌクレオチド。 40

【発明の詳細な説明】

【0001】

(技術分野)

本発明は、分泌タンパク質の核酸配列及びアミノ酸配列に関する。本発明はまた、これらの配列を利用した細胞増殖異常、自己免疫／炎症性の疾患、心血管疾患、神経系疾患、および発達障害の診断・治療・予防に関する。本発明はさらに、分泌タンパク質の核酸配列及びアミノ酸配列の発現における外来性化合物の効果についての評価に関する。

【0002】

(発明の背景)

タンパク質の輸送及び分泌は、細胞機能のために必須である。タンパク質の輸送は、タンパク質のアミノ末端のシグナルペプチドによって媒介されて、輸送若しくは分泌される。シグナルペプチドは、約10から20の疎水性アミノ酸からなり、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は分泌経路を通して進むか、ER、ゴルジ体、若しくはリソソームのような分泌細胞器官に残るかのいずれかであり得る。分泌経路を通して移動するタンパク質は、細胞外空間へ分泌されるか、原形質膜内に残るかのいずれかである。原形質膜内に保持されるタンパク質は、一つ若しくは複数の膜貫通ドメインを有し、各々が約20の疎水性アミノ酸残基よりなる。分泌タンパク質は、通常、不活性の前駆体として合成され、分泌経路を介して移動する際に翻訳後プロセッシングイベントによって活性化される。そのようなイベントには、グリコシル化、タンパク質分解、及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含まれる。タンパク質の運搬の間に起こりうるその他のイベントは、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及びフォールディング、及びレセプター若しくは細孔複合体を有するタンパク質の相互作用を含む。アミノ末端シグナルペプチドを有する分泌タンパク質の例は以下に述べられ、細胞間情報〔シグナル〕伝達の重要な役割を有するタンパク質を含む。そのようなタンパク質は、膜貫通レセプター及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、成長及び分化因子、酵素、ニューロペプチド、及び血管作動性因子（*vasomediator*）、細胞表面マーカーおよび抗原認識分子が含まれる（Alberts, B. ら（1994）Molecular Biology of The Cell, Garland Publishing, New York, NY, 557-560, 582-592ページの概説を参照）。

【0003】

細胞表面のマーカーには、免疫系の白血球細胞上で識別される細胞表面抗原が含まれる。これら抗原は、系統的なモノクローナル抗体（mAb）ベースの「ショットガン（shotgun）」技術を用いて同定されてきた。これらの技術によって、未知の細胞表面白血球抗原に対する数百のmAbが作製された。それらの抗原は、様々な分化白血球細胞型及び未分化白血球細胞型での共通の免疫細胞化学的な局在パターンに基づいた「分化クラスター」へとグループ分けされた。あるクラスター内の抗原は、ある単一の細胞表面タンパク質を識別すると仮定され、また「分化クラスター（cluster of differentiation）」即ち「CD」名が指定されている。CD抗原によって識別されたタンパク質をコードする遺伝子のいくつかは、通常、分子生物学技術によってクローンされ、確認されている。CD抗原は、膜貫通タンパク質、および原形質膜に固着された細胞表面タンパク質の両方として特徴づけられ、原形質膜に固着された細胞表面タンパク質は、共有結合によってグリコシルホスファチジルイノシトール（GPI）のような脂肪酸含有糖脂質に結合されることで原形質膜に固着されている。（Barclay, A. N. ら（1995）The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, 17-20. ページの概説を参照。）

基質タンパク質（MP）は、組織の形成、成長、再形成、及び持続において機能し、炎症反応の重要な介在物質及び制御因子としても機能する、膜貫通及び細胞外タンパク質である。MPの発現及びバランスが、先天的疾患、後成的疾患、若しくは感染性疾患の結果生じる生化学的変化によって乱されることがある。加えてMPは、免疫学的応答において白血球の遊走、増殖、分化、及び活性化に影響を与える。MPは、一つ若しくは複数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、及びフィブロネクチン様ドメインが含まれる。加えて、MPは重度にグリコシル化される場合があり、接着相互作用の役割をしている可能性のある、アルギニン-グリシン-アスパラギン酸（RGD）のトリペプチドモチーフを含んでいる場合がある。MPは、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、及びそのタンパク分解誘導体ソマトメジンBのような細胞外タンパク質を含み、また細胞接着分子（CAM）、カドヘリン、及びイ

ンテグリンのような細胞接着受容体を含む (Ayad, S. ら、(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. 及び Nelson, W. J. (1997) *BioEssays* 19 : 47-55 の概説を参照.)。

ムチンは高度にグリコシル化された糖タンパク質であり、粘液ゲルの主な構造的成分である。ムチンの生理的機能は細胞保護、機械的保護、粘性維持および細胞認識である。MUC6 はヒト胃液ムチンで、胆嚢、膵臓、精嚢、および女性生殖器官にも見られる (Toribara, N. W. ら (1997) *J. Biol. Chem.* 272 : 16398-16403)。MUC6 遺伝子はヒト染色体 11 にマッピングされている (Toribara, N. W. ら (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、抗細菌作動体分子の誘導に関与し得る (Theopold, U. ら (1996) *J. Biol. Chem.* 271 : 12708-12715)。

【0004】

タフテリン (tuffelin) はこれまでに同定されている四つの異なったエナメル質タンパク質の一つである。他の三つの既知のエナメル質タンパク質はアメロゲニン、エナメルインおよびアメロブラスチンである。これらの成分タンパク質からのエナメル細胞外基質の構築は、無機質置換を行える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. ら (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. ら (1998) *Connect Tissue Res.* 39 : 177-184)。

【0005】

オルファクトメジン (olfactomedin) 関連タンパク質は、保存された C 末端モチーフを有する細胞外基質の分泌糖タンパク質である。これらは、多種の組織および線虫からヒトにいたるまで広い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有する遺伝子ファミリーを含んでいる。この 5 つのうちの一つである TIGR/ミオシリン (myocilin) タンパク質は目に発現され、緑内障の病原と関連している (Kulkarni, N. H. ら (2000) *Genet. Res.* 76 : 41-50)。Yokoyama ら (1996) による研究は、神経芽細胞腫細胞系 cDNA ライブラリで、ラットの神経細胞のオルファクトメジンに関連し、ER に局在性するタンパク質に、AMY と呼ぶ 135 アミノ酸のタンパク質が 96 % の配列同一性を有することを発見した。これは AMY の神経組織における重要な役割を示唆している (Yokoyama, M. ら、(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に基質関連性機能があることを示唆している (Danielson, P. E. ら (1994) *J. Neurosci. Res.* 38 : 468-478)。

【0006】

Mac-2 結合タンパク質は 90 キロダルトンの血清タンパク質 (90 K) 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、もう 1 つの分泌糖タンパク質である。これは、ヒトマクロファージ結合性レクチンである Mac-2 に特異的に結合する。構造的に、成熟型タンパク質は長さが 567 アミノ酸で、18 アミノ酸のリーダーが、その前に付いている。16 のシステインと 7 つの潜在的 N 結合グリコシル化部位がある。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のシステインリッチドメインによって定義される古いタンパク質スーパーファミリーに極めて類似するドメインを示す (Koths, K. ら、(1993) *J. Biol. Chem.* 268 : 12708-12715)。

8 : 1 4 2 4 5 - 1 4 2 4 9) 9 0 K はエイズの患者の垂集団の血清中において上昇し、原発腫瘍サンプルおよび腫瘍細胞系において異なったレベルで発現される。U l l r i c h ら (1 9 9 4) は 9 0 K が宿主防御系を刺激し、インターロイキン-2 の分泌を誘発することを実証した。この免疫刺激は発癌性形質転換、ウイルス感染または病原性侵襲の結果であると提案されている (U l l r i c h , A . , e t a l . (1 9 9 4) J . B i o l . C h e m . 2 6 9 : 1 8 4 0 1 - 1 8 4 0 7) 。

【0007】

セマホリン (s e m a p h o r i n g) は少なくとも 3 0 の異なったメンバーからなる軸索誘導分子の大集団であり、脊椎動物、無脊椎動物および特定のウイルスにも見られる。すべてのセマホリンには、長さがほぼ 5 0 0 アミノ酸であるセマドメインが含まれている。セマホリン受容体であるニューロピリン (n e u r o p i l i n) は i n v i t r o で神経突起の成長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 C U B ドメイン、ジスコイジンドメイン、M A M ドメインの三つの異なったドメインからなる。ニューロピリンの C U B および M A M モチーフは、タンパク質間相互作用において役割を果たしていることが示されており、またセマ (s e m a) と C 末端ドメインを介してセマホリンの結合に関与することが示されている (R a p e r , J . A . (2 0 0 0) C u r r . O p i n . N e u r o b i o l . 1 0 : 8 8 9 4 の概説を参照)。プレキシン (p l e x i n) 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機構により細胞接着を媒介する。プレキシン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 (O h t a , K . ら (1 9 9 5) C e l l 1 4 : 1 1 8 9 - 1 1 9 9) 。発達段階の神経系において運動神経軸索や中枢神経系軸索の誘導を制御するよう作用するプレキシンもあることを示す証拠がある。完全なセマホリンドメインを有するプレキシンは、古典的セマホリンの祖先とセマホリンの結合パートナーの祖先の両方であり得る (W i n b e r g , M . L . ら (1 9 9 8) C e l l 9 5 : 9 0 3 - 9 1 6) 。

【0008】

ヒト妊娠特異的ベータ 1 - 糖タンパク質 (P S G) は、分子量が 7 2 キロダルトン、6 4 キロダルトン、6 2 キロダルトンおよび 5 4 キロダルトンの密接に関連する糖タンパク質のファミリーである。癌胎児性抗原と共に、これらは免疫グロブリンスーパーファミリー内でサブファミリーを構成する (P l o u z e k C . A . および C h o u J . Y . , E n d o c r i n o l o g y 1 2 9 : 9 5 0 9 5 8) 。 P S G の異なる垂集団がヒト胎盤の栄養胚葉および羊膜や漿膜によって産生されることが発見されている (P l o u z e k C . A . ら (1 9 9 3) P l a c e n t a 1 4 : 2 7 7 2 8 5) 。

【0009】

自己分泌型運動促進因子 (A M F) は、腫瘍細胞遊走を調節する運動性サイトカインの一つであり、したがって、それに関連するシグナル伝達経路の同定は非常に重要である。自己分泌型運動促進因子受容体 (A M F R) の発現は胸腺腫における腫瘍進行に関連することがわかっている (O h t a Y . ら (2 0 0 0) I n t . J . O n c o l . 1 7 : 2 5 9 2 6 4) 。 A M F R は分子量が 7 8 キロダルトンの細胞表面糖タンパク質である。

【0010】

ホルモンは、血液循環により運ばれ、標的細胞の表面またはその中の特異的受容体に結合する分泌分子である。これらは多様な生化学的成分や作用機構を持っているが、ホルモンは二つのカテゴリーに分類される。一つ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、細胞質内又は核内の受容体に結合し、遺伝子発現を変更する複合体を形成する。これらの分子の例には、レチノイン酸、チロキシン、およびプロゲステロン、エストロゲン、コルチゾール、アルドステロンのようなコレステロール性ステロイドホルモンがある。第二のカテゴリーには、原形質膜の内側へ信号を伝達する細胞表面の受容体に結合することで機能する親水性ホルモンが含まれる。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 (エピネフリン、ノルエピネフ

10

20

30

40

50

リン)、及びヒスタミンが、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、セクリチン、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、及びバソプレシンが含まれる。(例えば、Lodish ら、(1995) *Molecular Cell Biology*, Scientific American Books Inc., New York, NY, 856-864 頁参照)

プロオピオメラノコルチン(POMC)は、脳下垂体前葉によって合成されるホルモンであるコルチコトロピン(ACTH)の前駆体ポリペプチドで、副腎皮質の刺激において作用する。POMC はまたベータリポトロピン(β -LDH)の前駆体ポリペプチドでもある。各々のホルモンは異なった生物活性を持つ、さらに小さいペプチドを含む。 α -メラニン細胞刺激ホルモン(α -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、 γ -リポトロピン(γ -LDH)と β -エンドルフィン¹⁰は β -LPHのペプチド成分であるが、 β -MSHは γ -LPH中に含まれる。POMCのエキソン2および3における遺伝子突然変異に起因するACTH欠乏による副腎不全是、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性疾患を生じる(C¹¹hretien, M. ら(1979) *Canad. J. Biochem.* 57:1111-1121, Krude, H. ら(1998) *Nature Genet.* 19:155-157, Online Mendelian Inheritance in Man, OMIM. Johns Hopkins University, Baltimore, MD. OMIM Number: 176830: 2000年8月1²⁰日. World Wide Web URL: www.ncbi.nlm.nih.gov/omim/).

成長及び分化因子は、細胞間伝達で機能する分泌タンパク質である。複数の因子がオリゴマー形成を要求し、または活性のために膜タンパク質との結合を要求する。それらの因子とレセプター間の複合体相互作用は、細胞分裂、細胞分化、細胞シグナル伝達、及び細胞運動性を刺激又は抑制する細胞内信号伝達経路を開始させる。ほとんどの成長及び分化因子は、その局所的環境内の細胞に作用する(パラ分泌シグナル伝達)。成長及び分化因子の3つのブロードクラスが存在する。第一のクラスには、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、及び血小板由来成長因子のような大きなポリペプチド成長因子が含まれる。第二のクラスには、コロニー刺激因子(CSF)のような造血性成長因子が含まれる。造血性成長因子は、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、及びそれらの幹細胞前駆体のような血液細胞の増殖及び分化を刺激する。第三のクラスは、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、アンギオテンシンI³⁰I、血管作用性小腸ペプチド、及びブラジキニンのような小さなペプチド因子が含まれ、これらは増殖以外の細胞機能を制御するためのホルモンとして機能する。

【0011】

成長及び分化因子は、in vitroでの細胞の腫瘍性転換、及びin vivoでの腫瘍進行において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、腫瘍の血管新生及び転移に貢献しうる。造血中での成長因子の制御異常によって、貧血、白血病、及びリンパ腫起こる場合がある。インターフェロンのようなある成長因子は、in vitro及びin vivoの双方で腫瘍細胞に細胞毒性である。さらには、いくつかの成長因子及び成長因子レセプタが、構造的及び機能的の両面で腫瘍性タンパク質と関連している。さらに、成長因子はプロトオンコジーン及び腫瘍抑制因子遺伝子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 頁の概説を参照)。⁴⁰

【0012】

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らは、S⁵⁰

l i t 遺伝子の哺乳類相同体（ヒト S l i t - 1、S l i t - 2、S l i t - 3 およびラット S l i t - 1）を同定した。コードされたタンパク質は、共に保存されたタンパク質間相互作用ドメインである E F G 様モチーフおよびロイシンリッチリピートを含む推定分泌タンパク質である。S l i t - 1 m R N A、S l i t - 2 m R N A および S l i t - 3 m R N A s は脳、脊髄および甲状腺に各々発現される（I t o h, A. ら（1998）Brain Res. Mol. Brain Res. 62:175-186）。タンパク質の S l i t ファミリは神経組織内の g l y p i c a n - 1 の機能性リガンドであることが示されており、このことは、それらの相互作用は中枢神経系の組織発生時の特定の段階において重要であることを意味する（L i a n g, Y. ら（1999）J. Biol. Chem. 274:17885-17892）。

10

【0013】

神経ペプチド及び血管作動性因子（N P / V M）は、内在性シグナル伝達分子の大きなファミリーを構成する。このファミリーに含まれるのは、神経ペプチド及び神経ペプチドホルモンとして、ボンベシン、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、オピオイド、ガラニン、ソマトスタチン、タキキニン、ウロテンシン I I があり、平滑筋刺激に関与するペプチドとして、バソプレッシンペプチド、血管作用性小腸ペプチドがあり、また循環系によって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、ホルミルメチオニルペプチド、グルカゴン、コレシストキニン、及びガストリンがある。N P / V M は直接信号を伝達して、別の神経伝達物質及びホルモンの活性若しくは放出を調節し、またカスケードにおいて触媒酵素として作用する可能性がある。N P / V M の効果は、極端に短い時間から長時間持続まで、広範囲にわたる。（M a r t i n, C. R. ら（1985）Endocrine Physiology, Oxford University Press, New York, NY, 57-62 頁参照）

20

N P / V M は、多くの神経系疾患及び心血管疾患に関与する。例えば、神経ペプチド Y は、高血圧症、うっ血心不全、情動障害、食欲調節に関与する。ソマトスタチンは、腸、膵臓腺房細胞、及び膵臓ベータ細胞内の分泌を阻害と共に、成長ホルモン及び下垂体前葉の分泌を阻害する。ソマトスタチンレベルでの減少は、アルツハイマー病及びパーキンソン病で報告されている。バソプレッシンは、水分吸収及びナトリウム吸収のために腎臓内で作用し、また高濃度では、血管平滑筋の収縮、血小板活性、及び肝臓内グリコーゲン分解を刺激する。バソプレッシン及びその類似体は、尿崩症の治療のため臨床的に用いられる。エンドセリン及びアンジオテンシンは高血圧症に関与し、アンジオテンシンの血漿レベルを減少させるカプトプリルのような薬剤は、血圧を下げるのに用いられる（W a t s o n, S. 及び S. A r k i n s t a l l, （1994）The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, 194 ; 252 ; 284 ; 55 ; 111 の各頁）。

30

【0014】

神経ペプチドはまた痛覚（痛み）の役割を有していることが示された。血管作用性小腸ペプチドは、慢性的な神経障害痛において重要な役割を果たしているようである。オピオイドレセプター様 1 レセプターのための内在性リガンドであるノシセプチンは、主に非侵害受容性を有すると考えられており、持続性若しくは慢性的な痛みの異なった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（D i c k i n s o n, T. 及び F l e e t w o o d - W a l k e r, S. M. （1998）Trends Pharmacol. Sci. 19:346-348）。

40

シグナルペプチドを有する他のタンパク質には酵素活性を持つ分泌タンパク質が含まれる。そのような活性には、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、イソメラーゼ活性、若しくはリガーゼ活性が含まれる。例えば、基質メタロプロテイナーゼは、細胞外基質を分解する分泌加水分解酵素であり、したがって、腫瘍転換、組織形態形成、及び関節炎において重要な

50

役割を果たす (Reponen, P. ら、(1995) Dev. Dyn. 202 : 388-396 ; Firestein, G. S. (1992) Curr. Opin. Rheumatol. 4 : 348-354 ; Ray, J. M. 及び Stetler-Stevenson, W. G. (1994) Eur. Respir. J. 7 : 2062-2072 ; Mignatti, P. 及び Rifkin, D. B. (1993) Physiol. Rev. 73 : 161-195)。別の例は脂質合成またはエネルギー生成において使用する酢酸を活性化するアセチルCoA合成酵素である (Luong, A. ら (2000) J. Biol. Chem. 275 : 26458-26466)。アセチルCoA合成酵素が活性化されると、酢酸とCoAからアセチルCoAが形成される。アセチルCoA合成酵素はAMP結合ドメインシグネチャとして同定される配列類似性領域を共有する。アセチルCoA合成酵素は高血圧と関連性があることが示されている (H. Toh (1991) Protein Seq. Data Anal. 4 : 111-117 および Iwai, N. ら、(1994) Hypertension 23 : 375-380)。

【0015】

シグナルペプチドを有する他のタンパク質には、分泌経路を通過していくタンパク質の糖鎖形成 (グリコシル化) に関与する酵素が含まれる。ムチン型O結合糖鎖形成はタンパク質の糖鎖形成の主な形態である。ムチン型糖鎖形成の開始は、セリンおよびトレオニンアミノ酸の水酸基に単糖N-アセチルガラクトサミンを付加 (GalNAc α 1-O-Ser/Thr) することによって生じる。GalNAc O-糖鎖形成はムチンのような高分子量の分泌糖タンパク質において著しいが、種々の糖タンパク質においてもまた見られる (White, T. ら、J. Biol. Chem. (1995) 270 : 24156-24165)。さらに、セリン/トレオニンリッチ タンデムリピートはヒトムチンのコアタンパク質の特徴である。そのタンデムリピート領域はまた、モノクローナル抗体であるHMF G-1、HMF G-1およびSM-3によって認識される多くの抗原決定基をも含む。タンデムリピート領域内の糖鎖形成 (グリコシル化) 部位はMuc1 が単離された臓器によって異なってグリコシル化されることが発見された。可変的糖鎖形成活性の発見はMuc1 タンデムリピートにマップされる癌関連抗原決定基の分子的根拠の理解をさらに深めるために重要である (Gendler, S. J. ら (1990) J. Biol. Chem. 265 : 15286-15293)

抗原認識分子は、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から守るために発達させてきた高度な、複雑な免疫機構において重要な役割を果たすものである。免疫系の主要特徴は「自己」分子から外来性分子または抗原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性タンパク質かつ膜貫通タンパク質によって主に媒介される。大部分のこれらのタンパク質は、保存された構造のドメインの一つ以上の反復を含むメンバーである、免疫グロブリン (Ig) スーパー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) と呼ばれる構造でのジスルフィド結合によって連結されている逆平行の β シートから構成されている。IgスーパーファミリーのメンバーにはT細胞受容体、主要組織適合性 (MHC) タンパク質、抗体および「分化クラスター」またはCD抗原のような免疫細胞特異的表面マーカーが含まれる。これらの抗原は、系統的なモノクローナル抗体 (mAb) -ベースの「ショットガン (shotgun)」技術を用いて同定されてきた。これらの技術によって、未知の細胞表面白血球抗原に対する数百のmAbが作製された。それら抗原は、様々な分化白血球細胞型及び未分化白血球細胞型での共通の免疫細胞化学的な局在パターンに基づいた「分化クラスター」へとグループ分けされた。あるクラスター内の抗原は単一の細胞表面タンパク質を同定と仮定され、また「分化クラスター (cluster of differentiation)」即ち「CD」名が指定されている。CD抗原によって識別されたタンパク質をコードする遺伝子のいくつかは、通常の方法生物学技術によってクローンされ、確認されている。CD抗原は、膜貫通タンパク質、および原形質膜に固着された細胞表面タンパク質の両方として特徴付けられ、原形質膜に固着された細胞表面タン

パク質は、共有結合によってグリコシルホスファチジルイノシトール（GPI）のような脂肪酸含有糖脂質に結合されることで原形質膜に固着されている（Barclay, A. N. ら（1995）The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, 17-20頁の概説を参照）。

【0016】

MHCタンパク質はT細胞に外来性抗原を結合し、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
10
関与する。例えば、ウイルスに感染した細胞は細胞内の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発現する。MHC I／抗原複合体はキラーT細胞によって認識され、キラーT細胞が感染した細胞とその中のウイルスを破壊する。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性タンパク質を摂取し、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細胞がサイトカインおよび免疫反応を刺激する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、外来性抗原に結合した自己MHC分子に応答するのと同じように、移植受容者のT細胞が移植された臓器上の外来性MHC分子に応答する場合に発生する。（Alberts, B. ら（1994）Molecular Biology of the Cell, Garland Publishing, New York, NY, 1229-1246
20
頁の概説を参照）

抗体、即ち免疫グロブリンは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性抗原に結合して、それを中和する。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二つの同じポリペプチド重鎖（H鎖）と二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特徴的なY型を与える。抗体はH鎖組成に基づいて分類される。五つの抗体のクラスである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、どちらもいずれのH鎖対と対として会合でき得る。血液循環中に見られる抗体の最も一般的なク
30
ラスの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、またはこの基本的構造の多量体である。

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鎖では約110のアミノ酸、H鎖では約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、ある特定のクラスのH鎖またはL鎖の内ではほとんど同じである。可変領域は約110のアミノ酸でH鎖とL鎖の両方からなる。しかし、特定のクラスのH鎖とL鎖の中で、可変領域のアミノ酸配列は可変である。。H鎖またはL鎖の可変領域内に配列多様性が大きい高頻度可変領域が三ヶ所あり、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一緒になって抗原認識部位を形成する（前出のAlberts, 1206-1213 頁と 1216-1217 頁参照
40
）。

【0017】

H鎖とL鎖は両方とも反復されるIgドメインを含む。例えば、典型的なH鎖は四つのIgドメインを含んでおり、そのうちの三つは定常領域内にあり、一つは可変領域内にあっ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鎖は二つのIgドメインを含み、そのうちの一つは定常領域内にあり、他の一つは可変領域内にある。

【0018】

免疫系は体内に入る外来性分子を認識し、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して、あらゆる範囲の抗体を準備して防御しなければならない。このような抗体の多様性は可変領域と定常領域をコードする遺伝子区分
50

の体細胞性再編成によって作られる。これらの遺伝子区分は、各々の遺伝子区分が隣接する、極めて保存されたDNA配列間で生じる部位特異的組換えによって連結される。何百という異なった遺伝子区分があるため、何百万という固有の遺伝子が組み合わされて作成され得る。その上、これらの区分の不正確な連結とこれらの区分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さらに寄与している。

【0019】

多くのイソメラーゼが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の同化作用や異化作用の経路での段階を触媒する。一つの種類のイソメラーゼはペプチジルプロリルシーストランスイソメラーゼ（PPIアーゼ）として知られている。PPIアーゼはタンパク質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（FKBP）とサイクロフィリン（CyP）である。FKBPは強力な免疫抑制薬FK506とラパミシンを結合し、それによって、T細胞におけるシグナル伝達経路を抑制する。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。算出された分子量にしたがって名付けられているFKBPファミリーの五つのメンバー（FKBP12、FKBP13、FKBP25、FKBP52およびFKBP65）があり、それらが異なったタンパク質複合体と結合する細胞の異なった領域に局在している（Coss, M. ら（1995）J. Biol. Chem. 270:29336-29341; Schreiber, S. L.（1991）Science 251:283-287）。

10

20

【0020】

CyPのペプチジルプロリルイソメラーゼ活性は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イソメラーゼ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構築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。例えば、ショウジョウバエにおいて、CyP NinaAはロドプシンの正しい局在化に必要であるが、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タンパク質と結合し、この作用はシクロスポリンによって抑制されることがわかっている。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ている可能性がある。最後にCyp40は転写因子c-Mybに結合し、これを不活性化するが、この作用はシクロスポリンによって逆転されることが示されている。この効果は、転写、形質転換および分化の調節でのCyPの関与を意味している（Bergsma, D. J. ら（1991）J. Biol. Chem. 266:23204-23214; Hunter, T.（1998）Cell 92:141-143; および Leveron, J. D. および Ness, S. A.（1998）Mol. Cell 1:203-211）。

30

【0021】

プロリンリッチなγ-カルボキシグルタミン酸（Gla）タンパク質（PRGP）は、ビタミンK依存性の単一膜貫通領域をもつ膜内在性タンパク質ファミリーのメンバーである。これらのタンパク質は、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の一つまたは二つのコピーがあり、この配列PPXYはシグナル伝達、細胞周期進行、およびタンパク質代謝回転のような細胞の多様な機能に関与する種々のタンパク質に存在するモチーフである。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を含むタンパク質には血液凝固に関与する血漿タンパク質が含まれている。これらのタンパク質はプロトロンビン、タンパク質C、SおよびZ、並びに血液凝固第VII因子、第IX因子および第X因子である。オステオカルシン（骨-Glaタンパク質、BGP）および基質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.

40

50

Biochem. 19:1-7; C. Vermeer (1990) Biochem. J. 266:625-636)。

【0022】

ショウジョウバエ遺伝子 *crossveinless 2* は推定シグナル配列または膜貫通配列を持ち、また分子内結合と分子間結合の形成を調節すると知られる Von Willebrand (フォンウィレブランド) Factor D ドメインの一部と、BMP 様 (骨形態発生タンパク質) リガンドを結合すると知られる 5 ヶ所のシステインリッチドメインを有することを特徴としている。これらの機能は、*crossveinless 2* が細胞外経路または分泌経路において直接リガンドのシグナル伝達を増強させる作用をし、したがって、血管特異化においてある役割を果たすとされている BMP 様シグナル伝達経路に関与することを示唆する (Conley, C. A. ら、(2000) Development 127:3947-3959)。脊椎動物およびショウジョウバエの胚における背部および腹部の形態化には位置情報勾配を生成する細胞外タンパク質の保存された系を必要とする。

【0023】

シグナルタンパク質を含む他のタンパク質は、痙攣薬ペンチレンテトラゾールによって、発現が増加される脳特異的 cDNA である癲癇発作関連遺伝子 SEZ-6 によってコードされる。SEZ-6 タンパク質は脳および小脳において発現される。SEZ-6 は、シグナルペプチドと単一膜貫通ドメインに加えて、5 つの短いコンセンサスリピート (SCR または *sushi* ドメイン) および二つの CUB (補体 C1r/s 様リピート) ドメインを含む (Shimizu-Nishikawa, K. ら (1995) Biochem. Biophys. Res. Commun. 216:382-389)。新規の分泌タンパク質およびそれらをコードするポリヌクレオチドの発見により、新規の組成物を提供することで当分野の要望に応えることができる。この新規の組成物は、細胞増殖異常、自己免疫/炎症性の疾患、心血管疾患、神経の疾患、および発生障害の診断・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及びアミノ酸配列の発現における外来性化合物の効果についての評価にも有用である。

【0024】

(発明の要約)

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」、「SECP-24」、「SECP-25」、「SECP-26」、「SECP-27」、「SECP-28」、「SECP-29」、「SECP-30」、「SECP-31」、「SECP-32」、「SECP-33」、「SECP-44」と呼ぶ分泌タンパク質である精製されたポリペプチドを提供する。或る実施態様において本発明は、(a) SEQ ID NO: 1-44 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: 1-44 を有する群から選択したアミノ酸配列と少なくとも 90% が同一であるアミノ酸配列を有する天然のポリペプチド、(c) SEQ ID NO: 1-44 を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または (d) SEQ ID NO: 1-44 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抗原性断片を含む群から選択した実質上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: 1-44 のアミノ酸配列を含む実質上単離されたポリペプチドを提供する。

【0025】

また、本発明は (a) SEQ ID NO: 1-44 を有する群から選択したアミノ酸配列、(b) SEQ ID NO: 1-44 を有する群から選択したアミノ酸配列と少なくとも

90%の相同性を有する天然のアミノ酸配列、(c) SEQ ID NO: 1-44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: 1-44を有する群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドをコードするような実質上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: 1-44を有す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: 45-88を有する群から選択される。

【0026】

本発明は更に、(a) SEQ ID NO: 1-44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: 1-44を有する群から選択したアミノ酸配列と少なくとも90%の相同性を有するアミノ酸配列からなる天然のポリペプチド、(c) SEQ ID NO: 1-44を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1-44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抗原性断片から構成される群から選択されたポリペプチドをコードするポリヌクレオチドと機能的に結合したプロモーター配列を有する組換えポリヌクレオチドを提供する。一実施態様では、本発明は組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様では、本発明は組換えポリヌクレオチドを含む遺伝形質転換体を提供する。

【0027】

また、本発明は(a) SEQ ID NO: 1-44を有する群から選択したアミノ酸配列、(b) SEQ ID NO: 1-44を有する群から選択したアミノ酸配列と少なくとも90%の相同性を有する天然のアミノ酸配列、(c) SEQ ID NO: 1-44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: 1-44を有する群から選択したアミノ酸配列の免疫抗原性断片を含む実質上単離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受容する過程とを有し、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結合したプロモーター配列を有する。

【0028】

本発明は更に、(a) SEQ ID NO: 1-44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、(b) SEQ ID NO: 1-44を有する群から選択したアミノ酸配列と少なくとも90%同一であるアミノ酸配列を有する天然のポリペプチド、(c) SEQ ID NO: 1-44を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1-44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抗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実質上単離された抗体を提供する。

【0029】

本発明は更に、(a) SEQ ID NO: 45-88を有する群から選択したポリヌクレオチド配列、(b) SEQ ID NO: 45-88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、(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、または(e) (a) ~ (d) のRNA等価物を含む実質上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少なくとも60の連続したヌクレオチドを有する。

【0030】

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: 45-88を有する群から選択したポリヌクレオチド配列、(b) SEQ ID NO: 45-88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、

(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、または (e) (a) ~ (d) の RNA 等価物を含む実質上単離されたポリヌクレオチドを提供する。検出方法は、(a) サンプル中の標的ポリヌクレオチドに相補的な配列からなる少なくとも 20 の連続したヌクレオチドを含むプローブを用いて該サンプルをハイブリダイズする過程と、(b) ハイブリダイゼーション複合体の存在・不存在を検出し、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、プローブと標的ポリヌクレオチド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、プローブは標的ポリヌクレオチドに特異的にハイブリダイズする。一実施態様では、プローブは少なくとも 60 の連続したヌクレオチドを含む。

【0031】

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: 45-88 を有する群から選択したポリヌクレオチド配列、(b) SEQ ID NO: 45-88 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、(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、または (e) (a) ~ (d) の RNA 等価物を含む群から選択されたポリヌクレオチド配列を有する。検出方法は、(a) ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、(b) 標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含む。

【0032】

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成分を提供し、有効量のポリペプチドは、(a) SEQ ID NO: 1-44 を有する群から選択したアミノ酸配列、(b) SEQ ID NO: 1-44 を有する群から選択したアミノ酸配列と少なくとも 90 % の相同性を有する天然のアミノ酸配列、(c) SEQ ID NO: 1-44 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: 1-44 を有する群から選択したアミノ酸配列の免疫抗原性断片を含む。一実施例では、SEQ ID NO: 1-44 からなる一群から選択されたアミノ酸配列を含む組成物を提供する。更に、本発明は、患者にこの組成物を投与することを含む、機能的 SECP の発現の低下に関連した疾患やその症状の治療方法を提供する。

【0033】

本発明はまた、(a) SEQ ID NO: 1-44 を有する群から選択したアミノ酸配列、(b) SEQ ID NO: 1-44 を有する群から選択したアミノ酸配列と少なくとも 90 % の相同性を有する天然のアミノ酸配列、(c) SEQ ID NO: 1-44 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: 1-44 を有する群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドのアゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドを有するサンプルを化合物に曝す過程と、(b) サンプル中のアゴニスト活性を検出する過程とを含む。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的 SECP の発現の低下に関連した疾患やその症状の治療方法を提供する。

【0034】

本発明は更に、(a) SEQ ID NO: 1-44 を有する群から選択したアミノ酸配列、(b) SEQ ID NO: 1-44 を有する群から選択したアミノ酸配列と少なくとも 90 % の相同性を有する天然のアミノ酸配列、(c) SEQ ID NO: 1-44 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: 1-44 を有する群から選択したアミノ酸配列の免疫抗原性断片を含むポリペプチドのアンタゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドを含むサンプルを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とを含む。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成分を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的 S E C P の過剰な発現に関連した疾患やその症状の治療方法を提供する。

【0035】

本発明は更に、(a) S E Q I D N O : 1-44 を有する群から選択したアミノ酸配列、(b) S E Q I D N O : 1-44 を有する群から選択したアミノ酸配列と少なくとも 90% の相同性を有する天然のアミノ酸配列、(c) S E Q I D N O : 1-44 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) S E Q I D N O : 1-44 を有する群から選択したアミノ酸配列の免疫抗原性断片を含む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 つの試験化合物に結合させる過程と、(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。 10

【0036】

更に本発明は、(a) S E Q I D N O : 1-44 からなる一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) S E Q I D N O : 1-44 からなる一群から選択されたアミノ酸配列と 90% 以上の配列同一性を有するアミノ酸配列からなる天然のポリペプチドと、(c) S E Q I D N O : 1-44 からなる一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) S E Q I D N O : 1-44 とからなる一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一群から選択されたポリペプチドの活性を調節する化合物をスクリーニングする方法を提供する。このスクリーニング方法は、(a) このポリペプチドを、その活性が許容される条件下で少なくとも 1 つの化合物と結合させるステップと、(b) この試験化合物の存在下でのこのポリペプチドの活性を評価するステップと、(c) この試験化合物の存在下でのこのポリペプチドの活性と、この試験化合物の不在下でのこのポリペプチドの活性とを比較するステップとを含み、この試験化合物の存在下でのこのポリペプチドの活性の変化が、このポリペプチドの活性を調節する化合物の存在を示唆するという特徴を有する。 20

【0037】

本発明は更に、標的ポリヌクレオチドの変異発現の有効性を確認するために化合物をスクリーニングする方法を提供する。標的ポリヌクレオチドは、S E Q I D N O : 45-88 を有する群から選択した配列を含む。スクリーニング方法は、(a) 標的ポリヌクレオチドを含むサンプルを化合物に曝す過程と、(b) 標的ポリヌクレオチドの変異発現を検出する過程とを含む。 30

【0038】

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) S E Q I D N O : 45-88 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) S E Q I D N O : 45-88 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% の相同性を有するポリヌクレオチド配列を含む天然のポリヌクレオチド、(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、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、上記プローブと生物学的サンプル中の標的ポリヌクレオチドの間に特定のハイブリダイゼーション複合体が形成されるような条件下で発生し、上記標的ポリヌクレオチドは、(i) S E Q I D N O : 45-88, 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) S E Q I D N O : 45-88 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% の相同性を有するポリヌクレオチド配 40 50

列を含む天然のポリヌクレオチド、(i i i) (i) のポリヌクレオチドに相補的な配列を有するポリヌクレオチド、(i v) (i i) のポリヌクレオチドに相補的なポリヌクレオチド、(v) (i) ~ (i v) のRNA等価物を含む群から選択する。或いは、標的ポリヌクレオチドは、上記(i) ~ (v)を含む群から選択したポリヌクレオチド配列の断片と、(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。

【0039】

(本発明の記載について)

10

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

【0040】

請求の範囲及び明細書中で用いている単数形の「或る」及び「その(この)」の表記は、文脈から明らかにそうでないとされる場合を除いて複数のものを指す場合もあることに注意しなければならない。従って、例えば「或る宿主細胞」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

20

【0041】

本明細書中で用いる全ての専門用語及び科学用語は、特に定義されている場合を除き、当業者に一般に理解されている意味と同じ意味を有する。本明細書で説明するものと類似あるいは同等の任意の装置、材料及び方法を用いて本発明の実施または試験を行うことができるが、ここでは好適な装置、材料、方法について説明する。本発明で言及する全ての刊行物は、刊行物中で報告されていて且つ本発明に関係があるであろう細胞、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

【0042】

(定義)

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

【0043】

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。

40

【0044】

用語「対立遺伝子変異配列」は、ATRS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の突然変異から作製し得る。また、変異RNAまたはポリペプチドからも作製し得る。ポリペプチドの構造または機能は、変異することもしないこともある。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。一般に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、ヌクレオチドの自然欠失、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。

【0045】

50

S E C Pをコードする「変異」核酸配列は、様々なヌクレオチドの欠失、挿入、或いは置換が起こっても、S E C Pと同じポリペプチド或いはS E C P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、S E C P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適合或いは予期しないハイブリダイゼーション、並びにS E C Pをコードするポリヌクレオチドの特定のオリゴヌクレオチドプローブを用いて容易に検出可能な或いは検出困難な多形性を含む。コードされたタンパク質も変異され得り、サイレント変化を生じS E C Pと機能的に等価となるアミノ酸残基の欠失、挿入、或いは置換を含み得る。意図的なアミノ酸置換は、生物学的或いは免疫学的にS E C Pの活性が保持される範囲で、残基の極性、電荷、溶解度、疎水性、親水性、及び／または両親媒性についての類似性に基づいて成され得る。例えば、負に帯電したアミノ酸にはアスパラギン酸及びグルタミン酸があり、正に帯電したアミノ酸にはリジン及びアルギニンがある。親水性値が近似し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、セリンとスレオニンがある。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、ロイシンとイソロイシンとバリン、グリシンとアラニン、フェニルアラニンとチロシンがある。

10

【0046】

「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。

20

【0047】

「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。

【0048】

用語「アンタゴニスト」は、S E C P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、S E C Pに直接相互作用するか、或いはS E C Pが関与する生物学的経路の成分と作用して、S E C P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。

【0049】

「抗体」の語は、エピトープの決定基と結合することができる、そのままの免疫グロブリンやその断片、例えばF a、F（a b'）₂ 及びF v断片を指す。S E C P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、好みに応じて担体タンパク質に抱合することも可能である。通常用いられる担体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イガイのヘモシアニン（KLH）等がある。結合ペプチドは、動物を免疫化するために用いる。

30

【0050】

「抗原決定基」の語は、特定の抗体と接触している分子の領域（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（即ち免疫応答を誘導するために用いられる免疫原）と競合し得る。

40

【0051】

「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成分を指す。アンチセンス成分には、DNAや、RNAや、ペプチド核酸（PNA）や、ホスホロチオ酸、メチルホスホン酸またはベンジルホスホン酸等の修飾されたバックボーン連鎖を有するオリゴヌクレオチドや、2'-メトキシエチル

50

糖または2'-メトキシエトキシ糖等の修飾された糖類を有するオリゴヌクレオチドや、或いは5-メチルシトシン、2-デオキシウラシルまたは7-デアザ-2'-デオキシグアノシン等の修飾された塩基を有するオリゴヌクレオチドが含まれる。アンチセンス分子は、化学合成または転写を含む任意の方法で製造することができる。相補的アンチセンス分子は、ひとたび細胞に導入されたら、細胞が形成した天然の核酸配列と塩基対を形成し、転写または翻訳を妨害する二重鎖を形成する。「負」または「マイナス」という表現は、ある参考DNA分子のアンチセンス鎖を意味し、「正」または「プラス」という表現は同センス鎖を意味することがある。

【0052】

「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。

10

【0053】

「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'-A-G-T-3'」は、相補配列「3'-T-C-A-5'」に結合する。

【0054】

「所定のポリヌクレオチド配列を含む成分」及び「所定のアミノ酸配列を含む成分」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。この成分には、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)及びその他の構成エレメント(例えばデンハート液、脱脂粉乳、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。

20

【0055】

「コンセンサス配列」は、不要な塩基を分離するためにDNA配列の解析を繰り返し行い、XL-PCRキット(PE Biosystems, Foster City CA)を用いて5'及び/または3'の方向に伸長され、再度シーケンシングされた核酸配列、またはGELVIEW断片構築システム(GCG, Madison, WI)またはPhrap(University of Washington, Seattle WA)等の断片構築用のコンピュータプログラムを用いて1つ或いはそれ以上の重複するcDNAやEST、またはゲノムDNA断片から構築された核酸配列を指す。伸長及び構築の両方を行ってコンセンサス配列を決定する配列もある。

30

【0056】

「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。

40

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	10
Gln	Asn, Glu, His	
Glu	Asp, Gln, His	
Gly	Ala	
His	Asn, Arg, Gln, Glu	
Ile	Leu, Val	
Leu	Ile, Val	20
Lys	Arg, Gln, Glu	
Met	Leu, Ile	
Phe	His, Met, Leu, Trp, Tyr	
Ser	Cys, Thr	
Thr	Ser, Val	
Trp	Phe, Tyr	30
Tyr	His, Phe, Trp	
Val	Ile, Leu, Thr	

【0057】

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βシートやα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)側鎖の大部分を保持する。

【0058】

40

「欠失」は、結果的に1個若しくは数個のアミノ酸またはヌクレオチドが失われてなくなると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。

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

50

【0059】

「検出可能な標識」は、測定可能な信号を発生することができ、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。

「示差発現」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。

【0060】

用語「断片」は、S E C P または S E C P 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、画定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、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、分子の特定領域から選択的に選択し得る。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。

【0061】

SEQ ID NO: 45-88の断片には、固有のポリヌクレオチド配列領域が含まれる。

この領域は、SEQ ID NO: 45-88を特異的に同定するものであり、例えばこの断片を得たゲノム中の他の配列とは異なるものである。SEQ ID NO: 45-88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: 45-88に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: 45-88の断片の正確な長さ及び断片に対応するSEQ ID NO: 45-88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。

【0062】

SEQ ID NO: 1-44の断片はSEQ ID NO: 45-88の断片によってコードされている。SEQ ID NO: 1-44の断片はSEQ ID NO: 1-44を特異的に同定する固有のアミノ酸配列の領域を含む。例えば、SEQ ID NO: 1-44の断片は、SEQ ID NO: 1-44を特異認識する抗体を産出するための免疫抗原性ペプチドとして有用である。SEQ ID NO: 1-44の断片及び断片に対応するSEQ ID NO: 1-44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。

【0063】

「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。

【0064】

「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、または配列同一性を意味する。

【0065】

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

10

20

30

40

50

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。

【0066】

ポリヌクレオチド配列間の一致率は、MEGALIGN version 3.12e 配列アラインメントプログラムに組込まれているようなCLUSTAL Vアルゴリズムのデフォルトのパラメータを用いて決定できる。このプログラムは、LASERGENE ソフトウェアパッケージ（一組の分子生物学的分析プログラム）（DNASTAR, Madison WI）の一部である。このCLUSTAL Vは、Higgins, D. G. 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の対のアライメントの場合、デフォルトパラメータは、Ktuple=2、gap penalty=5、window=4、「diagonals saved」=4と設定する。「重み付けされた」残基重み付け表が、デフォルトとして選択された。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「類似率」として一致率を報告する。

10

【0067】

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. ら (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対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 と blastp（後述）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、ギャップ及びデフォルト設定に設定された他のパラメータと共に用いる。例えば、2つのヌクレオチド配列を比較するために、デフォルトパラメータとして設定された「BLAST 2 Sequences」ツールVersion 2.0.12（2000年4月21日）を用いてblastnを実行してもよい。デフォルトパラメータの設定例を以下に示す。

20

30

【0068】

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 および Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

10

【0069】

一致率は、完全に画定された（例えば特定の配列番号で画定された）配列長さと比較して測定し得る。或いは、より短い長さ、例えばより大きな画定された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。

20

【0070】

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。

【0071】

ポリペプ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以上のポリペプチド配列間で一致する残基の割合を意味する。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。

30

【0072】

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を用いる対方式のポリペプチド配列のアライメントの場合、デフォルトパラメータは、K t u p l e = 1、g a p p e n a l t y = 3、w i n d o w = 5、及び「d i a g o n a l s s a v e d」= 5と設定する。PAM250マトリクスが、デフォルトの残基重み付け表として選択される。ポリヌクレオチドアラインメントと同様に、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。

40

【0073】

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対で比較をする場合、ある者は、デフォルトパラメータで設定された「BLAST 2 Sequences」ツールVersion 2.0.12（Apr-21-2000）でblastpを使用するであろう。デフォルトパラメータの設定例を以下に

50

示す。

【0074】

Matrix: BLOSUM62

Open Gap: 11 および Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

10

【0075】

一致率は、完全に画定された（例えば特定の配列番号で画定された）ポリペプチド配列の長さと比較して測定し得る。一致率は、配列或いは、より短い長さ、例えばより大きな画定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。

20

【0076】

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

【0077】

「ヒト化抗体」の語は、非抗体結合領域におけるアミノ酸配列はヒト抗体により近づくように変異させた抗体分子であって、本来の結合能力はそのまま保持しているような抗体分子を指す。

30

【0078】

「ハイブリダイゼーション」は、所定のハイブリダイゼーション条件下で塩基対を形成することによって、一本鎖ポリヌクレオチドが相補的鎖とアニーリングするプロセスを指す。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同性を共有することを示すものである。特異的ハイブリダイゼーション複合体は許容される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、更に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、本技術分野における当業者が慣例的に決定する。許容条件はハイブリダイゼーション実験の間は一定でよいが、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中に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℃で、約6×SSC、約1%（w/v）のSDS、並びに約100μg/mlのせん断して変性したサケ精子DNAが含まれる。

40

【0079】

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（Tm）より約5～20℃低くなるように選択する。このTm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。Tmを計算する式及び核酸のハイブリダイゼーシ

50

ョン条件はよく知られており、Sambrook ら (1989) Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NYに記載されており、特に2巻の9章を参照されたい。

【0080】

本発明のポリヌクレオチド間の高いストリンジェンシーのハイブリダイゼーションでは、約0.2 x SSC及び約0.1%のSDSの存在下、約68℃で1時間の洗浄過程を含む。別法では、65℃、60℃、55℃、または42℃の温度で行う。SSC濃度は、約0.1%のSDS存在下で、約0.1~2 x SSCの範囲で変化し得る。通常は、遮断剤を用いて非特異ハイブリダイゼーションを阻止する。このような遮断試薬には、例えば、約100~200 µg/mlのせん断され変性したサケ精子DNAが含まれる。例えばRNAとDNAのハイブリダイゼーションのような特定条件下では、有機溶剤、例えば約35~50% v/vの濃度のホルムアミドを用いることもできる。洗浄条件の有用なバリエーションは、当業者には自明であろう。ハイブリダイゼーションは、特に高ストリンジェント条件下では、ヌクレオチド間の進化的な類似性を示唆し得る。このような類似性は、ヌクレオチド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。

10

【0081】

「ハイブリダイゼーション複合体」の語は、相補的塩基対間の水素結合の形成力によって2つの核酸配列間に形成された複合体を指す。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質であって細胞若しくはその核酸が固定される基質)に固定されているような2つの核酸配列間に形成され得る。

20

【0082】

「挿入」及び「付加」の語は、1個若しくは数個のアミノ酸残基またはヌクレオチド配列を各々付加す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。

【0083】

「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。

30

【0084】

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

【0085】

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

40

【0086】

「エレメント」または「アレイエレメント」の語は、マイクロアレイ上に定義された固有の位置にあるハイブリダイズ可能なポリヌクレオチド、ポリペプチドその他の化合物を指す。

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

【0087】

「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

50

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。

【0088】

「機能的に結合した」は、第1核酸配列が第2核酸配列と機能的な関係があるように配置された状態を指す。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結合している。機能的に結合したDNA配列は非常に近接するか、或いは連続し得る。そして、2つのタンパク質コード領域を結合する必要がある場合は、同一のリーディングフレーム内にある。

【0089】

「ペプチド核酸」（PNA）は、アンチセンス分子または抗遺伝子物質であって、リジン末端とするアミノ酸残基のペプチドバックボーンに結合した、少なくとも約5ヌクレオチドの長さのオリゴヌクレオチドからなるものを指す。末端のリジンは、成分に溶解性を与える。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拡張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。

【0090】

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

【0091】

「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鎖に延在し得る。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。

【0092】

本発明に用いるようなプローブ及びプライマーは通常、既知の配列の少なくとも15の連続したヌクレオチドを含んでいる。特異性を高めるために長めのプローブ及びプライマー、例えば開示した核酸配列の少なくとも20、25、30、40、50、60、70、80、90、100または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。

【0093】

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. ら (1989) Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NY、Ausubel, F. M. ら, (1987) Current Protocols in Molecular Biology, Greene Publi. Assoc. & Wiley-Intersciences, New York NY、Innis ら (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, Wh

10

20

30

40

50

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。

【0094】

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キロベ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（マサチューセッツ州ケンブリッジのWhitehead Institute/MITゲノム研究センターから一般向けに入手可能）ではユーザーが「ミスプライミング・ライブラリ」をインプットすることができ、ここでプライマー結合部位として避けたい配列はユーザーが指定する。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、各自のソースから得てユーザー固有のニーズを満たすように変更してもよい。）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、固有な、および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。

【0095】

「組換え核酸」は天然配列ではない配列であるか或いは人為的に組み合わせなければ分離しているような配列の2以上のセグメントを人為的に組み合わせで産出した配列を有する配列である。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、例えばの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異核酸も含む。しばしば組換え核酸には、プロモーター配列に機能的に結合した核酸配列が含まれる。このような組換え核酸は、ベクターの不可欠なエレメントであって例えばある細胞を形質転換するために用いられるようなものであり得る。

【0096】

或いはこのような組換え核酸は、ウイルスベクターの一部であって、例えばワクシニアウイルスに基づくものであり得る。そのようなワクシニアウイルスは哺乳動物に接種され、その組換え核酸が発現されて、その哺乳動物ない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。

【0097】

「調節エレメント」は、通常は遺伝子の未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'の未翻訳領域（UTR）を含む。調節エレメントは、転写、翻訳またはRNA安定性を調節する宿主またはウイルスタンパク質と相互作用する。

【0098】

「レポーター分子」とは、核酸、アミノ酸または抗体を標識するのに用いられる化学的または生化学的成分である。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。

【0099】

DNA配列に関する「RNA等価物」は、窒素塩基チミンが全てウラシルに置換されていることと、糖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構成されていることを除いて、参照DNA配列と同一のヌクレオチド線形配列から構成されている。

【0100】

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

【0101】

「特異結合」または「特異的に結合する」の語は、タンパク質またはペプチドと、アゴニスト、抗体、アンタゴニスト、小分子、任意の天然成分または合成結合成分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、標識された遊離した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、標識されていないA）を含むポリペプチドの存在が、抗体に結合している標識されたAの量を低減させる。

【0102】

「実質上精製された」の語は、自然環境から取り除かれ、或いは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、自然に会合するその他の構成エレメントの少なくとも約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%が遊離しているものを指す。

【0103】

「置換」は、1個若しくは数個のアミノ酸またはヌクレオチドを各々別のアミノ酸またはヌクレオチドに置換することを意味する。

【0104】

「基質」は、任意の好適な固体または半固体の支持体を指すものであって、膜、フィルタ、チップ、スライド、ウェハ、ファイバー、磁性または非磁性ビーズ、ゲル、管、プレート、ポリマー、微細粒子、毛管が含まれる。基質は、凹み、溝、ピン、チャンネル、孔等、様々な表面形態を有することができ、基質表面にはポリヌクレオチドやポリペプチドが結合する。

【0105】

「転写イメージ」は、所与の時間、条件での固有の細胞タイプまたは組織による遺伝子発現の集合的パターンを指す。

【0106】

「形質転換」は、外来性のDNAが受入細胞に入り込むプロセスを表す。形質転換は、本技術分野で知られている種々の方法に従って自然条件または人工条件下で生じ得るものであり、外来性の核酸配列を原核宿主細胞または真核宿主細胞に挿入する任意の既知の方法を基にし得る。形質転換の方法は、形質転換する宿主細胞の種類によって選択する。限定するものではないが形質転換方法には、バクテリオファージまたはウイルス感染、電気穿孔法（エレクトロポレーション）、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、限られた時間に一時的に導入DNA若しくは導入RNAを発現する細胞も含まれる。

【0107】

ここで用いる「遺伝形質転換体」とは任意の有機体であり、限定するものではないが動植物を含み、有機体の1個若しくは数個の細胞が、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形質転換技術によって導入された異種核酸を有する。核酸の細胞への導入は、直接または間接的に、細胞の前駆物質に導入することによって、計画的な遺伝子操作によって、例えば微量注射法によって或いは組換えウイルスの導入によって行う。遺伝子操作の語は、古典的な交雑育種或いは *in vitro* 受精を指すものではなく、組換えDNA分子の導入を指すものである。本発明に基づいて予期される遺伝形質転換体には、バクテリア、シアノバクテリア、真菌及び動植物がある。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をこのような有機体に移入する技術はよく知られており、前出の Sambrook ら (1989) 等の参考文献に与えられている。

10

【0108】

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」変異体として説明し得る。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交互スプライシングによって通常多数の或いは僅かな数のポリヌクレオチドを有することになる。対応するポリペプチドは、追加機能ドメイン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、種相互に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。多形性変異体は、与えられた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。また、多形性変異体は、1つのヌクレオチド塩基によってポリヌクレオチド配列が変化する「単一ヌクレオチド多型性」(SNP)を含み得る。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。

20

30

【0109】

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画定される。ここで、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。

【0110】

(発明)

本発明は、新規のヒト分泌タンパク質(SECP)及びSECPをコードするポリヌクレオチドの発見に基づき、これらの組成物を利用した細胞増殖異常、自己免疫/炎症性の疾患、心血管疾患、神経の疾患、および発生障害の診断、治療、及び予防に関する。

40

【0111】

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及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)と相関する。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

50

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。

【0112】

表2は、GenBank タンパク質 (genpept) データベースに対する BLAST 分析によって同定されたような、本発明のポリペプチドとの相同性を有する配列を示している。列1および列2はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号 (Polypeptide SEQ ID NO :) およびそれに対応する Incyte ポリペプチド配列番号 (Incyte Polypeptide ID) を示す。列3は、GenBank の最も近い相同体の GenBank の識別番号 (Genbank ID NO :) を示す。列4は、各ポリペプチドとその GenBank 相同体との間の一致を表す確率スコアを示す。列5は、GenBank 相同体のアノテーションを示し、更に該当箇所には適当な引用文も示す。これらを引用することを以て本明細書の一部とする。

10

【0113】

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と2は発明した個々のポリペプチドをポリペプチド配列識別番号 (SEQ ID NO) と対応する 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 (signature) 配列、ドメイン、およびモチーフを含むアミノ酸残基を示す。特にシグナルペプチド (「Signal _ peptide」または「Signal _ cleavage」として示す) の位置を示す。列7は、タンパク質の構造/機能の分析のための分析方法を示し、該当箇所にはさらに分析方法に利用した検索可能なデータベースを示す。

20

【0114】

表2及び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 : 1 はヒト UDP-GalNAc : ポリペプチド N-アセチルガラクトサミン転移酵素 (GenBank ID g971461) と51%の同一性を有することが Basic Local Alignment Search Tool (BLAST) によって確認された。(表2参照) BLAST 確率スコアは $1.5e-141$ であり、これは保存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 : 1 はまた、シグナルペプチドと膜貫通ドメインを含むことが隠れマルコフモデル (HMM) に基づく方法によって確認された(表3参照)。同様に SPSCAN 分析もまた、SEQ ID NO : 1 に N 末端シグナルペプチドがあることを示した。総合すると、これらの証拠は SEQ ID NO : 1 は分泌経路に N-アセチルガラクトサミン転移酵素として存在することを示す。

30

別の例において、SEQ ID NO : 2 はマウス てんかん関連遺伝子作成物6のタイプ2前駆体 (GenBank ID g1139548) に90%同一であり、これは Basic Local Alignment Search Tool (BLAST) によって確認された。(表2参照) BLAST 確率スコアは0.0であるが、これは観測されたポリペプチド配列が偶然に得られる確率を示す。SEQ ID NO : 2 はまた、5つのスシ (sushi) ドメインと2つの CUB 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) さらに、SEQ ID NO : 2 はシグナルペプチドと一つの膜貫通ドメインを有することが、HMMER 分析によって同定された。

40

例えば、SEQ ID NO : 3 はニワトリのリゾチーム (GenBank ID g4467410) と43%の同一性を有することが Basic Local Alignment

50

Search Tool (BLAST) によって確認された。(表2参照) BLAST 確率スコアは 5.2×10^{-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: 3 はまた、G-リゾチームシグナチャドメインを有し、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モチーフの PRINTS 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) PFAM、PRODOM および DOMO データベースよりのデータは、SEQ ID NO: 3 がリゾチームであるという、さらに実証的な証拠を提供する。

例えば、SEQ ID NO: 17 は SPScan および隠れマルコフモデル (HMM) に基づいた分析によって確認されるようにシグナルペプチドを有している。例えば、SEQ ID NO: 17 はヒト免疫グロブリンλ軽鎖 (GenBank ID g33702) と 86% 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 2.2×10^{-6} 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: 17 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、MOTIFS、及び PROFILSCAN 解析よりのデータは、SEQ ID NO: 17 が分泌免疫グロブリンである、さらに実証的な証拠を提供する。使用できる証拠によって SEQ ID NO: 19 もまた、分泌免疫グロブリンであることが示された。

【0115】

例えば、SEQ ID NO: 38 はヒト免疫グロブリンλ軽鎖 (GenBank ID g33718) と 95% の同一性を有することが Basic Local Alignment Search Tool (BLAST) によって示された。(表2参照) BLAST 確率スコアは 5.2×10^{-114} 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: 38 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、MOTIFS、及び PROFILSCAN 解析よりのデータは、SEQ ID NO: 38 が分泌タンパク質であり、特に免疫グロブリンである、さらに実証的な証拠を提供する。SEQ ID NO: 4-16、SEQ ID NO: 18-37 および SEQ ID NO: 39-44 は同じような方法で解析し、注釈を付けた。SEQ ID NO: 1-44 を解析するためのアルゴリズム及びパラメータは表7に記述する。

【0116】

表4に示すように、本発明の完全長ポリヌクレオチド配列は、cDNA 配列またはゲノム DNA 由来のコード (エキソン) 配列を用いて、或いはこれら2種類の配列を任意に組み合わせて構築した。列1および列2はそれぞれ、本発明の各ポリヌクレオチドのポリヌクレオチド配列識別番号 (Polynucleotide SEQ ID NO:) およびそれに対応する Incyte ポリヌクレオチドコンセンサス配列番号 (Incyte Polynucleotide ID) を示す。列3は、各ポリヌクレオチド配列の長さ (塩基対単位) を示す。列4は、例えば、SEQ ID NO: 45-88 を同定するため、或いは SEQ ID NO: 45-88 と、関連するポリヌクレオチド配列とを区別するためのハイブリダイゼーションまたは増幅技術に有用なポリヌクレオチド配列の断片を示す。列5はcDNA 配列、ゲノムDNAから予想されたコード配列 (エキソン) 及び／またはcDNA 及びゲノムDNAを共に有する配列集合に対応する識別番号を示している。これらの配列は、本発明の完全長ポリヌクレオチド配列の構築に用いた。表4の列6および列7はそれぞれ、列5の配列に対応するcDNA 配列およびゲノム配列の開始ヌクレオチド (5') 位置および終了ヌクレオチド (3') 位置を示す。

【0117】

表4の列5の識別番号は、特に例えば Incyte cDNA とそれに対応する cDNA ライブラリを指す場合もある。例えば、6735891H1 は Incyte cDNA 配列の識別番号であり、LIVRTUT13 はそれが由来する cDNA ライブラリの識別番号である。cDNA ライブラリが示されていない Incyte cDNA は、プールされている cDNA ライブラリ（例えば、71013085V1）に由来する。または、列5の識別番号は、完全長ポリヌクレオチド配列の構築に用いられた GenBank の cDNA または EST（例えば、g1496797）の識別番号の場合もある。或いは列5の識別番号は、ゲノム DNA の Genscan 分析により予想されるコード領域を指す場合もある。この Genscan 推定コード配列は、配列を組み立てる前に編集する場合がある。（例4を参照）。または、列5の識別番号は、「exon-stitching」アルゴリズムによって合わされた cDNA および Genscan 推定エキソンの両方からなる群を指す場合もある（実施例5を参照）。または、列5の識別番号は、「exon-stretching」アルゴリズムによって cDNA および Genscan 推定エキソンの両方からなる群の場合もある（実施例5を参照）。場合によっては、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の列5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、関連する Incyte cDNA 識別番号は示さなかった。

【0118】

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る 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及びベクターを表5に示し、表6で説明している。

【0119】

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

【0120】

本発明はまた、SECPをコードするポリヌクレオチドを提供する。特定の実施例において、本発明は、SECPをコードする SEQ ID NO: 45-88 からなる一群から選択された配列を含むポリヌクレオチド配列を提供する。配列表に示した SEQ ID NO: 45-88 のポリヌクレオチド配列は、窒素系塩基のチミンがウラシルに置換され、糖鎖の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等価 RNA 配列を含む。

【0121】

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70%のポリヌクレオチド配列同一性、或いは少なくとも85%のポリヌクレオチド配列同一性、更には少なくとも95%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: 45-88 からなる群から選択された配列を含むポリヌクレオチド配列の変異体を含み、これは、SEQ ID NO: 45-88 からなる群から選択された核酸配列に少なくとも約70%、或いは少なくとも約85%、または少なくとも約95%のポリヌクレオチド配列同一性を有する。前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できる。

【0122】

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、既知の自然発生する任意の遺伝子のポリヌクレオチド配列と最小の類似性しか有しないものも含まれることを、当業者は理解するであろう。したがって本発明には、選択可

能なコドンに基づく組合せを選ぶことによって産出され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変異体を網羅する。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、全ての変異が明確に開示されていると考慮する。

【0123】

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有するSECP或いはその誘導体をコードするヌクレオチド配列を作るとは有利となり得る。宿主が特定のコドンを利用する頻度に基づいて、特定の真核又は原核宿主に発生するペプチドの発現率を高めるようにコドンを選択することが可能である。コードされたアミノ酸配列を変えないで、SECP及びその誘導体をコードするヌクレオチド配列を実質的に変更する別の理由は、天然の配列から作られる転写物より例えば長い半減期など好ましい特性を備えるRNA転写物を作ることにある。

10

【0124】

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、当分野でよく知られている試薬を用いて、この合成配列を任意の様々な入手可能な発現ベクター及び細胞系中に挿入し得る。

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に突然変異を誘導し得る。

20

更に本発明には、種々のストリンジェントな条件下で、請求項に記載されたポリヌクレオチド配列、特に、SEQ ID NO: 45-88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を参照）。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。

【0125】

DNAシーケンシングの方法は当分野でよく知られており、本発明の何れの実施例も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, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amersham, Pharmacia Biotech, Piscataway NJ) を用いることができる。或いは、例えばELONGASE増幅システム (Life Technologies, Gaithersburg MD) において見られるように、ポリメラーゼと校正エキソヌクレアーゼを併用することができる。好適には、MICROLAB 2200液体転移システム (Hamilton, Reno, NV)、PTC 200サーマルサイクラー (MJ Research, Watertown MA) 及びABI CATALYST 800サーマルサイクラー (Applied Biosystems) 等の装置を用いて配列の調製を自動化する。次に、ABI 373 或いは 377 DNAシーケンシングシステム (Applied Biosystems)、MEGABACE 1000

30

40

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)。

50

【0126】

当分野で周知のPCR法をベースにした種々の方法で、部分的なヌクレオチド配列を利用して、SECPをコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば、使用し得る方法の1つである制限部位PCR法は、ユニバーサルプライマー及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から未知の配列を増幅する方法である（Sarkar, G. (1993) PCR Methods Applic 2: 318-322等を参照）。別の方法に逆PCR法があり、これは広範な方向に伸長させたプライマーを用いて環状化した鋳型から未知の配列を増幅する方法である。鋳型は、既知のゲノム遺伝子座及びその周辺の配列からなる制限断片から得る（Triglia, T. ら (1988) Nucleic Acids Res 16: 8186等を参照）。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. ら (1991) PCR Methods Applic 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連結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. ら (1991) Nucleic Acids Res. 19: 3055-3060等を参照）。更に、PCR、ネステッドプライマー及びPromoterFinder（商標）ライブラリ（Clontech, Palo Alto CA）を用いてゲノムDNAをウォーキングすることができる。この手順は、ライブラリをスクリーニングする必要がなく、イントロン／エキソン接合部を見付けるのに有用である。全てのPCRベースの方法に対して、市販されているソフトウェア、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、温度約68℃～72℃で鋳型に対してアニーリングするようにプライマーを設計し得る。

10

20

【0127】

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。ゲノムライブラリは、5'非転写調節領域への配列の伸長に有用であろう。

30

【0128】

市販されているキャピラリー電気泳動システムを用いて、シークエンシングまたは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、キャピラリーシークエンシングは、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蛍光色素と、放出された波長の検出に利用するCCDカメラとを有し得る。出力／光の強度は、適切なソフトウェア（Applied Biosystems社のGENOTYPER、SEQUENCE NAVIGATOR等）を用いて電気信号に変換し得る。サンプルのロードからコンピュータ分析及び電子データ表示までの全プロセスがコンピュータ制御可能である。キャピラリー電気泳動法は、特定のサンプルに少量しか存在しないようなDNA小断片のシークエンシングに特に適している。

40

【0129】

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のクローン化及び発現に利用可能である。

【0130】

50

種々の目的でSECPをコードする配列を変えるために、当分野で一般的に知られている方法を用いて、本発明のヌクレオチド配列を組換えることができる。この目的には、遺伝子産物のクローン化、プロセッシング及び／または発現の調節が含まれるが、これらに限定されるものではない。遺伝子断片及び合成オリゴヌクレオチドのランダムなフラグメンテーション及びPCR再アセンブリによるDNAシャッフリングを用い、ヌクレオチド配列を組み換えることが可能である。例えば、オリゴヌクレオチド伸介特定部位突然変異誘導を利用して、新規な制限部位の生成、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行う突然変異を導入し得る。

【0131】

本発明のヌクレオチドを、MOLECULAR BREEDING (Maxygen Inc., Santa Clara CA; 米国特許第5,837,458号; Chang, C.-C. 他 (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. 他 (1999) Nat. Biotechnol. 17:259-264; Cramer, A. 他 (1996) Nat. Biotechnol. 14:315-319) などのDNAシャ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伸介再組換えを用いて遺伝子変異体のライブラリを生成するプロセスである。ライブラリはその後、その遺伝子変異体を所望の特性に同定するような選択またはスクリーニングにかける。次にこれらの好適な変異体をプールし、更に反復してDNAシャッフリング及び選択／スクリーニングを行ってもよい。このように、遺伝の多様性は「人為的」品種改良及び急速な分子の進化を経て創生される。例えば、ランダムポイント突然変異を有する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することができる。或いは、所定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て、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、指図された制御可能な方法で最大化させることができる。

【0132】

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. ら (1980) Nucl. Acids Res. Symp. Ser. 7:215-223; 及びHorn, T. 他 (1980) Nucl. Acids Res. Symp. Ser. 225-232を参照）。別法として、化学的方法を用いて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うことができる（Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, 55-60ページ、Roberge, J.Y. ら (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／または化学的方法を用いた他のタンパク質または任意のその一部からの配列との組み合わせにより、天然のポリペプチド配列を有するポリペプチドまたは変異体ポリペプチドを作製することが可能である。

【0133】

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

【0134】

生物学的に活性な S E C P を発現させるために、S E C P をコードするヌクレオチド配列またはその誘導体を好適な発現ベクターに挿入する。この発現ベクターは、好適な宿主に挿入されたコーディング配列の転写及び翻訳の調節に必要なエレメントを含む。これらのエレメントには、ベクター及び S E C P をコードするポリヌクレオチド配列におけるエンハンサー、構成型及び発現誘導型のプロモーター、5' 及び 3' の非翻訳領域などの調節配列が含まれる。このような要素は、長さ及び特異性が様々である。特定の開始シグナルによって、S E C P をコードする配列のより効果的な翻訳を達成することが可能である。このようなシグナルには、A T G 開始コドンと、コザック配列などの近傍の配列が含まれる。S E C P をコードする配列及びその開始コドン、上流の調節配列が好適な発現ベクターに挿入された場合は、更なる転写調節シグナルや翻訳調節シグナルは必要なくなるであろう。しかしながら、コーディング配列或いはその断片のみが挿入された場合は、インフレームの A T G 開始コドンを含む外来性の翻訳調節シグナルが発現ベクターに含まれるようにすべきである。外来性の翻訳要素及び開始コドンは、様々な天然物及び合成物を起源とし得る。発現の効率は、用いられる特定の宿主細胞系に好適なエンハンサーを包含することによって高めることができる (Scharf, D. ら (1994) *Results Probl. Cell Differ.* 20:125-162. 等を参照)。

10

20

30

40

50

【0135】

当業者に周知の方法を用いて、S E C P 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ある (Sambrook, J. ら (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, Plainview NY の 4, 8, 16-17 章、Ausubel, F. M. ら. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY の 9, 13, 16 章等を参照)。

【0136】

種々の発現ベクター／宿主系を利用して、S E C P 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 または pBR322 プラスミド) で形質転換させた植物細胞系、動物細胞系などの微生物がある。(前出の Sambrook、前出の Ausubel、Van Heek e, G. および S. M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509、; Engelhard, E. K. ら (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227、Sandig, V. ら (1996) *Hum. Gene Ther.* 7:1937-1945、Takamatsu, N. (1987) *EMBO J.* 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. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659、Harrington, J. J. ら (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輸送することができる (Di Nicola, M. ら (1998) *Cancer Gen. Ther.* 5(6):350-356、Yu, M. ら (1993) *Proc. Na*

tl. Acad. Sci. USA 90 (13) : 6340-6344、Buller, R. M. ら (1985) Nature 317 (6040) : 813-815; McGregor, D. P. ら (1994) Mol. Immunol. 31 (3) : 219-226、Verma, I. M. and N. Somia (1997) Nature 389 : 239-242等を参照)。本発明は使用される宿主細胞によって限定されるものではない。

【0137】

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、増殖には、PBLUESCRIPT (Stratagene, La Jolla CA) またはpSPORT1プラスミド (GIBCO BRL) などの多機能の大腸菌ベクターを用いることができる。ベクターの複数のクローニング部位にSECPをコードする配列をライゲーションするとlacZ遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。更にこれらのベクターは、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救出、入れ子状態の欠失の生成にも有用であろう (Van Heeke, G. and S. M. Schuster (1989) J. Biol. Chem. 264 : 5503-5509. 等を参照)。例えば、抗体の産生のためなどに多量のSECPが必要な場合は、SECPの発現をハイレベルで誘導するベクターが使用できる。例えば、強力な誘導T5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。

【0138】

SECPの発現に酵母の発現系の使用が可能である。 α 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、酵母菌サッカロミセスセレビジエまたは*Pichia pastoris*に使用可能である。更に、このようなベクターは、発現タンパク質を分泌或いは細胞内への保持のいずれかに誘導し、外来配列の宿主ゲノムへの組込みを可能にし、増殖を安定化する。(前出のAusubel (1995)、Bitter, G. A. 他 (1987) Methods Enzymol. 153 : 516544、Scorer, C. A. 他 (1994) Bio/Technology 12 : 181-184等を参照)。

【0139】

植物系もSECPの発現に使用可能である。SECPをコードする配列の転写は、ウイルスプロモーター、例えば単独或いは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及び19Sプロモーターによって促進される (Takamatsu, N. (1987) EMBO J. 6 : 307-311)。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 (Coruzzi, G. ら (1984) EMBO J. 3 : 16711680; Broglie, R. ら (1984) Science 224 : 838843; および Winter, J. ら (1991) Results Probl. Cell Differ. 17 : 85105等参照)。これらの構成物は、直接DNA形質転換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。(The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, 191196ページを参照)

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結合し得る。ウイルスのゲノムの非必須のE1またはE3領域への挿入により、感染した宿主細胞にSECPを発現する生ウイルスを得ることが可能である (Logan, J. 及びShenk, T. 50

(1984) *Proc. Natl. Acad. Sci.* 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。

【0140】

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の断片を輸送することもできる。治療目的のために約6kb~10MbのHACを作製し、従来の輸送方法(リポソーム、ポリカチオンアミノポリマー、またはベシクル)で供給する(Harrington, J. J. ら (1997) *Nat Genet.* 15:345-355. 等を参照)。

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及び／または内在性の発現要素や、同じ或いは別のベクターの上の選択マーカー遺伝子を含む。ベクターの導入後、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ることができる。選択可能マーカーの目的は選択培地への抵抗性を与えることであり、選択可能マーカーが存在することにより、導入された配列をうまく発現するような細胞の成長及び回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に適した組織培養技術を用いて増殖可能である。

【0141】

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。選択系には、以下のものに限定はしないが、単純ヘルペスウイルスチミジンキナーゼ遺伝子及びアデニンホスホリボシルトランスフェラーゼ遺伝子が含まれ、それぞれtk⁻またはapr⁻細胞において使用される(Wigler, M. ら (1977) *Cell* 11:223-232、Lowy, I. ら (1980) *Cell* 22:817-823等を参照)。また、選択の基礎として代謝拮抗物質、抗生物質或いは除草剤への耐性を用いることができる。例えばdhfrはメトトレキサートに対する耐性を与え、neoはアミノグリコシッドネオマイシン及びG-418に対する耐性を与え、als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. ら (1980) *Proc. Natl. Acad. Sci. USA* 77:3567-3570、Colbere-Garapin, F. ら (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は、文献に記載されている(Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:8047-8051等を参照)。可視マーカー、例えばアントシアニン、緑色蛍光タンパク質(GFP; Clontech)、βグルクロニダーゼ及びその基質β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特定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することが可能である(Rhodes, C. A. (1995) *Methods Mol. Biol.* 55:121-131等を参照)。

【0142】

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。

【0143】

一般に、S E C Pをコードする核酸配列を含み、S E C Pを発現する宿主細胞は、当業者に周知の種々の方法を用いて特定することが可能である。これらの方法には、DNA-DNA 或いはDNA-RNA ハイブリダイゼーションや、PCR法、核酸或いはタンパク質の検出及び／または数量化のための膜系、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生物学的試験法または免疫学的アッセイが含まれるが、これらに限定されるものではない。

【0144】

特異的なポリクローナル抗体またはモノクローナル抗体のどちらかを用いるS E C Pの発現の検出及び計測のための免疫学的な方法は、当分野で周知である。このような技法には、酵素に結合した免疫吸着剤検定法 (E L I S A)、ラジオイムノアッセイ (R I A) 10、フローサイトメーター (F A C S) などがある。S E C P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、競合の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 20のアッセイは、当分野で公知である (Hampton, R. ら (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Sect. IV, Coligan, J. E. ら (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)。

【0145】

多岐にわたる標識方法及び結合方法が、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。S E C Pをコードするポリヌクレオチドに関連する配列を検出するための、標識されたハイブリダイゼーションプローブ或いはPCRプローブを生成する方法には、オリゴ標識化、ニックトランスレーション、末端標識化、または標識されたヌクレオチドを用いるPCR増幅が含まれる。別法として、S E C Pをコードする配列、またはその任意の断片をmRNAプローブを生成する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 30ており、市販もされており、T7、T3またはSP6等の好適なRNAポリメラーゼ及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等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或いは標識には、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、発色剤等がある。

【0146】

S E C P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、或いはその両者に依存する。S E C P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するS E C P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。 40

【0147】

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、脂質化及びアシル化がある。タンパク質の「プレプロ」または「プロ」形を切断するような翻訳後処理を利用して、タンパク質のターゲティング、折りたたみ及び／または活性を特定する 50

ことも可能である。翻訳後の活性のための固有の細胞装置及び特徴のある機構を有する種々の宿主細胞（例えばCHO、HeLa、MDCK、MEK293、WI38等）は、American Type Culture Collection (ATCC, Manassas, VA) から入手可能であり、外来タンパク質の正しい修飾及び処理を確実にするように選択し得る。

【0148】

本発明の別の実施例では、SECPをコードする自然の或いは変更された、または組換えの核酸配列を上記した任意の宿主系の融合タンパク質の翻訳となる異種配列に結合させる。例えば、市販の抗体によって認識できる異種部分を含むキメラSECPタンパク質が、SECPの活性のインヒビターに対するペプチドライブラリのスクリーニングを促進し得る。また、異種タンパク質部分及び異種ペプチド部分も、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及び精製方法は、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及び精製を促進することもできる。

10

20

【0149】

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結合したタンパク質コード配列の転写及び翻訳を結合する。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。

30

【0150】

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つまたは複数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。試験化合物の例には、抗体、オリゴヌクレオチド、タンパク質（例えば受容体）または小分子が挙げられる。

【0151】

一実施例では、このように同定された化合物は、例えばリガンドやその断片などのSECPの天然のリガンド、または天然の基質、構造的または機能的な擬態性または自然結合パートナーに密接に関連している (Coligan, J. E. 他 (1991) Current Protocols in Immunology 1 (2) の5章等を参照)。同様に、化合物は、SECPが結合する天然受容体、或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し得る。何れの場合も、既知の技術を用いてこの化合物を合理的に設計することができる。一実施例では、このような化合物に対するスクリーニングには、分泌タンパク質或いは細胞膜上のタンパク質の何れか一方として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、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を含む細胞膜断片を試験化合物と接触させて、SECPまたは化合物の何れかの結合、刺激または阻害を分析する。

40

【0152】

50

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

【0153】

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。このような化合物には、アゴニスト、アンタゴニスト、或るいは部分的または逆アゴニスト等が含まれる。一実施例では、SECPが少なくとも1つの試験化合物と結合する、SECPの活性が許容される条件下でアッセイを実施し、試験化合物の存在下でのSECPの活性が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。別法では、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。これらアッセイの何れかにおいて、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つから複数の試験化合物をスクリーニングすることができる。

10

20

【0154】

別の実施例では、胚性幹細胞（ES細胞）における相同組換えを用いて動物モデル系内で、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「ノックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの作製に有用である（米国特許第5,175,383号及び第5,767,337号等を参照）。（米国特許第5,175,383号及び第5,767,337号等を参照）。例えば129/SvJ細胞株等のマウスES細胞は初期のマウス胚に由来し、培地で増殖させることができる。このES細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（neo: C_ap_ec_ch_i, M. R. (1989) *Science* 244:1288-1292）等のマーカー遺伝子で破壊した目的の遺伝子を含むベクターで形質転換する。このベクターは、相同組換えにより宿主ゲノムの対応する領域に組み込まれる。別法では、Cre-loxP系を用いて相同組換えを行い、組織特異的または発生段階特異的に目的遺伝子をノックアウトする（Mar_th, J. D. (1996) *Clin. Invest.* 97:1999-2002; Wag_ner, K. U. 他 (1997) *Nucleic Acids Res.* 25:4323-4330）。形質転換したES細胞を同定し、例えばC57BL/6マウス系等から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を決め、これを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、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。

30

【0155】

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、中胚葉及び外胚葉の細胞の種類を含む少なくとも8つの別々の細胞系統に分化する可能性を有する。これらの細胞系統は、例えば神経細胞、造血系統及び心筋細胞に分化する（Thom_son, J. A. 他 (1998) *Science* 282:1145-1147）。

40

【0156】

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或る領域を動物ES細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細

50

胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について研究し、可能性のある医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばSEC Pを乳汁内に分泌するなどSEC P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)。

【0157】

(治療)

SEC Pのある領域と分泌蛋白のある領域との間に、例えば配列及びモチーフの文脈における化学的及び構造的類似性が存在する。さらに、SEC Pの発現は生殖器、内分泌、免疫系、胃腸、線維芽細胞、肺、脳および神経の各組織に密接に関連している。従って、SEC Pは、細胞増殖異常、自己免疫/炎症性疾患、心血管疾患、神経疾患、および発育障害においてある役割を果たすと考えられる。SEC P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、SEC Pの発現または活性を低下させることが望ましい。また、SEC P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、SEC Pの発現または活性を亢進させることが望ましい。

【0158】

従って、一実施例において、SEC P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SEC P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/炎症性の疾患、心血管疾患、神経の疾患、および発生障害が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病(MCTD)、骨髓線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髓、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、精巣、胸腺、甲状腺、及び子宮の癌が含まれ、自己免疫/炎症性の疾患の中には、炎症及び日光性角化症、後天性免疫不全症候群(AIDS)及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー(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、リンパ球毒素性一時性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、膵炎、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、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、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、心血管疾患の中には、うっ血性心不全、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、僧帽弁輪状石灰化(mitral annular calcification)、僧帽弁脱出、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、肺移植の合併症、動静脈瘻、アテローム硬化、高血圧、脈管炎、レイノー病、静脈奇形、動脈解離、静脈瘤、血栓静脈炎及び静脈血栓、血管の腫瘍、血栓崩壊の合併症、バルーン血管形成術(balloon angioplasty)、血管置換術(vascular replacement)、大動脈冠動脈バイパス術移植手術(coronary artery bypass graft surgery)が含まれ、先天性肺異常、無気肺、肺鬱血および肺水腫、肺塞栓症、肺出血、肺梗塞、肺高血圧症、血管硬化症、閉塞性肺疾患、拘束性肺疾患、慢性閉塞性肺疾患、肺気腫、慢性気管支炎、気管支喘息、気管支拡張症、細菌性肺炎、ウイルスおよびマイコプラズマ

ラズマ性肺炎、肺膿瘍、肺結核、びまん性間質性疾患、塵肺症、サルコイドーシス、特発性肺線維症、剥離性間質性肺炎、過敏性肺臓炎、肺好酸球増加細気管支炎閉塞性器質化肺炎、びまん性肺出血症候群、グッドパスチャー症候群、特発性肺ヘモシデリン沈着症、肺併発膠原性血管疾患、肺胞蛋白症、肺腫瘍、炎症性および非炎症性胸水、気胸、肋膜腫瘍、薬剤性肺疾患、放射線性肺疾患、および肺移植合併症、神経の疾患の中には、癲癇、虚血性脳血管障害、脳卒中、大脳新生物、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病及びその他の錐体外路障害、筋萎縮性側索硬化及びその他の運動ニューロン障害、進行性神経性筋萎縮症、色素性網膜炎、遺伝性運動失調、多発性硬化症及び他の脱髄疾患、細菌性及びウイルス性髄膜炎、脳膿瘍、硬膜下蓄膿症、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎及び神経根炎、ウイルス性中枢神経系疾患と、クールー及びクロイツフェルト-ヤコブ病、Gerstmann-Strausler-Scheinker症候群を含むプリオン病と、致死性家族性不眠症、神経系性栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜血管芽腫(cerebellorretinal 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及び他の発達障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経疾患、脊髄疾患筋ジストロフィー、および他の神経筋障害、末梢神経系疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性、及び中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分性及び不安性の障害、分裂病性疾患を含む精神病と、季節性障害(SAD)と、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、帯状疱疹後神経痛、及びトゥレット病と、進行性核上麻痺、大脳皮質基底核変性症及び家族性前頭側頭性健忘症および癲癇発作とが含まれ、また発達障害も含まれ、その中には尿細管性アシドーシス、貧血、クッシング症候群、軟骨形成不全性小人症、デュシェンヌ-ベッカー型筋ジストロフィー、癲癇、性腺形成異常、WAGR症候群(ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神薄弱)、スミス-マジェニス症候群(Smith-Magenis syndrome)、脊髄形成異常症候群、遺伝性粘膜上皮異形成、遺伝性角皮症、シャルコー-マリー-トゥース病及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、Sydenham舞踏病(Sydenham's chorea)及び脳性小児麻痺などの発作障害、脊髄二分裂、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感覚神経性聴力損失が含まれる。

【0159】

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SEC P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、SEC Pまたはその断片や誘導体を発現し得るベクターを患者に投与することも可能である。

【0160】

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SEC Pの発現または活性の低下に関連した疾患の治療または予防のために、実質的に精製されたSEC Pを含む組成物を好適な医薬用担体と共に患者に投与することも可能である。

【0161】

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SEC P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、SEC P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。

【0162】

更なる実施例では、SEC P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SEC Pのアンタゴニストを投与することが可能である。限定するものではないが、このような疾患の例には、上記した細胞増殖異常、自己免疫/炎症性の疾患、心血管疾患、神経の疾患、および発生障害が含まれる。一実施態様では、SEC P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SEC Pを発現する細胞または組織に薬剤を運ぶターゲティング或いは送達機構として間接的に用いられ得る。

【0163】

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 S E C P 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、S E C P 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。

【0164】

別の実施例では、本発明の任意のタンパク質、アンタゴニスト、抗体、アゴニスト、相補配列、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、当業者が従来の医薬原理に従ってを選択し得る。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。この方法を用いることにより少量の各薬剤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

【0165】

S E C P 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 S E C P 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 S E C P と特異的に結合するものの同定が可能である。S E C P の抗体も、当分野で一般的な方法を用いて製造することが可能である。このような抗体には、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、一本鎖、F a b 断片、及び F a b 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（即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。

【0166】

抗体の産生のためには、ヤギ、ウサギ、ラット、マウス、ヒト及びその他のものを含む種々の宿主が、S E C P 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等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、キーホールリンペットヘモシニアン、ジニトロフェノール等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、B C G（カルメット-ゲラン杆菌）及びコリネバクテリウム-パルヴムが特に好ましい。

20

【0167】

S E C P に対する抗体を誘発するために用いられるオリゴペプチド、ペプチド、または断片は、少なくとも約 5 個のアミノ酸からなり、一般的には約 10 個以上のアミノ酸からなるものが好ましい。これらのオリゴペプチド、ペプチドまたは断片は、天然のタンパク質のアミノ酸配列の一部と同一であり且つ小さな天然の分子の全アミノ酸配列を含むことが望ましい。S E C P アミノ酸の短いストレッチは、K L H などの別のタンパク質の配列と融合し、キメラ分子に対する抗体が産生され得る。

30

【0168】

S E C P 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 E B V-ハイブリドーマ技術がある（K o h l e r, G. ら. (1975) Nature 256: 495-497、K o z b o r, D. ら. (1985) J. Immunol. Methods 81: 31-42、C o t e, R. J. ら. (1983) Proc. Natl. Acad. Sci. USA 80: 2026-2030、C o l e, S. P. ら. (1984) Mol. Cell Biol. 62: 109-120等を参照）。

40

【0169】

更に、「キメラ抗体」作製のために発達したヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの技術が、好適な抗原特異性及び生物学的活性を備える分子を得るために用いられる（例えば、M o r r i s o n, S. L. 他. (1984) Proc. Natl. Acad. Sci. 81-4851-4855; N e u b e r g e r, M. S. 他. (1984) Nature 312: 604-608; T a k e d a, S

50

ら、(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特異性一本鎖抗体を生成する。関連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)。

【0170】

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、或いは免疫グロブリンライブラリのスクリーニングまたは文献に開示されているような高特異結合試薬のパネルのスクリーニングによっても行い得る(Orlandi, R. ら (1989) *Proc. Natl. Acad. Sci. USA* 86: 3833-3837、Winter, G. ら (1991) *Nature* 349: 293-299等を参照)。

【0171】

SECPに対する特異的な結合部位を含む抗体断片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、モノクローナル Fab 断片を所望の特異性と迅速且つ容易に同定することが可能となる(Huse, W. D. ら (1989) *Science* 256: 1275-1281等を参照)。

【0172】

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。隔離された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる競合的な結合、または免疫放射線活性のための数々のプロトコルが、当分野では周知である。通常このようなイムノアッセイには、SECPとその特異性抗体との間の複合体形成の計測が含まれる。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、競合的な結合アッセイも利用することができる(Pound、前出)。

【0173】

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

【0174】

ポリクローナル抗体試薬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医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 $1 \sim 2 \text{ mg/ml}$ の特異的な抗体、好ましくは $5 \sim 10 \text{ mg/ml}$ 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C a t t yの文献、同C o l i g a nらの文献等を参照）。

【0175】

本発明の別の実施例では、S E C Pをコードするポリヌクレオチド、またはその任意の断片や相補配列が、治療目的で使うことができる。ある実施態様では、S E C Pをコードする遺伝子のコーディング領域や調節領域に相補的な配列やアンチセンス分子（DNA及びRNA、修飾ヌクレオチド）を設計して遺伝子発現を変更することができる。このような技術は当分野では周知であり、センスまたはアンチセンスオリゴヌクレオチドまたは大きな断片が、S E C Pをコードする配列の制御領域から、またはコード領域に沿ったさまざまな位置から設計可能である。（A g r a w a l, S., ed. (1996) A n t i s e n s e T h e r a p e u t i c s, H u m a n a P r e s s I n c., T o t a w a N J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発現する発現プラスミドの形で細胞内に輸送することが可能である（S l a t e r, J. E. 他 (1998) J. A l l e r g y C l i. I m m u n o l. 102 (3): 469-475; およびS c a n l o n, K. J. 他 (1995) 9 (13): 1288-1296を参照。）アンチセンス配列はまた、例えばレトロウイルスやアデノ関連ウイルスベクター等のウイルスベクターを用いて細胞内に導入することもできる（M i l l e r, A. D. (1990) B l o o d 76: 271、前出のA u s u b e l, U c k e r t, W. および W. W a l t h e r (1994) P h a r m a c o l. T h e r. 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ップ及び当分野で公知のその他の系が含まれる（R o s s i, J. J. (1995) B r. M e d. B u l l. 51 (1): 217-225; B o a d o, R. J. ら (1998) J. P h a r m. S c i. 87 (11): 1308-1315、M o r r i s, M. C. ら (1997) N u c l e i c A c i d s R e s. 25 (14): 2730-2736. 等を参照）。

【0176】

本発明の別の実施例では、S E C P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（C a v a z z a n a-C a l v o, M. 他 (2000) S c i e n c e 288: 669-672）により特徴付けられる重度の複合型免疫欠損症（S C I D）-X1の場合）、先天性アデノシンデアミナーゼ（A D A）欠損症に関連する重度の複合型免疫欠損症（B l a e s e, R. M. 他 (1995) S c i e n c e 270: 475-480、B o r d i g n o n, C. 他 (1995) S c i e n c e 270: 470-475）、嚢胞性繊維症（Z a b n e r, J. 他 (1993) C e l l 75: 207-216; C r y s t a l, R. G. 他 (1995) H u m. G e n e T h e r a p y 6: 643-666、C r y s t a l, R. G. 他 (1995) H u m. G e n e T h e r a p y 6: 667-703）、サラセミア（t h a l a s s a m i a）、家族性高コレステロール血症、第V I I I 因子若しくは第I X 因子欠損に起因する血友病（C r y s t a l, R. G. (1995) S c i e n c e 270: 404-410、V e r m a, I. M. および S o m i a, N. (1997) N a t u r e 389: 239-242）を治療し、(i i) 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）、(i i i) 細胞内の寄生生物（例えばヒト免疫不全ウイルス（H I V）（B a l t i m o r e, D. (1988) N a t u r e 335: 395-396、P o e s c h l a, E. 他 (1996) P r o c. N a t l. A c a d. S c i. U S A. 93: 11395-11399）、B型若しくはC型肝炎ウイルス（H B V、H C V）、C a n d i d

10

20

30

40

50

a a l b i c a n s及びP a r a c o c c i d i o i d e s b r a s i l i e n s i s等の真菌寄生菌、並びに熱帯熱マラリア原虫及びクルーズトリパノソーマ等の原虫寄生生物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。

【0177】

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。i n v i v o或いはe x v i t r o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。

【0178】

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ベクター(Invitrogen, Carlsbad CA)、PCMV-SCRIPT、PCMV-TAG、PEGSH/PERV(Stratagene, La Jolla CA)、PTE-T-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、チミジンキナーゼ(TK)、若しくはβ-アクトチン遺伝子等)、(ii)誘導性プロモーター(例えば、市販されているT-RXプラスミド(Invitrogen)に含まれている、テトラサイクリン調節性プロモーター(Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. U. S. A. 89:5547-5551; Gossen, M. 他 (1995) Science 268:1766-1769; Rossi, F. M. V. and H. M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456))、エクジソン誘導性プロモーター(市販されているプラスミドPVGRXR及びPINDに含まれている:Invitrogen)、FK506/ラパマイシン誘導性プロモーター、またはRU486/ミフェプリストン誘導性プロモーター(Rossi, F. M. V. and H. M. Blau, 前出)、または(iii)正常な個体に由来するSECPをコードする内在性遺伝子の天然のプロモーター若しくは組織特異的プロモーターを用いることが可能である。

【0179】

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、当業者は経験にそれほど頼らないでもポリヌクレオチドを培養中の標的細胞に導入することが可能になる。別法では、リン酸カルシウム法(Graham, F. L. および A. J. Eb (1973) Virology 52:456-467)若しくは電気穿孔法(Neumann, B. ら (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、標準化された哺乳動物の形質移入プロトコルの修飾が必要である。

【0180】

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PFBNEO)はStratagene社から市販されており、刊行データ(Riviere, I. ら, (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系(VPCL)において増殖され、VPCLは、標的細胞上の受容体に対する向性を有するエンベロープ遺伝子またはVSVg等の乱交雑エンベロープタンパク質を発現する(Armentano, D. ら (1987) J. Virol. 61:1647-1650、Bender, M. A. ら (1987) J. Virol. 61:1639-1646、Adam, M. A. および A. D. Miller (1988) J. Virol. 62:3802-3806、Dull, T. ら (1998) J. Virol. 72:8463-8471、Zufferey, R. ら (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系を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、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. ら (1997) J. Virol. 71:7020-7029、Bauer, G. ら (1997) Blood 89:2259-2267、Bonyhadi, M. L. (1997) J. Virol. 71:4707-4716、Ranga, U. ら (1998) Proc. Natl. Acad. Sci. U. S. A. 95:1201-1206、Su, L. (1997) Blood 89:2283-2290)。

【0181】

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或いは複数の遺伝子異常を有する細胞にSECPをコードするポリヌクレオチドを送達する。アデノウイルス系ベクター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脾臓の無損傷の脾島内に導入するために可変性であることが証明された(Csete, M. E. ら (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、Antinuzzi, P. A. ら (1999) Annu. Rev. Nutr. 19:511-544 及び Verma, I. M. および 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。

【0182】

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或いは複数の遺伝子異常を有する標的細胞にSECPをコードするポリヌクレオチドを送達する。単純疱疹ウイルス(HSV)系のベクターは、HSV親和性の中枢神経細胞にSECPを導入する際に特に重要である。ヘルペス系ベクターの作製及びパッケージングは、当業者に公知である。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. ら (1

999) Exp. Eye Res. 169:385-395)。HSV-1 ウイルスベクターの作製についても、De Luca 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 つの外在性遺伝子を有するゲノムを含む組換え HSV d92 についての記載がある。上記特許はまた、ICP4、ICP27 及び ICP22 のために除去される組換え 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. ら (1999) J. Virol. 73:519-532 及び Xu, H. ら (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部分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。

【0183】

別法では、 α ウイルス (正の一本鎖 RNA ウイルス) ベクターを用いて SEC P 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 α 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が SFV ゲノムに基づいていることが分かった (Garoff, H. および K.-J. Li (1998) Curr. Opin. Biotech. 9:464-469)。 α ウイルス RNA の複製中に、通常はウイルスのキャプシッ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キャプシッドタンパク質が過剰産生される。同様に、SEC P をコードする配列を α ウイルスゲノムのカプシドをコードする領域に導入することによって、ベクター導入細胞において多数の SEC P をコードする RNA が産生され、高いレベルで SEC P が合成される。通常は α ウイルスの感染が数日以内での細胞溶解に関係する一方で、シンドビスウイルス (SIN) の変異体を有するハムスター正常腎臓細胞 (BHK-21) の持続的な感染を確立する能力は、 α ウイルスの溶解複製を遺伝子治療に適用できるように好適に変更可能であることを示唆している (Dryga, S. A. ら (1997) Virology 228:74-83)。様々な宿主に α ウイルスを導入できることから、様々なタイプの細胞に SEC P を導入することできる。或る集団におけるサブセットの細胞の特定形質導入は、形質導入前に細胞の選別を必要とし得る。 α ウイルスの感染性 cDNA クローンの処置方法、 α ウイルスの cDNA 及び RNA の形質移入方法及び α ウイルスの感染方法は、当業者に公知である。

【0184】

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。転写開始部位とは例えば開始部位から数えて約 -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ような二重らせんの能力を阻害するので、三重らせん塩基対形成は有用である。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. ら (1994) in: Huber, B. E. および B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY, 163-177 ページ等を参照)。相補配列またはアンチセンス分子もまた、転写物がリボソームに結合するのを阻止することによって mRNA の翻訳を阻止するように設計することができる。

【0185】

リボザイムは酵素性RNA分子であり、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、ヌクレオチド鎖切断に先立つ相補的標的RNAへのリボザイム分子の配列特異性ハイブリダイゼーションに参与している。例えば、人工的に作製されたハンマーヘッド型リボザイム分子が、SECPをコードする配列の内ヌクレオチド鎖分解性の切断を特異的且つ効果的に触媒できる可能性がある。

【0186】

任意のRNA標的内の特異的リボザイム切断部位を、GUA、GUU、GUC配列を含めたりリボザイム切断部位に対する標的分子をスキャンすることによって先ず同定する。一度同定すると、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次構造の特徴であること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とのハイブリダイゼーションのアクセス可能性をテストすることによって行うことができる。

10

【0187】

本発明の相補リボ核酸分子及びリボザイムは、当分野でよく知られている任意の核酸分子の合成方法を用いて作製することができる。任意の方法には、固相フォスフォアミダイト化学合成のようなオリゴヌクレオチドを化学的に合成する方法がある。或いは、SECPをコードするDNA配列の*in vitro*及び*in vivo*転写によってRNA分子を作成し得る。このようなDNA配列は、T7やSP6等の好適な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、細胞系、細胞または組織内に導入することができる。

20

【0188】

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には、分子の5'末端、3'末端、あるいはその両方においてフランキング配列を追加したり、分子の主鎖内においてホスホジエステラーゼ結合ではなくホスホロチオネートまたは2'-O-メチルを使用したり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り、これら全ての分子に拡大することができる。それには、内在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、チオール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)、ワイブトシン(wybutosine)等を加えることでできる。

30

【0189】

本発明の更なる実施例は、SECPをコードするポリヌクレオチドの発現の変化に有効な化合物をスクリーニングする方法を含む。限定するものではないが特異ポリヌクレオチドの発現変化を起こすのに有効な化合物には、オリゴヌクレオチド、アンチセンスオリゴヌクレオチド、三重らせん形成オリゴヌクレオチド、転写因子その他のポリペプチド転写制御因子、及び特異ポリヌクレオチド配列と相互作用し得る非高分子化学的実体がある。有効な化合物は、ポリヌクレオチド発現のインヒビターまたはエンハンサーのいずれかとして作用することによりポリヌクレオチド発現を改変し得る。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効であり、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効であり得る。

40

【0190】

特異ポリヌクレオチドの変異発現における有効性に対して、少なくとも1個から複数個の試験化合物をスクリーニングし得る。試験化合物は、当分野で通常知られている任意の方法により得られる。このような方法には、ポリヌクレオチドの発現を変異させる場合と、既存の、市販のまたは専売の、天然または非天然の化合物ライブラリから選択する場合と、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

50

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。S E C Pをコードするポリヌクレオチドを含むサンプルは、得られた少なくとも1つの試験化合物に曝露する。サンプルには例えば、無傷細胞、透過化処理した細胞、試験管内無細胞再構成系または再構成生化学系があり得る。S E C Pをコードするポリヌクレオチドの発現の改変は、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、S E C P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によって、ポリヌクレオチドの発現を変える際に試験化合物が有効であることが示される。特異ポリヌクレオチドの変異発現に有効な化合物に対して、例えば Schizosaccharomyces pombe 遺伝子発現系 (Atkins, D. ら (1999) 米国特許第5, 932, 435号、Arndt, G. M. ら (2000) Nucleic Acids Res. 28: E15) またはHeLa細胞等のヒト細胞系 (Clarke, M. L. ら (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の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 10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310
320
330
340
350
360
370
380
390
400
410
420
430
440
450
460
470
480
490
500
510
520
530
540
550
560
570
580
590
600
610
620
630
640
650
660
670
680
690
700
710
720
730
740
750
760
770
780
790
800
810
820
830
840
850
860
870
880
890
900
910
920
930
940
950
960
970
980
990
1000
1010
1020
1030
1040
1050
1060
1070
1080
1090
1100
1110
1120
1130
1140
1150
1160
1170
1180
1190
1200
1210
1220
1230
1240
1250
1260
1270
1280
1290
1300
1310
1320
1330
1340
1350
1360
1370
1380
1390
1400
1410
1420
1430
1440
1450
1460
1470
1480
1490
1500
1510
1520
1530
1540
1550
1560
1570
1580
1590
1600
1610
1620
1630
1640
1650
1660
1670
1680
1690
1700
1710
1720
1730
1740
1750
1760
1770
1780
1790
1800
1810
1820
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
2020
2030
2040
2050
2060
2070
2080
2090
2100
2110
2120
2130
2140
2150
2160
2170
2180
2190
2200
2210
2220
2230
2240
2250
2260
2270
2280
2290
2300
2310
2320
2330
2340
2350
2360
2370
2380
2390
2400
2410
2420
2430
2440
2450
2460
2470
2480
2490
2500
2510
2520
2530
2540
2550
2560
2570
2580
2590
2600
2610
2620
2630
2640
2650
2660
2670
2680
2690
2700
2710
2720
2730
2740
2750
2760
2770
2780
2790
2800
2810
2820
2830
2840
2850
2860
2870
2880
2890
2900
2910
2920
2930
2940
2950
2960
2970
2980
2990
3000
3010
3020
3030
3040
3050
3060
3070
3080
3090
3100
3110
3120
3130
3140
3150
3160
3170
3180
3190
3200
3210
3220
3230
3240
3250
3260
3270
3280
3290
3300
3310
3320
3330
3340
3350
3360
3370
3380
3390
3400
3410
3420
3430
3440
3450
3460
3470
3480
3490
3500
3510
3520
3530
3540
3550
3560
3570
3580
3590
3600
3610
3620
3630
3640
3650
3660
3670
3680
3690
3700
3710
3720
3730
3740
3750
3760
3770
3780
3790
3800
3810
3820
3830
3840
3850
3860
3870
3880
3890
3900
3910
3920
3930
3940
3950
3960
3970
3980
3990
4000
4010
4020
4030
4040
4050
4060
4070
4080
4090
4100
4110
4120
4130
4140
4150
4160
4170
4180
4190
4200
4210
4220
4230
4240
4250
4260
4270
4280
4290
4300
4310
4320
4330
4340
4350
4360
4370
4380
4390
4400
4410
4420
4430
4440
4450
4460
4470
4480
4490
4500
4510
4520
4530
4540
4550
4560
4570
4580
4590
4600
4610
4620
4630
4640
4650
4660
4670
4680
4690
4700
4710
4720
4730
4740
4750
4760
4770
4780
4790
4800
4810
4820
4830
4840
4850
4860
4870
4880
4890
4900
4910
4920
4930
4940
4950
4960
4970
4980
4990
5000
5010
5020
5030
5040
5050
5060
5070
5080
5090
5100
5110
5120
5130
5140
5150
5160
5170
5180
5190
5200
5210
5220
5230
5240
5250
5260
5270
5280
5290
5300
5310
5320
5330
5340
5350
5360
5370
5380
5390
5400
5410
5420
5430
5440
5450
5460
5470
5480
5490
5500
5510
5520
5530
5540
5550
5560
5570
5580
5590
5600
5610
5620
5630
5640
5650
5660
5670
5680
5690
5700
5710
5720
5730
5740
5750
5760
5770
5780
5790
5800
5810
5820
5830
5840
5850
5860
5870
5880
5890
5900
5910
5920
5930
5940
5950
5960
5970
5980
5990
6000
6010
6020
6030
6040
6050
6060
6070
6080
6090
6100
6110
6120
6130
6140
6150
6160
6170
6180
6190
6200
6210
6220
6230
6240
6250
6260
6270
6280
6290
6300
6310
6320
6330
6340
6350
6360
6370
6380
6390
6400
6410
6420
6430
6440
6450
6460
6470
6480
6490
6500
6510
6520
6530
6540
6550
6560
6570
6580
6590
6600
6610
6620
6630
6640
6650
6660
6670
6680
6690
6700
6710
6720
6730
6740
6750
6760
6770
6780
6790
6800
6810
6820
6830
6840
6850
6860
6870
6880
6890
6900
6910
6920
6930
6940
6950
6960
6970
6980
6990
7000
7010
7020
7030
7040
7050
7060
7070
7080
7090
7100
7110
7120
7130
7140
7150
7160
7170
7180
7190
7200
7210
7220
7230
7240
7250
7260
7270
7280
7290
7300
7310
7320
7330
7340
7350
7360
7370
7380
7390
7400
7410
7420
7430
7440
7450
7460
7470
7480
7490
7500
7510
7520
7530
7540
7550
7560
7570
7580
7590
7600
7610
7620
7630
7640
7650
7660
7670
7680
7690
7700
7710
7720
7730
7740
7750
7760
7770
7780
7790
7800
7810
7820
7830
7840
7850
7860
7870
7880
7890
7900
7910
7920
7930
7940
7950
7960
7970
7980
7990
8000
8010
8020
8030
8040
8050
8060
8070
8080
8090
8100
8110
8120
8130
8140
8150
8160
8170
8180
8190
8200
8210
8220
8230
8240
8250
8260
8270
8280
8290
8300
8310
8320
8330
8340
8350
8360
8370
8380
8390
8400
8410
8420
8430
8440
8450
8460
8470
8480
8490
8500
8510
8520
8530
8540
8550
8560
8570
8580
8590
8600
8610
8620
8630
8640
8650
8660
8670
8680
8690
8700
8710
8720
8730
8740
8750
8760
8770
8780
8790
8800
8810
8820
8830
8840
8850
8860
8870
8880
8890
8900
8910
8920
8930
8940
8950
8960
8970
8980
8990
9000
9010
9020
9030
9040
9050
9060
9070
9080
9090
9100
9110
9120
9130
9140
9150
9160
9170
9180
9190
9200
9210
9220
9230
9240
9250
9260
9270
9280
9290
9300
9310
9320
9330
9340
9350
9360
9370
9380
9390
9400
9410
9420
9430
9440
9450
9460
9470
9480
9490
9500
9510
9520
9530
9540
9550
9560
9570
9580
9590
9600
9610
9620
9630
9640
9650
9660
9670
9680
9690
9700
9710
9720
9730
9740
9750
9760
9770
9780
9790
9800
9810
9820
9830
9840
9850
9860
9870
9880
9890
9900
9910
9920
9930
9940
9950
9960
9970
9980
9990
10000

【0191】

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一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、リボソーム注入またはポリカチオンアミノポリマーによる送達は、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. ら (1997) Nat. Biotechnol. 15: 462-466. 等を参照)。

【0192】

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このような治療が必要な全ての対象に適用できる。

【0193】

本発明の追加実施例は、通常薬剤として許容できる賦形剤で処方される活性成分を有する成分の投与に関連する。賦形剤には例えば、糖、でんぷん、セルロース、ゴム及びタンパク質がある。様々な処方通常知られており、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、S E C P、S E C Pの抗体、擬態、アゴニスト、アンタゴニスト、またはS E C Pのインヒビターなどからなる。

【0194】

本発明に用いられる成分は、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、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。

【0195】

肺から投与する成分は、液状または乾燥粉末状で調製し得る。このような成分は通常、患者が吸入する直前にエアロゾル化する。小分子 (例えば伝統的な低分子量有機薬) の場合には、速効製剤のエアロゾル送達は当分野で公知である。高分子 (例えばより大きなペ

10

20

30

40

50

プチド及びタンパク質)の場合には、当該分野において肺の肺胞領域を介しての肺送達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. ら, 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。

【0196】

本発明での使用に適した成分には、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成分が含まれる。有効投与量の決定は、当業者の能力の範囲内で行う。

【0197】

SECPまたはその断片を含む高分子を直接細胞内に輸送するべく、特殊な形態に組成物が調製されるのが好ましい。例えば、細胞不透過性高分子を含むリポソーム製剤は、細胞融合及び高分子の細胞内送達を促進し得る。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は、マウスモデル系の脳を含む全ての組織の細胞に形質導入することがわかっている(Schwarze, S. R. ら (1999) Science 285:1569-1572)。

10

【0198】

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、先ず治療の有効投与量を推定することができる。動物モデルはまた、好適な濃度範囲及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及び投与経路を決定することができる。

20

【0199】

医学的に効果的な薬用量は、症状や容態を回復させる、たとえば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、インヒビターなどの活性処方成分の量に関連する。治療有効度及び毒性は、細胞培養または動物実験における標準的な薬剤手法によって、例えばED₅₀ (集団の50%の医薬的有效量)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。毒性効果の薬用効果に対する投与量の比は、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成分が望ましい。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調剤するのに用いられる。このような組成物が含まれる投与量は、毒性を殆ど或いは全く含まず、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、患者の感受性及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

【0200】

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。効果的なレベルの活性成分を与え、或いは所望の効果を維持するべく、投与量及び投与を調節する。被験者に関する要素としては、疾患の重症度、患者の通常の状態、患者の年齢、体重及び性別、投与の時間及び頻度、薬剤の配合、反応感受性及び治療に対する応答等を考慮する。長時間作用の成分は、特定の製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。

40

【0201】

通常の投与量は、投与の経路にもよるが約0.1~100,000µgであり、合計で約1gまでとする。特定の投与量及び送達方法に関するガイダンスは文献に記載されており、現場の医者は通常それを利用することができる。当業者は、タンパク質またはインヒビターに対する処方とは異なる、ヌクレオチドに対する剤形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、状態、位置等に特異的なものとなる。

【0202】

(診断)

50

別の実施例では、S E C Pに特異的に結合する抗体が、S E C Pの発現によって特徴付けられる疾患の診断、またはS E C PやS E C Pのアゴニストまたはアンタゴニスト、インヒビターで治療を受けている患者をモニターするためのアッセイに用いられる。診断目的に有用な抗体は、上記の治療の箇所で記載した方法と同じ方法で調合される。S E C P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S E C P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。

【0203】

10

S E C Pを測定するためのE L I S A, R I A, 及びF A C S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S E C P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S E C P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S E C Pに対する抗体とを結合させることによって決定する。標準複合体形成量は、種々の方法、例えば測光法で定量できる。被験者のS E C Pの発現の量、対照及び疾患、生検組織からのサンプルが基準値と比較される。基準値と被験者との偏差が疾患を診断するパラメータとなる。

【0204】

本発明の別の実施例によれば、S E C Pをコードするポリヌクレオチドを診断のために用いることもできる。用いることができるポリヌクレオチドには、オリゴヌクレオチド配列、相補的R N A及びD N A分子、そしてP N Aが含まれる。このポリヌクレオチドを用いて、疾患と関連し得るS E C P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、S E C Pの存在の有無、更に過剰な発現を調べ、治療中のS E C P値の調節を監視する。

20

【0205】

一実施形態では、S E C P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P C Rプローブを用いたハイブリダイゼーションによって、S E C Pをコードする核酸配列を同定することが可能である。例えば5'調節領域である高度に特異的な領域か、例えば保存されたモチーフであるやや特異性の低い領域から作られているかのプローブの特異性と、ハイブリダイゼーション或いは増幅のストリンジェントは、プローブがS E C P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コードする自然界の配列のみを同定するかどうかによって決まるであろう。

30

【0206】

プローブはまた、関連する配列の検出に利用され、S E C P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、D N AあるいはR N Aが可能であり、S E Q I D N O: 45-88の配列、或いはS E C P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。

【0207】

40

S E C PをコードするD N Aに対して特異的なハイブリダイゼーションプローブの作製方法には、S E C P及びS E C P誘導体をコードするポリヌクレオチド配列をm R N Aプローブの作製のためのベクターにクローニングする方法がある。m R N A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R N Aポリメラーゼ及び好適な標識されたヌクレオチドを加えることによって、i n v i t r oでR N Aプローブを合成するために用いられ得る。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。

【0208】

50

S E C Pをコードするポリヌクレオチド配列を用いて、S E C P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／炎症性の疾患、心血管疾患、神経の疾患、および発生障害が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病（M C T D）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、精巣、胸腺、甲状腺、及び子宮の癌が含まれ、自己免疫／炎症性の疾患の中には、炎症及び日光性角化症、後天性免疫不全症候群（A I D S）及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー（A P E C E D）、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、膵炎、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、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、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、心血管疾患の中には、うっ血性心不全、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、僧帽弁輪状石灰化（m i t r a l a n n u l l a r c a l c i f i c a t i o n）、僧帽弁脱出、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、肺移植の合併症、動静脈瘻、アテローム硬化、高血圧、脈管炎、レイノー病、静脈奇形、動脈解離、静脈瘤、血栓静脈炎及び静脈血栓、血管の腫瘍、血栓崩壊の合併症、バルーン血管形成術（b a l l o o n a n g i o p l a s t y）、血管置換術（v a s c u l a r r e p l a c e m e n t）、大動脈冠動脈バイパス術移植手術（c o r o n a r y a r t r y b y p a s s g r a f t s u e g e r y）が含まれ、先天性肺異常、無気肺、肺鬱血および肺水腫、肺塞栓症、肺出血、肺梗塞、肺高血圧症、血管硬化症、閉塞性肺疾患、拘束性肺疾患、慢性閉塞性肺疾患、肺気腫、慢性気管支炎、気管支喘息、気管支拡張症、細菌性肺炎、ウイルスおよびマイコプラズマ性肺炎、肺膿瘍、肺結核、びまん性間質性疾患、塵肺症、サルコイドーシス、特発性肺線維症、剥離性間質性肺炎、過敏性肺臓炎、肺好酸球増加細気管支炎閉塞性器質化肺炎、びまん性肺出血症候群、グッドパスチャー症候群、特発性肺ヘモシデリン沈着症、肺併発膠原性血管疾患、肺胞蛋白症、肺腫瘍、炎症性および非炎症性胸水、気胸、肋膜腫瘍、薬剤性肺疾患、放射線性肺疾患、および肺移植合併症、神経の疾患の中には、癲癇、虚血性脳血管障害、脳卒中、大脳新生物、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病及びその他の錐体外路障害、筋萎縮性側策硬化及びその他の運動ニューロン障害、進行性神経性筋萎縮症、色素性網膜炎、遺伝性運動失調、多発性硬化症及び他の脱髄疾患、細菌性及びウイルス性髄膜炎、脳膿瘍、硬膜下蓄膿症、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎及び神経根炎、ウイルス性中枢神経系疾患と、クールー及びクロイツフェルト-ヤコブ病、G e r s t m a n n - S t r a u s s l e r - S c h e i n k e r症候群を含むプリオン病と、致死性家族性不眠症、神経系性栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜血管芽腫（c e r e b e l l o r e t i n a l h e m a n g i o b l a s t o m a t o s i s）、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及び他の発達障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経疾患、脊髄疾患筋ジストロフィー、および他の神経筋障害、末梢神経系疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性、及び

中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分性及び不安性の障害、分裂病性疾患を含む精神病と、季節性障害（SAD）と、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、帯状疱疹後神経痛、及びトゥレット病と、進行性核上麻痺、大脳皮質基底核変性症及び家族性前頭側頭性健忘症および癲癇発作とが含まれ、また発達障害も含まれ、その中には尿細管性アシドーシス、貧血、クッシング症候群、軟骨形成不全性小人症、デュシェンヌ・ベッカー型筋ジストロフィー、癲癇、性腺形成異常、WAGR症候群（ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神薄弱）、スミス・マジェニス症候群（Smith-Magenis syndrome）、脊髄形成異常症候群、遺伝性粘膜上皮異形成、遺伝性角皮症、シャルコー・マリー・トゥース病及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、Syndenham舞蹈病（Syndenham's chorea）及び脳性小児麻痺などの発作障害、脊髄二分裂、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感覚神経性聴力損失が含まれる。SECPをコードするポリヌクレオチド配列は、サザン法やノーザン法、ドットブロット法、或いはその他の膜系の技術、PCR法、ディップスティック（dipstick）、ピン（pin）、ELISA式アッセイ、及び変異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するマイクロアレイに使用することが可能である。このような定性方法または定量方法は、当分野で公知である。

10

【0209】

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推定するため、或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。

20

【0210】

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。既知量の実質的に精製されたポリヌクレオチドを用いて行った実験から得た値を正常な対象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。

30

【0211】

疾患の存在が確定されて治療プロトコルが開始されると、患者の発現レベルが正常な被検者に観察されるレベルに近づき始めたかどうかを測定するため、ハイブリダイゼーションアッセイを通常ベースで繰り返すことができる。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。

40

【0212】

癌に関しては、個体からの生体組織における異常な量の転写物（過少発現または過剰発現）の存在は、疾患の発生素質を示したり、実際に臨床的症状が現れる前に疾患を検出する方法を提供したりし得る。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。

【0213】

50

S E C Pをコードする配列から設計されたオリゴヌクレオチドのさらなる診断への利用には、P C Rの利用が含まれ得る。こ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、或いは*i n v i t r o*で産出し得る。オリゴマーは、好ましくはS E C Pをコードするポリヌクレオチドの断片、或いはS E C Pをコードするポリヌクレオチドと相補的なポリヌクレオチドの断片を含み、最適な条件下で、特定の遺伝子や条件を識別するために利用される。また、オリゴマーは、やや緩いストリンジェント条件下で、近縁のD N A或いはR N A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。

【0214】

或る実施態様において、S E C Pをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いて、一塩基多型 (S N P) を検出し得る。S N Pは、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S N Pの検出方法には、S S C P (s i n g l e - s t r a n d e d c o n f o r m a t i o n p o l y m o r p h i s m) 及び蛍光S S C P (f S S C P) 法がある。S S C Pでは、S E C Pをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いたポリメラーゼ連鎖反応 (P C R) でD N Aを増幅する。D N A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。D N A内のS N Pは、一本鎖形状のP C R生成物の2次及び3次構造に差異を生じさせる。差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。f S S C P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D N Aシーケンシング機などの高処理機器でアンプリマー (a m p l i m e r) の検出が可能になる。更に、インシリコS N P (i n s i l i c o S N P, i s S N P) と呼ばれる配列データベース分析法は、一般的なコンセンサス配列に配列されるような個々の重畳するD N A断片の配列を比較することにより、多形性を同定し得る。これらのコンピュータベースの方法は、D N Aの実験室での調整及び統計モデル及びD N A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M A S S A R R A Yシステム (S e q u e n o m, I n c., S a n D i e g o C A) を用いた質量分析によりS N Pを検出し、特徴付ける。

【0215】

S E C Pの発現を定量するために用い得る方法には、ヌクレオチドの放射標識またはビオチン標識、調節核酸の相互増幅 (c o a m p l i f i c a t i o n) 及び標準曲線から得た結果の補間もある (例えば、M e l b y, P. C. 他 (1993) J. I m m u n o l. M e t h o d s 159:235244; D u p l a a, C. 他 (1993) A n a l. B i o c h e m. 212:229236を参照。) 目的のオリゴマーまたはポリヌクレオチドが種々の希釈液中に存在し、分光光度法または比色反応によって定量が迅速になるような高処理フォーマットのアッセイを行うことによって、複数のサンプルの定量速度を加速することができる。

【0216】

更に別の実施例では、本明細書で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、突然変異及び多形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剤の活性を開発及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、有効的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。

【0217】

別の実施例では、S E C P、S E C Pの断片、S E C P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、薬剤－標的相互作用及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。

【0218】

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り遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所定の条件下で所定の時間に発現した遺伝子の数及び相対存在量を定量することにより分析される（Sei l l i a m e r 他、米国特許第5, 840, 484号の”C o m p a r a t i v e G e n e T r a n s c r i p t A n a l y s i s”を参照。この特許に言及することを以って本明細書の一部とする）。従って、特定の組織または細胞タイプの転写または逆転写全体に本発明のポリヌクレオチドまたはその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。

【0219】

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。転写イメージは従って、組織または生検サンプルの場合には i n v i v o、または株化細胞の場合には i n v i t r oでの遺伝子発現を反映する。

【0220】

本発明のポリヌクレオチドの発現のプロフィールを作製する転写イメージはまた、工業的または天然の環境化合物の毒性試験のみならず、i n v i t r oモデル系及び薬剤の前臨床評価に関連して使用し得る。全ての化合物は、作用及び毒性のメカニズムを暗示する、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と称されるような特徴的な遺伝子発現パターンを惹起する（Nuway sir, E. F. ら、（1999）Mol. C a r c i n o g. 24:153-159、Steiner, S. および N. L. A n d e r s o n （2000）T o x i c o l. L e t t. 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一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、発現のゲノム全域にわたる測定が最高品質のシグネチャを提供する。自己の発現が任意の試験された化合物により変化しない遺伝子が同様に重要であっても、このような遺伝子の発現レベルを用いて残りの発現データを規準化する。規準化手順は、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が毒性メカニズムの解釈に役立つが、毒性の予測につながるシグネチャの統計的に一致させるのに遺伝子機能の知識は必要とされない（例えば2000年2月29日に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、中毒学的スクリーニングの際に毒性シグネチャを用いて、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。

【0221】

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つ若しくは複数のプローブでハイブリダイズし、それによって本発明のポリヌクレオチドに対応する転写レベルを定量し得る。処理した生物学的サン

プル中の転写レベルを、非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。

【0222】

別の実施例は、本発明のポリペプチド配列を用いて組織または細胞タイプのプロテオームを分析することに関連する。プロテオームの語は、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、個々に更に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即ちプロフィールは、所与の条件下で所与の時間に発現したタンパク質の数及び相対存在量を定量することにより分析し得る。従って細胞のプロテオームのプロフィールは、特定の組織または細胞タイプのポリペプチドを分離及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、通常クーマシーブルーまたはシルバーまたは蛍光染色などの物質を用いてゲルで染色することにより、分散した、独特な位置にある点としてゲル中で可視化される。各タンパク質スポットの光学密度は、通常サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処理または非処理のいずれかの生物学的サンプルから得られ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。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が得られる。

10

20

【0223】

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要素へのタンパク質結合レベルを検出することによりタンパク質発現レベルを定量する（Lueking, A. ら. (1999) Anal. Biochem. 270:103-111, Mendoz, L. G. ら. (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を用いて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。

30

【0224】

プロテオームレベルでの毒性シグネチャ（特徴）も中毒学的スクリーニングに有用であり、転写レベルでの毒性シグネチャと並行に分析するべきである。或る組織の或るタンパク質に対しては、転写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. および J. Seilhamer (1997) Electrophoresis 18:533-537）、転写イメージにはそれ程影響しないがタンパク質のプロフィールを変化させるような化合物の分析においてプロテオーム毒性サインは有用たり得る。更に、体液中での転写の分析はmRNA急速な分解により困難であるので、タンパク質のプロフィール作成はこのような場合により信頼でき、情報価値がある。

40

【0225】

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。処理された生物学的サンプル中で発現したタンパク質は、各タンパク質の量を定量し得るように分離する。各タンパク質の量を、非処理生物学的サンプル中の対応するタンパク質の量と比較する。両サンプルのタンパク質量の差は、処理サンプル中の試験化合物に対する反応を示す。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。

50

【0226】

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。生物学的サンプルから得たタンパク質は、本発明のポリペプチドに特異的な抗体を用いてインキュベートする。抗体により認識されたタンパク質の量を定量する。処理された生物学的サンプル中のタンパク質の量を、非処理生物学的サンプル中のタンパク質の量と比較する。両サンプルのタンパク質量の差は、処理サンプル中の試験化合物に対する反応を示す。

【0227】

マイクロアレイは、本技術分野でよく知られている方法を用いて調製し、使用し、そして分析する (Brennan, T. M. ら (1995) の米国特許第5, 474, 796号、Schena, M. ら (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619、Baldeschweiler らの (1995) PCT出願第WO95/251116号、Shalon, D. らの (1995) PCT出願第WO95/35505号、Heller, R. A. ら (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155、Heller, M. J. らの (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。

【0228】

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、コード配列全体で非コード配列が好ましい。例えば、多重遺伝子ファミリーのメンバー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 (Harrington, J. J. ら (1997) Nat Genet. 15:345-355、Price, C. M. (1993) Blood Rev. 7:127-134、Trask, B. J. (1991) Trends Genet. 7:149-154等を参照)。一度マッピングすると、本発明の核酸配列を用いて例えば病状の遺伝を特定の染色体領域の遺伝または制限酵素断片長多型 (RFLP) と関連させるような遺伝子連鎖地図を作成し得る。(例えば、Lander, E. S. およびD. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照。)

蛍光原位置ハイブリッド形成法 (FISH) は、他の物理的及び遺伝地図データと関連し得る (前出のHeinz-Ulrich, ら (1995) in Meyers, 965-968ページ、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との相関性、或いは特定の疾患に対する素因が、このような疾患と関連するDNA領域の決定に役立つため、更なる位置を決定するクローニングが行われる。

【0229】

確定した染色体マーカーを用いた結合分析等の物理的マッピング技術及び染色体標本原位置ハイブリッド形成法を用いて、遺伝地図を拡張することができる。例えばマウスなど別の哺乳動物の染色体上に遺伝子を配置することにより、正確な染色体の遺伝子座がわかっていない場合でも関連するマーカーを明らかにし得る。この情報は、位置クローニングその他の遺伝子発見技術を用いて疾患遺伝子を探す研究者にとって価値がある。疾患または

症候群が、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、該領域に対するいかなるマッピングも、更なる調査のための関連遺伝子或いは調節遺伝子を表すことができる (Gatti, R. A. ら (1988) Nature 336: 577-580 等を参照) 転座、反転等に起因する、健常者、保有者、感染者の三者間における染色体位置の相違を発見するために本発明のヌクレオチド配列を用いてもよい。

【0230】

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、細胞内に位置することになる。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。

【0231】

別の薬剤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, らの (1984) PCT出願番号 WO84/03564 等を参照)。この方法においては、多数の異なる小さな試験用化合物を固体基質上で合成する。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に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。

【0232】

別の実施例では、SECPと結合可能な中和抗体がSECPと結合するため試験用化合物と特に競合する、競合的薬剤スクリーニングアッセイを用いることができる。この方法では、抗体を用いて、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も検出できる。

【0233】

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

【0234】

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

本明細書において開示した全ての特許、特許出願及び刊行物、特に米国特許第60/214,601号、第60/212,890号、第60/222,372号、第60/213,466号、第60/231,435号および第60/232,889は、言及することをもって本明細書の一部となす。

【0235】

(実施例)

1 cDNAライブラリの作製

Incyte cDNAは、LIFESeq GOLDデータベース (Incyte Genomics, Palo Alto CA) に記載されたcDNAライブラリに由来するものであり、表4の列5に列記した。

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た組織もあり、また、ホモジナイズしてフェノールまたは好適な変性剤の混合液に溶解した組織もある。変性剤の混合液は、例えばフェノールとグアニジニウムイソチオシアネートの単相溶液であるTRIZOL (Life Technologies) 等である。結果として得られた溶解物は、塩化セシウムクッション上で遠心分離するかクロロホルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方、或いは別の方法を用いて、

10

20

30

40

50

溶解物からRNAを沈殿させた。

【0236】

RNAの純度を高めるため、RNAのフェノールによる抽出及び沈殿を必要な回数繰り返した。場合によっては、DNアーゼでRNAを処理した。殆どのライブラリでは、オリゴd(T)連結常磁性粒子(Promega)、OLIGOTEXラテックス粒子(QIAGEN, Valencia 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。

【0237】

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公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, unit 5.1-6.6等を参照)。逆転写は、オリゴd(T)またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖cDNAに連結反応させ、好適な制限酵素でcDNAを消化した。殆どのライブラリに対して、cDNAのサイズ(300~1000bp)選択は、SEPHACRYL S1000、SEPHAROSE CL2BまたはSEPHAROSE CL4Bカラムクロマトグラフィー(Amersham Pharmacia Biotech)、或いは調製用アガロースゲル電気泳動法を用いて行った。好適なプラスミド、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)またはpINCY(Incyte Genomics, Palo Alto CA)、或いはその誘導体等の、ポリリンカーの適合する制限酵素部位にcDNAをつないだ。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DH5α、DH10BまたはELECTROMAX DH10Bを含む大腸菌細胞に形質転換した。

【0238】

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。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及びQIAWELL 8 Ultra Plasmid 精製システム、R. E. A. L. Prep 96プラスミドキットの中から少なくとも1つを用いて、プラスミドを精製した。沈殿させた後、0.1 ml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。

【0239】

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)。宿主細胞の溶解及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 II蛍光スキャナ(Labsystems Oy, Helsinki, Finland)を用いて蛍光分析的に定量した。

【0240】

10

20

30

40

50

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) に与えられた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対呼び出し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よく知られている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, unit 7.7 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。

10

20

30

40

50

【0241】

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プログラムを用いた。次に、BLAST、FASTA 及び BLIMPS に基づくプログラムを用いて、プログラム中の注釈を得るべく、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 及び PFAM 等の Hidden Markov Model (HMM) ベースのタンパク質ファミリーデータベースの選択に対する Incyte cDNA 配列またはその翻訳を問い合わせた HMM は、遺伝子ファミリーのコンセンサス 1 次構造を解析する確率的アプローチである (Eddy, S. R. (1996) Curr. 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。或いは、GenBank cDNA、GenBank EST、ステッチされた配列、ストレッチされた配列または Genscan 予測コード配列 (実施例 4 及び 5を参照) を用いて Incyte cDNA の集団を完全長まで伸長させた。Phred、Phrap 及び Consed に基づくプログラムを用いて構築し、GenMark、BLAST 及び FASTA 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対してスクリーニングした。対応する完全長ポリペプチド配列を誘導するべく完全長ポリヌクレオチド配列を翻訳した。或いは、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。引き続き、GenBank タンパク質データベース (genpept)、SwissProt、BLOCKS、PRINTS、DOMO、PRODOM 及び Prosite 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に対する問合せによって完全長ポリペプチド配列を分析した。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) 及び LASERGENE ソフトウェア (DNASTAR) を用いて分析した。ポリヌクレオチド及びポリペプチド配列アラインメントは、アラインメントした配列と配列の一致率も計算する MEGALIGN マルチシークエンスアラインメントプログラム (DNASTAR) に組み込まれてい

るようなCLUSTALアルゴリズムによって特定されるデフォルトパラメータを用いて生成する。

【0242】

Incyte cDNA及び完全長配列の分析及びアセンブリに利用したツール、プログラム及びアルゴリズムの概略と、適用可能な説明、引用文献、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は好適な引用文献であり、全ての文献はそっくりそのまま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（スコアが高ければ高いほど2配列間の相同性が高くなる）。

10

【0243】

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: 45-88のポリヌクレオチド配列断片の同定にも利用できる。ハイブリダイゼーション及び増幅技術に有用である約20～約4000ヌクレオチドの断片を表4の列4に示した。

【0244】

4 ゲノムDNAからのコード配列の同定及び編集

推定分泌蛋白は、公共のゲノム配列データベース（例えば、gbpriやgbhtg）においてGenscan遺伝子同定プログラムを実行して初めに同定された。Genscanは、様々な生物からゲノムDNA配列を分析する汎用遺伝子同定プログラムである（Burge, C. および S. Karlin (1997) J. Mol. Biol. 268: 78-94 及び Burge, C. 及び S. Karlin (1998) Cuff. Opin. Struct. Biol. 8: 346-354 参照）。プログラムは予測エキソンを連結し、メチオニンから停止コドンに及ぶ構築されたcDNA配列を形成する。Genscanの出力は、ポリヌクレオチド及びポリペプチド配列のFASTAデータベースである。Genscanが一度に分析する配列の最大範囲は、30kbに設定した。これらのGenscan推定cDNA配列の内、どの配列が分泌蛋白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蛋白について問合せて分析した。潜在的な分泌蛋白が、分泌蛋白としてアノテーションが付けられたインサイトcDNA配列に対する相同性を基に同定された。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gbpri及びgbhtgと比較した。必要であれば、genpeptからヒットしたトップのBLASTと比較することによりGenscan予測配列を編集し、余分なまたは取り除かれたエキソンなどのGenscanにより予測された配列のエラーを修正する。BLAST分析はまた、任意のIncyte cDNAまたはGenscan予測配列の公共cDNA適用範囲の発見に用いられるので、転写の証拠を提供する。Incyte cDNA適用範囲が利用できる場合には、この情報を用いてGenscan予測配列を修正または確認した。完全長ポリヌクレオチド配列は、実施例3に記載された構築プロセスを用いて、Incyte cDNA配列及び／または公共のcDNA配列でGenscan予測コード配列を構築することにより得た。或いは、完全長ポリヌクレオチド配列は編集または非編集のGenscan予測コード配列に完全に由来する。

20

30

40

【0245】

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)

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つ若しくは複数のゲノム配列から予測されたGenscanエキソンを含むクラスタに分解した。cDNA及びゲノム情報を統合するべくグラフ理論及び動的プログラミングに基づくアルゴリズムを用いて各クラスタを分析し、引き続いて確認、編集または伸長して完全長配列を産出するような

50

潜在的スプライス変異体を生成した。間隔全体の長さがクラスタ中の2以上の配列に存在するような配列を同定し、そのように同定された間隔は推移により等しいと考えられた。例えば、1つのcDNA及び2つのゲノム配列に間隔が存在する場合、3つの間隔は全て等しいと考えられる。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) に沿って現われるようにステッチアルゴリズムで縫い合わせ、可能な最も長い配列および変異配列を作製する。1種類の親配列に沿って発生した間隔と間隔との連鎖 (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。結果として得られるステッチ配列は、BLAST分析により公共データベース genpept 及び gbpr i に翻訳されて比較された。Genscanにより予測された不正確なエキソンは、genpept からヒットしたトップのBLASTと比較することにより修正した。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。

10

【0246】

ストレッチ配列 (Stretched Sequence)

部分DNA配列は、BLAST分析に基づくアルゴリズムにより完全長まで伸長された。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に記載され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

20

【0247】

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: 45-88を構築するために用いた配列を、BLAST及びSmith-Watermanアルゴリズムを用いて、Incyte LIFESEQデータベース及び公共のドメインデータベースの配列と比較した。SEQ ID NO: 45-88と一致するこれらのデータベースの配列を、Phrap (表7) などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。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前もってマッピングされたかを測定した。マッピングされた配列がクラスタに含まれている結果、個々の配列番号を含めてそのクラスタの全配列が地図上の位置に割り当てられた。

30

40

【0248】

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。センチモルガン間隔の地図上の位置は、染色体のpアームの末端に関連して測定する。(センチモルガン (cM) は、染色体マーカー間の組換え頻度に基づく計測単位である。平均して、1cMは、ヒト中のDNAの1メガベース (Mb) にほぼ等しい。尤も、この値は、組換えのホットスポット及びコールドスポットに起因して広範囲に変化する。) cM距離は、配列が各クラスタ内に含まれるような放射線ハイブリッドマーカーに対して境界を提供するようなGenethonによってマッピングされた遺伝マーカーに基づく。NCBI「GeneMap99」 (<http://www.ncbi.nlm.nih.gov/genemap>) などの一般個人が入手可能なヒト遺伝子マップおよびその他の情報源を用いて、上記し

50

た区間が既に同定されている疾患遺伝子マップ内若しくは近傍に位置するかを決定できる。

このようにして、SEQ ID NO: 48は72. から77. 4センチモルガン内で染色体15にマップされた。

【0249】

このようにして、SEQ ID NO: 54は6. 20から9. 40センチモルガン内で染色体20にマップされた。SEQ ID NO: 61は、0. 00～19. 50センチモルガン以内の間隔で染色体22にマッピングした。

【0250】

このようにして、SEQ ID NO: 82は0. 0から19. 5センチモルガン内で染色体22にマップされた。SEQ ID NO: 85は、84. 7～92. 5センチモルガン以内および137. 5から145. 7センチモルガン以内で染色体12にマッピングされた。SEQ ID NO: 85については一つ以上のマップ位置が報告され、異なったマップ位置を持つ配列が、一つのクラスターに構築されたことを示す。極めて類似しているが、完全に同一でない配列が一つのクラスターに構築される場合にこの状態が発生する。

【0251】

このようにして、SEQ ID NO: 66は65. 60から72. 60センチモルガン内で染色体16にマップされた。このようにして、SEQ ID NO: 67は59. 50から65. 00センチモルガン内で染色体11にマップされた。このようにして、SEQ ID NO: 69は132. 70から144. 40センチモルガン内で染色体6にマップされた。

7 ポリヌクレオチド発現の分析

ノーザン分析は、転写された遺伝情報の存在を検出するために用いられる実験技術であり、特定の細胞種または組織からのRNAが結合される膜への標識されたヌクレオチド配列のハイブリダイゼーションを含む。(前出のSambrook, 7章、同Ausubel, F. M. ら, 4章及び16章等を参照。)

BLASTに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する。ノーザン分析は、多数膜系ハイブリダイゼーションよりも非常に速い。更に、特定の同一を厳密な或いは相同的なものとして分類するか否かを決定するため、コンピュータ検索の感度を変更することができる。検索の基準は積スコアであり、次式で定義される。

【数1】

(BLAST スコア×配列一致率)

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

【0252】

積スコアは、2つの配列間の類似度及び配列が一致する長さの両方を考慮している。積スコアは、0～100の規準化された値であり、次のようにして求める。BLASTスコアにヌクレオチドの配列一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除する。高スコアリングセグメント対(HSP)に一致する各塩基に+5のスコアを割り当て、各不適性塩基対に-4を割り当てることにより、BLASTスコアを計算する。2つの配列は、2以上のHSPを共有し得る(ギャップにより隔離され得る)。2以上のHSPがある場合には、最高BLASTスコアの塩基対を用いて積スコアを計算する。積スコアは、断片的重畳と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%重畳しているか、他端が88%一致し、100%重畳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%重畳しているか、他端が79%一致し、100%重畳しているかの

いずれかの場合に得られる。

【0253】

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。例えば或る完全長配列は、Incycyte cDNA配列（実施例3を参照）と少なくとも一部は重畳するように構築される。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、以下の生物／組織カテゴリー即ち心血管系、結合組織、消化器系、胎芽構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液及び免疫系、肝、筋骨格系、神経系、脾臓、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。同様に、各ヒト組織は、以下の疾患／病状カテゴリー即ち癌、細胞系、発達、炎症、神経性、外傷、心血管、鬱血、その他の1つに分類される。

10

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（Incycyte Genomics, Palo Alto CA）から得ることができる。

【0254】

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、完全長分子の適切な断片から設計したオリゴヌクレオチドプライマーを用いて該断片を伸長させて生成した。一方の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、他方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。開始プライマーの設計は、長さが約22～30ヌクレオチド、GC含有率が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。

20

【0255】

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、付加的プライマー或いはプライマーのネステッドセットを設計した。

30

【0256】

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、 Mg^{2+} と $(NH_4)_2SO_4$ とβ-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℃，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。

40

【0257】

各ウェルのDNA濃度は、1X TE及び0.5 μlの希釈していないPCR産物に溶解した100 μlのPICO GREEN定量試薬（0.25 (v/v) PICO GREEN

50

N ; M o l e c u l a r P r o b e s , E u g e n e O R) を不透明な蛍光光度計プレート (C o m i n g C o s t a r , A c t o n M A) 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の濃度を定量するべくプレートをFluoroskan II (L a b s y s t e m s O y , H e l s i n k i , F i n l a n d) でスキャンした。反応混合物のアリコート5~10 μ l を1%アガロースミニ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決定した。

【0258】

伸長させたヌクレオチドは、脱塩及び濃縮して384穴プレートに移し、CviJIコラウイルスエンドヌクレアーゼ (M o l e c u l a r B i o l o g y R e s e a r c h , M a d i s o n W I) を用いて消化し、pUC18ベクター (A m e r s h a m P h a r m a c i a B i o t e c h) 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・シークエンシングのために、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AgarACE (P r o m e g a) で消化した。伸長させたクローンをT4リガーゼ (N e w E n g l a n d B i o l a b s , B e v e r l y M A) を用いてpUC18ベクター (A m e r s h a m P h a r m a c i a B i o t e c h) に再連結し、PfuDNAポリメラーゼ (S t r a t a g e n e) で処理して制限部位の張出部 (o v e r h a n g) 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℃で一晩培養した。

【0259】

細胞を溶解して、Taq DNAポリメラーゼ (A m e r s h a m P h a r m a c i a B i o t e c h) 及びPfu DNAポリメラーゼ (S t r a t a g e n e) 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。上記したようにPICOGREEN試薬 (M o l e c u l a r P r o b e s) 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 エネルギー転送ファクター シークエンシングプライマー、及びDYENAMIC DIRECT キット (A m e r s h a m P h a r m a c i a B i o t e c h) またはABI PRISM BIGDYE ターミネーターサイクル シークエンシング反応キット (T e r m i n a t o r c y c l e s e q u e n c i n g r e a d y r e a c t i o n k i t) (A p p l i e d B i o s y s t e m s) を用いてシーケンシングした。

【0260】

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。

【0261】

9 個々のハイブリダイゼーションプローブの標識及び使用

SEQ ID NO: 45-88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事実上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 (N a t i o n a l B i o s c i e n c e s) 等の最新ソフトウェアを用いて設計し、各オリゴマー50 pmolと、 $[\gamma\text{-}^{32}\text{P}]$ アデノシン3リン酸 (A m e r s h a m P h a r m a c i a B i o t e c h) 250 μ Ciと、T4ポリヌクレオチドキナーゼ (D u P o n t N E N , B o s t o n M A) を混合することにより標識

10

20

30

40

50

する。標識したオリゴヌクレオチドは、SEPHADEX G-25超細繊維分子サイズ排除デキストラン ビードカラム (Amersham Pharmacia Biotech) を用いて十分に精製する。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されたヒトゲノムDNAの典型的な膜ベースのハイブリダイゼーション解析において、毎分 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。

【0262】

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 (Nytan Plus, Schleicher & Schuell, Durham NH) に移す。ハイブリダイゼーションは、40℃で16時間行う。非特異的シグナルを除去するため、例えば0.1×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ー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。

【0263】

10 マイクロアレイ

マイクロアレイの表面上でアレイエレメントの連鎖または合成は、フォトリソグラフィ、圧電印刷 (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質は、均一且つ非多孔性の固体とするべきである (Schena (1999) 前出)。推奨する基質には、シリコン、シリカ、スライドガラス、ガラスチップ及びシリコンウエハがある。或いは、ドットプロット法またはスロットプロット法に類似のアレイを利用して、熱的、紫外線の、化学的または機械的結合手順を用いて基質の表面にエレメントを配置及び結合させてもよい。通常のアレイは、当業者に周知の利用可能な方法や機械を用いて作製でき、任意の適正数のエレメントを有し得る (Schena, M. ら (1995) Science 270: 467-470, Shalton, D. ら (1996) Genome Res. 6: 639-645, Marshall, A. および J. Hodgson (1998) Nat. Biotechnol. 16: 27-31. を参照)。

【0264】

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レメントと成り得る。ハイブリダイゼーションに好適な断片またはオリゴマーを、レーザGENEソフトウェア (DNASTAR) 等の本技術分野で公知のソフトウェアを用いて選択することが可能である。アレイエレメントは、生物学的サンプル中でポリヌクレオチドを用いてハイブリダイズされる。生物学的サンプル中のポリヌクレオチドは、検出を容易にするために蛍光標識またはその他の分子タグに抱合される。ハイブリダイゼーション後、生物学的サンプルからハイブリダイズされていないヌクレオチドを除去し、蛍光スキャナを用いて各アレイエレメントにおいてハイブリダイゼーションを検出する。或いは、レーザ脱離及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エレメントにハイブリダイズする各ポリヌクレオチドの相補性の度合及び相対存在度は、算定し得る。一実施例におけるマイクロアレイの調整及び使用について、以下に詳述する。

【0265】

組織または細胞サンプルの調製

グアニジウムチオシアネート法を用いて組織サンプルから全RNAを単離し、オリゴ (dT) セルロース法を用いてポリ (A)⁺ 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/μl のオリゴ (dT) プライマー (21mer)、1×第1鎖バッファ、0.03 unit/μl のRNアーゼ阻害剤、500 μM の dATP、500 μM の dGTP、500 μM の dTTP、40 μM の dCTP、40 μM の dCTP-Cy3 (BDS) または dCTP-Cy5 (Amersham Phar

10

20

30

40

50

macia Biotech) を用いて逆転写する。逆転写反応は、GEMBRIGHT キット (Incyte) を用いて 200 ng ポリ (A)⁺ RNA 含有の 25 体積 ml 内で行う。特異的対照ポリ (A)⁺ RNA は、非コード酵母ゲノム DNA から i n v i t r o 転写により合成する。各反応サンプル (1 つは Cy 3、もう 1 つは Cy 5 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 ゲル濾過スピンカラム (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) を用いて精製する。結合後、2 つの反応サンプルは、1 ml のグリコーゲン (1 mg/ml) を用いて析出させたエタノール、60 ml の酢酸ナトリウム及び 300 ml の 100% エタノールである。サンプルは次に、Speed VAC (Savant Instruments Inc., Hollbrook NY) を用いて乾燥して仕上げ、14 µl の 5×SSC/0.2% SDS 中で再懸濁する。

【0266】

マイクロアレイの調製

本発明の配列を用いて、アレイエレメントを生成する。各アレイエレメントは、クローン化 cDNA インサートによりベクター含有細菌性細胞から増幅する。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 µg より大きい最終量までアレイエレメントを増幅する。増幅されたアレイエレメントは、SEPHACRYL-400 (Amersham Pharmacia Biotech) を用いて精製される。

【0267】

精製したアレイエレメントは、ポリマーコートされたスライドガラス上に固定する。顕微鏡スライドガラス (Corning) は、処理中及び処理後に 0.1% の SDS 及びアセトン中で超音波をかけ、蒸留水で非常に良く洗って洗浄する。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、蒸留水中で広範囲にわたって洗浄し、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。

【0268】

米国特許第 5,807,522 号に記載されている方法を用いて、コーティングしたガラス基板にアレイエレメントを付加する。この特許に引用することを以て本明細書の一部とする。平均濃度が 100 ng/µl のアレイエレメント DNA 1 µl 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。

【0269】

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。

【0270】

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×SSC、0.2% SDS ハイブリダイゼーション緩衝液に Cy 3 及び Cy 5 標識した cDNA 合成産物を各 0.2 µg 含む 9 µl 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 cm² 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。チェンバーのコーナーに $140\mu\text{l}$ の $5\times\text{SSC}$ を加えることにより、チェンバー内部を湿度 100% に保持する。アレイを含むチェンバーは、 60°C で約 6.5 時間インキュベートする。アレイは、第1洗浄緩衝液中($1\times\text{SSC}$, $0.1\%\text{SDS}$)において 45°C で 10 分間洗浄し、第2洗浄緩衝液中($0.1\times\text{SSC}$)において 45°C で 10 分間各々 3 度洗浄して乾燥させる。

【0271】

検出

レポーター標識ハイブリダイゼーション複合体は、Cy3の励起のためには 488nm 、Cy5の励起のためには 632nm でスペクトル線を発生し得るInnova 70混合ガス 10W レーザ(Coherent, Inc., Santa Clara CA)を備えた顕微鏡で検出する。 $20\times$ 顕微鏡対物レンズ(Nikon, Inc., Melville NY)を用いて、アレイ上に励起レーザ光の焦点を当てる。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、対物レンズを通過してラスタースキャンする。本実施例で用いた $1.8\text{cm}\times 1.8\text{cm}$ のアレイは、 $20\mu\text{m}$ の解像度でスキャンした。

10

【0272】

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

20

【0273】

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

30

【0274】

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

40

【0275】

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、GEMTOOLS遺伝子発現分析プログラム(Incyte)である。

【0276】

1.1 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SEC

50

Pの発現を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、これより小さな或いは大きな配列の断片の場合でも本質的に同じ方法を用いることができる。Oligo4.06ソフトウェア(National Biosciences)及びSECPのコーディング配列を用いて、適切なオリゴヌクレオチドを設計する。転写を阻害するためには、最も独特な5'配列から相補的オリゴヌクレオチドを設計し、これを用いてプロモーターがコーディング配列に結合するのを阻害する。翻訳を阻害するためには、相補的なオリゴヌクレオチドを設計して、リボソームがSECPをコードする転写物に結合するのを阻害する。

【0277】

12 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、抗生物質耐性及びcDNAの転写レベルを高める誘導性のプロモーターを含む好適なベクターにcDNAをサブクローニングする。このようなプロモーターには、lacオペレーター調節エレメントに関連するT5またはT7バクテリオファージプロモーター及びtrp-lac(tac)ハイブリッドプロモーターが含まれるが、これらに限定するものではない。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β-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)を感染させて行う。バキュロウイルスの非必須ポリヘドリン遺伝子を、相同組換え或いは転移プラスミドの媒介を伴う細菌の媒介による遺伝子転移のどちらかによって、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、強い多角体プロモーターによって高いレベルのcDNAの転写が行われる。組換えバキュロウイルスは、多くの場合はSpodoptera frugiperda(Sf9)昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる。(Engelhard, E. K. ら (1994) Proc. Natl. Acad. Sci. USA 91: 3224-3227, Sandig, V. ら (1996) Hum. Gene Ther. 7: 1937-1945. 等を参照)。

【0278】

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、タンパク質の活性及び抗原性を維持した状態で、固定化グルタチオン上で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)を用いて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。タンパク質の発現及び精製の方法は、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16、及び17のアッセイを行うことができる。

【0279】

13 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。選り抜きのベクターには、pCMV SPORTプラスミド(Life Technologies)及び

10

20

30

40

50

pCR 3.1プラスミド (Invitrogen) が含まれ、どちらもサイトメガロウイルスプロモーターを有する。リポソーム製剤或いは電気穿孔法を用いて、5～10 μgの組換えベクターをヒト細胞株、例えば内皮由来または造血由来の細胞株に一時的に形質移入する。更に、標識タンパク質をコードする配列を含む1～2 μgのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、cDNAの組換えベクターからの発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質 (GFP; Clontech)、CD64またはCD64-GFP融合タンパク質から選択できる。自動化された、レーザー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、プロピジウムヨウ化物によるDNA染色によって計測される核DNA内容物の変化、前方光散乱と90°側方光散乱によって計測される細胞サイズと顆粒状性の変化、プロモデオキシウリジンの取込量の低下によって計測されるDNA合成の下方調節、特異抗体との反応性によって計測される細胞表面及び細胞内におけるタンパク質の発現の変化、及び蛍光複合アネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変化とがある。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

20

【0280】

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、形質転換された細胞表面で発現し、ヒト免疫グロブリンG (IgG) の保存領域と結合する。形質転換された細胞と形質転換されない細胞とは、ヒトIgG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及び対象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。

30

【0281】

1.4 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. 1816-3088-495を参照) または他の精製技術で実質的に精製されたSECPを用いて、標準的なプロトコルでウサギを免疫化して抗体を作り出す。

【0282】

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、免疫抗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)。

40

【0283】

通常は、長さ約15残基のオリゴペプチドを、Fmocケミストリーを用いる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、免疫抗原性を高める (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-KLM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

50

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。

【0284】

1.5 特異的抗体を用いる天然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ーにより実質的に精製する。イムノアフィニティークラムは、CNBr-活性化SEPHAROSE (Amersham Pharmacia Biotech) のような活性化クロマトグラフィー用レジンと抗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樹脂をブロックし、洗浄する。

10

【0285】

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を優先的に吸着できる条件で（例えば、界面活性剤の存在下において高イオン強度のバッファーで）そのカラムを洗浄する。そのカラムを、抗体と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。

【0286】

1.6 SECPと相互作用する分子の同定

SECPまたは生物学的に活性であるSECP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。

20

【0287】

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 (yeast two-hybrid system) やMATCHMAKERシステム (Clontech) などの2-ハイブリッドシステムに基づいた市販のキットを用いて分析する。

30

【0288】

SECPはまた、ハイスループット型の酵母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)。

【0289】

1.7 SECPの活性の実証

SECPの活性を測定するアッセイは、最終体積が100 μl の溶液中に25 mM のトリス塩酸緩衝液 (pH 7.4)、0.25% Triton X-100、5 mM の塩化マンガン、5 mM のCDP-コリン、5 mM の2-メルカプトエタノール、0.05 mM のUDP- ^{14}C GalNAc (4,000 cpm/ナノモル)、250 μM ペプチド、および様々な量のSECPを含む酵素反応混合液なので構成される。反応混合液は37°Cで10分間インキュベートし、次に、Dowex 1イオン交換（蟻酸）クロマトグラフィーを行う。溶出されたペプチドを含む分画をシンチレーションカウンターにかける。ペプチド含有分画に存在する ^{14}C GalNAcの量はSECP活性に比例する。基質とSECP源の確認はC-18 クロマトグラフィー (C2C18 3.2 Smart System, Pharmacia Biotech Inc.) によって評価して、ペプチドの安定性を確認し、取り込まれた ^{14}C GalNAcが

40

50

ペプチドと会合していることを確認できる (Sorensen, T. ら (1995) J. Biol. Chem. 270:24166-24173)。

【0290】

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、SECPの量の変更が、放射性DNA前駆物質である [^3H] チミジンの存在中で静止状態3T3培養細胞へと加えられる。このアッセイのためのSECPは、組み替え型の手段を用いて得られても良く、また生化学的な標本より得られても良い。酸-沈殿可能DNAへの [^3H] チミジンの混合は、適切な時間間隔で測定され、混合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濃度レンジにわたる一次の用量-応答曲線は、成長調節活性を示している。ミリリットルあたりの活性の一つのユニットは、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-沈殿可能DNAへの [^3H] チミジンの最大の混合を示している。

10

【0291】

代替例として、SECP活性のためのアッセイは、培養細胞中の神経伝達物質の刺激作用若しくは抑制を測定する。培養CHO繊維芽細胞はSECPに晒されている。エンドサイトーシスのSECP取込後に、細胞を新鮮培養液で洗浄し、アフリカツメガエル筋細胞を固定した全体細胞が、SECPを含まない媒質中の繊維芽細胞の一つと接触させて処置する。膜電流は筋細胞より測定される。対照値に対し増加若しくは減少した電流は、SECPの神経調節性効果を示している (Morimoto, T. ら (1995) Neuron 15:689-696)。

20

代替例として、SECP活性のアッセイは、分泌性の膜結合細胞小器官におけるSECPの量を測定する。上述したように形質移入された細胞が採取され、溶解された。ライセートは、例えば当業者にショ糖密度勾配超遠心法として知られるような方法を用いて分別される。そのような方法は、ゴルジ体、ER、小膜結合小胞、及びその他の分泌細胞小器官のような、細胞内要素の隔離を可能とする。分割された細胞ライセート及び全体の細胞ライセートよりの免疫沈降は、SECP特定抗体を用いて実行され、また免疫沈降サンプルはSDS-PAGE及び免疫ブロット技術を用いて解析される。全体の細胞ライセートにおけるSECPに対する分泌オルガネラ中のSECP濃度は、分泌経路を通る通過するSECPの量に比例する。

30

【0292】

他の例において、SECPは血清から抗原を認識し、沈殿する。この活性は定量沈降反応によって測定される (Golub, E. S. ら (1987) Immunology: A Synthesis, Sinauer Associates, Sunderland, MA, pages 113-115)。SECPは当分野において既知の方法を用いて同位体的に標識される。種々の血清濃度で一定量の標識SECPに加える。SECP抗原複合物が溶液に沈殿し、遠心分離によって回収する。沈殿性SECP抗原複合物の量は沈殿物中に検出される放射性同位元素の量に比例する。沈殿性SECP抗原複合物の量を血清濃度に対してプロットする。異なった血清濃度の特徴的沈降曲線が得られ、沈殿性SECP抗原複合物の量は最初は血清中濃度の増加に比例して増加し、等価点を頂点とし、その後は血清中濃度の増加に比例して減少する。このように、沈殿性SECP抗原複合物の量は、抗原の制限量と過剰量に対する感受性を特徴とするSECP活性の測定方法となる。

40

別な方法として、SECP活性に対するアッセイは、細胞表面におけるSECPの発現を測定する。SECPをコードするcDNAを、非白血球細胞系に形質移入する。細胞表面タンパク質をビオチンで標識する (de la Fuente, M. A. ら (1997) Blood 90:2398-2405)。SECP特異的抗体を用いて免疫沈降を

50

行い、免疫沈降したサンプルを S D S - P A G E と免疫ブロット法を用いて分析する。標識された免疫沈降と未標識免疫沈降の比は、細胞表面に発現した S E C P の量に比例する。

【0293】

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の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特定の好適実施例に関連して説明を行ったが、本発明の範囲が、そのような特定の実施例に不当に制限されるべきではないことを理解されたい。実際に、分子生物学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改変は、特許請求の範囲内にあるものとする。

10

【0294】

(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。

【0295】

表2は、GenBank 識別番号及び本発明のポリペプチドに最も近いGenBank 相同体の注釈を示す。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。

【0296】

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。

20

【0297】

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。

【0298】

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。

【0299】

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。

30

【0300】

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

【表1】

表 1-1

Incyte エント ID	ポリペプチド SEQ ID NO:	Incyte ペプチド ID	ポリ SEQ ID NO:	Incyte ポリスクレオチド ID
2101688	1	2101688CD1	45	2101688CB1
5452330	2	5452330CD1	46	5452330CB1
4362432	3	4362432CD1	47	4362432CB1
5308104	4	5308104CD1	48	5308104CB1
3092736	5	3092736CD1	49	3092736CB1
3580257	6	3580257CD1	50	3580257CB1
3634758	7	3634758CD1	51	3634758CB1
4027923	8	4027923CD1	52	4027923CB1
4348533	9	4348533CD1	53	4348533CB1
4521857	10	4521857CD1	54	4521857CB1
4722253	11	4722253CD1	55	4722253CB1
4878134	12	4878134CD1	56	4878134CB1
5050133	13	5050133CD1	57	5050133CB1
5630124	14	5630124CD1	58	5630124CB1
5677286	15	5677286CD1	59	5677286CB1
6436791	16	6436791CD1	60	6436791CB1
1820972	17	1820972CD1	61	1820972CB1
3286805	18	3286805CD1	62	3286805CB1
3506590	19	3506590CD1	63	3506590CB1
003600	20	003600CD1	64	003600CB1
1251534	21	1251534CD1	65	1251534CB1
1402211	22	1402211CD1	66	1402211CB1
1623474	23	1623474CD1	67	1623474CB1
1706443	24	1706443CD1	68	1706443CB1
1748627	25	1748627CD1	69	1748627CB1
1818332	26	1818332CD1	70	1818332CB1
1822832	27	1822832CD1	71	1822832CB1
1832219	28	1832219CD1	72	1832219CB1
1899010	29	1899010CD1	73	1899010CB1
2008768	30	2008768CD1	74	2008768CB1
2070984	31	2070984CD1	75	2070984CB1

10

20

30

【表 2】

【表 3】

表 1-2

Incyte エント ID	ボリペプチド SEQ ID NO:	Incyte ペプチド ID	ボリ SEQ ID NO:	Incyte ボリヌクレオチド ID
2193240	32	2193240CD1	76	2193240CB1
2235177	33	2235177CD1	77	2235177CB1
2416227	34	2416227CD1	78	2416227CB1
2461076	35	2461076CD1	79	2461076CB1
1957517	36	1957517CD1	80	1957517CB1
866038	37	866038CD1	81	866038CB1
3869704	38	3869704CD1	82	3869704CB1
1415179	39	1415179CD1	83	1415179CB1
1664792	40	1664792CD1	84	1664792CB1
2079396	41	2079396CD1	85	2079396CB1
5390115	42	5390115CD1	86	5390115CB1
1403326	43	1403326CD1	87	1403326CB1
7690129	44	7690129CD1	88	7690129CB1

表 2

ポリペプチド SEQ ID No:	Incyte ポリペプチドID	GenBank ID NO:	塩基スコア	GenBank 相同体
1	2101688CD1	g971461	1.50E-141	UDP-GalNAc: ポリペプチド N-アセチル-ガラクトサミニルトランスフェラーゼ [ヒト] (White, T. 他 J. Biol. Chem. (1995) 270(41):24156-65)
2	5452330CD1	g1139548	0	癌細胞増殖関連遺伝子作成物 6、2 型前駆体 [マウス]](Shimizu-Nishikawa, K. 他 (1995) Biochem. Biophys. Res. Commun. 216:382-389)
3	4362432CD1	g4467410	5.20E-40	リゾチーム [ニワトリ] (Nakano, T. & Graf, T. (1992) Oncogene 7:527-534)
4	5308104CD1	g3878261	2.10E-92	S. Pombe BEN1/BUD5 に類似性 [線虫]
16	6436791CD1	g13274582	5.00E-39	胸腺萎縮症関連タンパク質 [マウス]
17	1820972CD1	g33702	2.20E-106	免疫グロブリンラムダ軽鎖 [ヒト]
18	3286805CD1	g431420	1.50E-283	マクロファージ特異的タンパク質 MPST [マウス] (Spillsbury, K. 他 (1995) Blood 85:1620-1629)
19	3506590CD1	g577056	1.00E-211	C ガンマ 3 [ヒト]
29	1899010CD1	g13384378	8.00E-43	推定リン酸トランスロケータ [イネ]
36	1957517CD1	g1572802	2.90E-65	Enterococcus faecalis TRAB [線虫]
37	866038CD1	g849238	1.90E-30	ポリボース遺伝子産タンパク質 1 に類似 [線虫]
38	3869704CD1	g33718	5.20E-114	免疫グロブリンラムダ軽鎖 [ヒト]
43	1403326CD1	g3983152	8.10E-56	Schlafen 3 リンパ様増殖調節タンパク質 [マウス] (Schwarz, D.A. 他 (1998) Immunity 9:657-668)
44	7690129CD1	g6715117	3.10E-219	ベックウィズ・ウィーデマン症候群染色体領域 11p15.5 に見られる MTR1 [ヒト] メラスタチン/TRP 関連タンパク質 (Prawitt, D. 他 (2000) Hum. Mol. Genet. 9:203-216)

10

20

30

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	2101688CD1	552	S200 S241 S313 S387 S399 S433 S45 S507 S84 S89 T130 T196 T237 T27 T35 T355 T41 T467 T5 Y408 Y74		グリコシル転移酵素: S114-F292 グリコシル転移酵素: I147-D157 (P<0.021) PD03162:N アセチルガラクトサミン転移酵素 転移酵素 ポリペプチド アセチルガラクトサミン転移酵素 UDPGALNAc:ポリペプチドグリコシル転移酵素 PROTEINUDP タンパク質 UDP N: Q256-P414 アセチルガラクトサミン転移酵素:ポリペプチド; DM03891 I37405 21-571: V26-W547 Signal_peptide: M1-R29 膜貫通ドメイン: L4-W25 Signal_cleavage: M1-R29 シグナルペプチド: M1-G19 シグナルペプチド: M1-G19 膜貫通ドメイン: I330-Y947 スシドメイン (SCR リピート): C357-C412, C532-C589, C710-C765, C771-C830, C838-C895 CUB ドメイン: C416-Y524, C593-F701 てんかん発作関連遺伝子生成物前駆体 シグナルタイプ PD024762:H18-A415 PD028803:V911-G984 スリリピート DM04887 P33730 1-610: T735-D901, F381-P450, T548-I631 DM04887 P16581 1-609: T732-Y904, I354-P450, E525-P610 DM04887 P27113 1-551: S722-R896, I354-P450 Signal_cleavage: M1-G19	HMMER_FFAM BLIMPS_FFAM BLAST_PRODOM HMMER_FFAM HMMER_FFAM BLAST_PRODOM HMMER_FFAM HMMER_FFAM BLAST_PRODOM BLAST_DOMO HMMER HMMER SPSCAN HMMER SPSCAN HMMER HMMER_FFAM HMMER_FFAM BLAST_PRODOM BLAST_DOMO SPSCAN
2	5452330CD1	994	S218 S249 S257 S263 S291 S378 S463 S501 S674 S724 S770 S780 S786 S820 S824 S842 S877 S919 S974 T38 T425 T553 T63 T647 T655 T709 T757 T812	N247 N289 N313 N399 N422 N436 N440 N541 N583 N707		
3	4362432CD1	212	S181 S190 S211			SPSCAN

【表 5】

10

20

30

表 3-2

SEQ ID NO.	Incyte 抗体 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
4	5308104CD1	308	S26 T153 T16 T188 T45		Signal peptide: M1-G19	HMNER
					トランスグリコシターゼ SLT ドメイン SLT: T82-A202	HMNER_PPAM
					プテリン 4 アルファカルピノールアミン脱水酵素 G124-K130	BLIMPS_PPAM
					リゾチーム G シグネチャ PR00749: G174-D195, D191-S211, C39-M59, N60-Q81, I84-I102, S103-F123, G124-K142, K157-K173	BLIMPS_PRINTS
5	3092736CD1	328	S154 S158 S201 S5 S79 S93 T225 T253 T55 T71 Y163		リゾチーム G 4 ベータ-N-アセチルラムミダーゼ グルコース型加水分解酵素 PD016787: G38-F212	BLAST_PRODOM
					リゾチーム G	BLAST_DOMO
					DM07376 P00718 1-184: C39-F212 DM07376 P27042 27-210: G38-F212	SPSCAN
					Signal cleavage: M1-G61	HMNER_PPAM
6	3580257CD1	69	S116 S121 S148 S155 S159 S221 S278 S317 S52 T57		ジェンラクトン () 加水分解酵素 フアマミリ DL: P235-H262	MOTIFS
					Tonb_Dependent_Receptor タンパク質 シグネチャ M1-S5	BLAST_PRODOM
					TRAXPINO プラスミド部分 BEM46 のタンパク質遺伝子間	BLAST_PRODOM
					傾成膜貫通 KRELHXT14 PD009919: T113-S216 飯島上の 34.9 KD タンパク質 飯島タンパク質 PD126088: F234-S302	BLAST_PRODOM
7	3634758CD1	158	T34 T55		K04G2.2 タンパク質 PD126091: N2-E40	BLAST_PRODOM
					Signal peptide: M1-A19 Signal cleavage: M1-G22	HMNER SPSCAN
7	3634758CD1	158			Signal cleavage: M1-A21 Signal cleavage: M1-G17	SPSCAN SPSCAN

10

20

30

【表 6】

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
8	4027923CD1	463	S113 S175 S360 S45 S86 T132 T157		Signal_peptide: M1-R37	HMME
9	4348533CD1	648	S179 S244 S265 S303 S327 S329 S337 S389 S551 S571 S586 S620 S639 T276 T425 T470 T49 T496 T599 T606 S10 T75	N161 N310 N313	Signal_cleavage: M1-N68 Leucine_Zipper: L178-L199	SPSCAN MOTIFS
10	4521857CD1	130	S10 T75		Signal_cleavage: M1-A38	SPSCAN
11	4722253CD1	279	S171 S230 S73 S77 T107 T243 T268	N191 N266 N71	膜貫通ドメイン: G20-Y40 Signal_cleavage: M1-A62	HMME SPSCAN
12	4878134CD1	458	S15 S229 S279 S321 S340 S381 S439 T127 T93	N198 N259 N319	膜貫通ドメイン: L22-L41 Rgd: R118-D120	HMME MOTIFS
13	5050133CD1	173	S130 S50		Signal_cleavage: M1-A31	SPSCAN
14	5630124CD1	335	S142 S191 S219 S295 S302 S324 S67 S74 T104 T190 T225 T243 T252 T275 T292 Y332 T42		Signal_peptide: M1-A39 Signal_cleavage: M1-G36	HMME SPSCAN
15	5677286CD1	71	T42		Signal_peptide: M1-A34 Signal_cleavage: M1-A66	HMME SPSCAN
16	6436791CD1	148	S143 S16 T18	N31	膜貫通ドメイン: L109-F126	HMME
17	1820972CD1	231	S140 S206 S219 S74		Signal_peptide: M1-S20 Signal_cleavage: M1-G16 免疫グロブリン; IG; 組換え性; 主基 DM02680 A39949 1-118: V115-C230	HMME SPSCAN BLAST_DOMO

【表 7】

10

20

30

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 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	3286805CD1	716			MHC フレームワークドメイン DM00397 S24319 1-128: M1-PL28	BLAST_DOMO
					B 細胞 μ 鎖結合 8HS20 タンパク質前駆体 PD174509: L23-V108	BLAST_PRODOM
					免疫グロブリンと MHC タンパク質シグネチャ BL00290: T150-S172, Y210-P227	BLIMPS_BLOCKS
					免疫グロブリンと MHC タンパク質シグネチャ Ig_mhc.prf: K190-S231	PROFILES CAN
					免疫グロブリンドメイン ig: V108, A146-V214	HMWER_FFAM
					Ig_Mhc: Y210-H216	MOTIFS
19	3506590CD1	519			Signal_peptide: M1-P22	HMWER
					膜貫通ドメイン: S653-I676	HMWER
					Signal_cleavage: M1-A17	SPSCAN
					Signal_cleavage: M1-C19	SPSCAN
					Signal_peptide: M1-C19	HMWER
					MHC ヒンジドメインドメイン DM01060 P01862 1-329: S142-K275, R285-G518	BLAST_DOMO
					IG ガンマ3 鎖 C 領域 重鎖ドメインタンパク質 HDC	BLAST_PRODOM
					免疫グロブリン 糖タンパク質 PD028815: E241-G309	
					免疫グロブリンと MHC タンパク質シグネチャ BL00290: S436-Q458, F495-S512	BLIMPS_BLOCKS
					免疫グロブリンと MHC タンパク質シグネチャ Ig_mhc.prf: T371-V420, D473-K519	PROFILES CAN

10

20

30

表 3-5

SEQ ID NO:	Incyte ポリヌクレオチド 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	分析方法データベース
20	003600CD1	172	T73 T71 T90 S128		免疫グロブリンドメイン Ig: R117, G162-V227, S326-V395, K432-V499 Ig Mhc: Y223-H229, F495-H501 シグナルペプチド: M6-L26 シグナル切断: M1-A28 膜貫通ドメイン: L12-N30	HMmer PFAM MOTIFS HMmer SPSCAN HMmer
21	1251534CD1	314			ロイシンジッパーモチーフ: L12-L33	MOTIFS
22	1402211CD1	542	S430 S131 S137 S186 T273 S371 S395 T417 T426 S454 T34 S44 T114 S319 T509	N2 N359 N408 N409 N424 N529	シグナルペプチド: M43-M67 膜貫通ドメイン: A250-I267 シグナルペプチド: M345-H366	HMmer HMmer HMmer
23	1623474CD1	715	T66 S121 T216 T334 S376 S380 S386 T436 T475 T524 S543 S585 S586 S647 T659 S704 T709 S5 T108 T222 T279 S372 S390 S395 S406 S429 S445 S455 S503 S590 S639	N238 N335 N61 N239 N461 N465 N535	Rgd モチーフ: R377-D379 シグナルペプチド: M187-V211 膜貫通ドメイン: I49-P67	MOTIFS HMmer HMmer
24	1706443CD1	469	Y228 T70 S102 S158 T283 T337 S364 T37 S168 S179 T182 S232 S316 S359 T436 S462 S466		Rgd モチーフ: R119-D121 シグナルペプチド: M1-G24 シグナル切断: M1-G24	MOTIFS HMmer SPSCAN

10

20

30

【表 9】

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
25	1748627CD1	274	T9 T90 T237 S241 S248 S62 S100 S136 S191 T35	N254 N270	シグナル切断: M1-A59	SPSCAN
26	1818332CD1	154	S120 S136 T41 S56 T76 S98 S138		シグナル切断: M1-A26	SPSCAN
27	1822832CD1	102	T13 T19	N16	シグナルペプチド: M34-P57 Rgd モチーフ: R21-D23	HMME
28	1832219CD1	113			シグナル切断: M1-G29	MOTIFS
29	1899010CD1	313	S127 S145 S300	N43 N92 N97 N98 N238	シグナルペプチド: M194-G211	SPSCAN
30	2008768CD1	195	S35 S49 T64 S78 S117		膜貫通ドメイン: H11-I35, F151-V171, W219-V237	HMME
31	2070984CD1	350	T77	N294	シグナルペプチド: M121-A139	HMME
32	2193240CD1	360	Y327 S220 S221 S7 S38 T135 S318	N159 N207 N218 N142	膜貫通ドメイン: L95-R116, N122-L145	SPSCAN
33	2235177CD1	559	S301 S412 S520 T11 T27 S29 S42 T76 T156 S165 S252 T277 T303 T336 T462 T120 T121 S292 S322 S397 T407 T418	N70 N171 N357 N325 N417	シグナル切断: M1-A66	HMME
34	2416227CD1	198	S136 S167 S137	N38 N68 N75 N92	シグナルペプチド: M191-A209	SPSCAN
					シグナル切断: M1-S18	HMME
					膜貫通ドメイン: F113-L133	HMME

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
35	2461076CD1	73	T40 S25 T41		シグナルペプチド: M1-G21 シグナル切断: M1-V19	HMMSR SPSCAN
36	1957517CD1	376	S87 T94 T196 S257 S326 S38 S224 S280	N36 N307		MOTIFS
37	866038CD1	216	T11 T15 S59 S114 S142 T146 S167 S172 S107 S157 T200 T209 S210 Y68		ロイシンジッパーモチーフ: L129-L150 シグナル切断: M1-G45	MOTIFS SPSCAN
38	3869704CD1	233	S112 S142 S208 S221 S74 T15 T36		シグナルペプチド: M1-A19 シグナル切断: M1-A19 免疫グロブリンと主要組織適合性ドメイン: Y212-H219 免疫グロブリンと主要組織適合性ドメイン ig_mhc.ppf: N191-S233 免疫グロブリンと主要組織適合性ドメイン BL00290: T152-S174, Y212-P229 免疫グロブリンドメイン ig: G34-S108, A148-V216 免疫グロブリン; 主要組織適合性 DM02680 A39949 1-118: V117-C232 免疫グロブリンフレームワークドメイン DM00397 S30526S 1-119: S20-F139 免疫グロブリン DM00001 S29258 119-206: T137-K225	HMMSR SPSCAN MOTIFS PROFILES SCAN BLIMPS_BLOCKS HMMSR_PFM
39	1415179CD1	163	T104 T86		シグナル切断: M1-S35 ミトコンドリア相体: P134-M142	SPSCAN MOTIFS
40	1664792	235	S33 T70 T93 T94 T121 T224		2P 受容体型ドメイン BL00682: C50-L56 シグナルペプチド M1-D18	BLIMPS_BLOCKS HMMSR

10

20

30

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
41	2079396CD1	94	S21 S45		シグナル切断: M1-S42 GTP 結合伸長因子シグネチャ M1-S52 ペルオキシダーゼシグネチャ I37-W90	SPSCAN PROFILES SCAN PROFILES SCAN
42	5390115CD1	85	S3 S8 T16 T63 T81		シグナル切断: M1-S47 膜貫通ドメイン: Y24-I44 P-ループ Atp_Gtp_A: Q599-T606	SPSCAN HMMER MOTIFS
43	1403326CD1	901	S120 S13 S139 S219 S259 S383 S521 S531 S588 S603 S641 S708 S80 S805 S853 S858 T154 T230 T25 T296 T344 T352 T354 T493 T505 T650 T688 T776 T795 T815 Y279 Y311 Y681 Y804 Y824			
44	7690129CD1	1040	S191 S254 S367 S539 S579 S679 S969 S971 S978 T112 T140 T182 T503 T535 T544 T729 T93	N116 N54 N818	Leucine Zipper: L695-L716 Rgd: R40-D42 R241-D243 膜貫通ドメイン: V606-F623, W753-A773, W844-V862 タンパク質染色体膜貫通メラスチン C05Q12.3 T01H8.5 I F54D1.5 IV PD151509: V730-A1018 PD018035: K8-W246 PD039592: Q382-E546	MOTIFS MOTIFS HMMER BLAST_PRODROM

10

20

30

【表 1 2】

表 4 - 1

ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID	配列長	選択された断片	配列断片	5'位置	3'位置
45	2101688CB1	2508	71-123	6735891H1 (LIVRTUT13) 7620180J1 (KIDNTUE01) 6874586H1 (EPIMUNN04) 7704542H1 (UTREFTUE01) 3593046H1 (293TFST01) 6018547H1 (HNT2UNN03) 6124211H1 (BRAHNN05) 7700489J1 (KIDPTDE01) 3470968F6 (BRAIDIT01)	883 1757 1988 192 1 1327 1053 409 1862	1375 2290 2508 634 304 2033 1632 938 2432
46	5452330CB1	4034	1493-1673, 1-1081, 2638-2908, 3129-3535	6982855F8 (BRAIFER05) 4775091H1 (BRAQNOT01) 5404047T6 (BRAHNOT01) 7293087F8 (BRAIFER06) 7583209H1 (BRAIFEC01) 6207435H1 (PTUNON01) 5404047F6 (BRAHNOT01) 7115489H1 (BRAENOR01) 690568H1 (BRAIFER05) 7293087R8 (BRAIFER06) 7579594H1 (BRAIFEC01) 7291338F8 (BRAIFER06) 4362432F6 (SKIRNOT01) 4362432T9 (SKIRNOT01) 71013085V1	1 3798 3471 526 301 3033 2931 2298 480 1078 2446 1717 1 228 1689	427 4034 4026 1205 861 3733 3445 2745 1092 1790 2962 2352 664 845 2273
47	4362432CB1	845	1-44, 685-845	6809635J1 (SKIRNOR01) 8044501J1 (OVARTUE01)	1 214	532 765
48	5308104CB1	2300	1-807, 2192-2300			

10

20

30

【表 1 3】

表 4-2

ポリスクレオチド SEQ ID NO:	Incyte リスクレオチド ID	ポ 配列長	選択された断片	配列断片	5'位置	3'位置
				1550768R6 (PROSNOT06)	1989	2283
				6804176H1 (COLNOR03)	1230	1814
				71014150V1	656	1210
				6880707H1 (BRAHTR03)	1100	1808
				503680H1 (TWLR3DT02)	2119	2300
49	3092736CB1	1587	1-180	SCGA02766V1	367	1073
				SCGA07870V1	685	1131
				1611754F6 (COLNTUT06)	1153	1587
49				2823991F6 (ADRETUT06)	1031	1532
				SCGA12762V1	1	524
50	3580257CB1	669	1-24	3580257F6 (293TF3T01)	133	669
				5107219H1 (PROSTUS19)	1	240
				51496797	1	495
51	3634758CB1	1463	1-51	4719037H1 (BRAHCT02)	1177	1432
				SKAF05002V1	1	521
				SKAF05483V1	379	866
				3243342H1 (BRAINT019)	1231	1463
				2729881H1 (OVARUT04)	1095	1333
				SKAF05152V1	604	1131
52	4027923CB1	1686	1-204, 1666-1686	2932289H1 (GBLANOT02)	963	1179
				1281432F6 (COLNNOT16)	620	1173
				2561353H1 (ADRETUT01)	1	276
				664136H1 (SCORNOT01)	1428	1686
				3585158H1 (293TF4T01)	325	639
				1281432T6 (COLNNOT16)	1051	1684
				6772967J1 (BRAUNOR01)	75	607
53	4348533CB1	2497	1556-1848, 1-150, 2371-2497, 762-909	6933091H1 (SINTTHR02)	1345	1901

10

20

30

【表 1 4】

表 4-3

ポリヌクレオチド SEQ ID NO:	Incyte リヌクレオチド ID	ポ ズ	配列長	選択された断片	断片断片	5'位置	3'位置
					g1617775	1	405
					2890155F6 (LUNGFET04)	1	483
					6781002J1 (OVARDIR01)	153	903
					2622331H1 (KERANOT02)	2139	2497
					2507578T6 (CONUTUT01)	1684	2359
					1728113F6 (PROSNOT14)	546	1160
					6945931H1 (FTUBTUR01)	1132	1795
54	4521857CB1		1783	1-733, 805-890	30031172H1 (TLYMNOT06)	900	1194
					4521857F6 (HNT2TX01)	1	537
					825638T1 (PROSNOT06)	1100	1762
					857689R1 (NGANNOT01)	1194	1777
					3644845F6 (LUNGNOT34)	490	903
					362417R6 (PROSNOT01)	1227	1783
55	4722253CB1		1461	1-499	4722253H1 (COLCTUT02)	933	1204
					7018504H1 (KIDNNOC01)	1	668
					2455753F6 (ENDANOT01)	577	1109
55					93053012	975	1461
					3028265F6 (HEARFET02)	1292	1461
56	4878134CB1		2116	1-1071	3396235H1 (BRAIDIT01)	1865	2116
					4501019F6 (BRAVXT02)	544	1119
					SBQA01857D1	467	1044
					3521653T6 (LUNGNON03)	1181	1713
					5031921H1 (OVARNON03)	1745	2041
					3766951H1 (BRSTNOT24)	1553	1851
					4501019T6 (BRAVXT02)	1050	1704
					70874715V1	1	544
57	5050133CB1		702	1-28, 651-702	91802638	321	702
					6871022H1 (BRAGNON02)	1	630

10

20

30

【表 15】

表 4 - 4

ポリスクレオチド SEQ ID NO:	Incyte リヌクレオチド ID	ボ 配列長	選択された断片	配列断片	5'位置	3'位置
58	5630124CB1	2613	1-975	3729290F6 (SMCCNON03)	199	686
				6821390J1 (SINTNOR01)	520	1293
				6431012H1 (LUNGNON07)	1175	1881
				6855495H1 (BRAIFEN08)	1	643
				3878611T6 (SPLANNOT11)	2075	2590
				1358001T6 (LUNGNOT09)	1940	2586
				481430R7 (LIVRECT01)	854	1381
				2252822R6 (OVARUTUT01)	2311	2613
				481430T7 (LIVRECT01)	1371	2013
59	5677286CB1	1778	1736-1778, 1-143, 672-767	70613827V1	1195	1777
				7053934H2 (BRACNOK02)	602	1280
				6340571H1 (BRANDIN01)	869	1324
				3620887T6 (BRSTNOT24)	1	642
				1810961F6 (PROSTUT12)	1421	1778
60	6436791CB1	1234	1-192	1212854T6 (BRSTTUT01)	556	1221
				3510032F6 (CONCNOT01)	1	590
				3943483F6 (SCORNOT04)	764	1234
61	1820972CB1	863	1-228, 843-863	60144357B1	227	833
				70636975V1	253	863
				1820972H1 (GBLATUT01)	1	267
62	3286805CB1	2521	1-155, 1165-2294	5030319F7 (COLCDIT01)	1300	1973
				7168560H1 (MCLRNOC01)	575	1020
				6959075H1 (SKINDIA01)	1	674
62				3286805F6 (HEAONOT05)	1706	2266
				6466032H1 (PLACFBB01)	764	1319
				7105400SV1	1949	2521
63	3506590CB1	1765	1-798	71409670V1	1051	1765

10

20

30

表 4-5

ポリヌクレオチド SEQ ID NO:	Incyte リヌクレオチド ID	ボ	配列長	選択された断片	配列断片	5'位置	3'位置
					7710638H1 (TESTTUE02)	524	1080
					70515763V1	1	556
					7733848H2 (COLDDIE01)	543	1266
64	003600CB1		1264	1-699	2718319H1 (THYRN0T09)	195	442
					2397316T6 (THPIAZT01)	714	1264
					003600R6 (HMCINGT01)	509	1042
					008108H1 (HMCINGT01)	269	498
					4592276H1 (MASTYXT01)	398	639
					2911566H1 (KIDNTUT15)	1	265
65	1251534CB1		3415	1-122, 2125-2328, 2982-3415, 834-1643	531771T6 (BRAINT03)	2338	2967
					4697183F6 (BRALN0T01)	235	863
					487800H1 (HNT2AGT01)	1633	1906
					4740283H1 (THYMNOR02)	1210	1470
					1717327T6 (UCMCN0T02)	2865	3415
					7761153H1 (THYMNOR02)	1	708
					6855869H1 (BRAIFEN08)	1969	2594
					6442673H1 (BRAEN0T02)	1814	2453
					3291485F6 (BONRFET01)	798	1410
					5404331H1 (BRAHNOT01)	1669	1969
					1251534H1 (LUNGFET03)	1448	1681
66	1402211CB1		2289	1707-2289	1650008F6 (PROSTUT09)	1296	1912
					429727T6 (BLADN0T01)	1626	2289
					2862734H1 (SININOT03)	1082	1342
					2129033R6 (KIDNNOT05)	1	512
					2499655F7 (ADRETUT05)	564	1084
					5397049H1 (LIVRTUT13)	1031	1293
					826301K1 (PROSNOT06)	1332	1928

10

20

30

【表 17】

表 4-6

ポリヌクレオチド SEQ ID NO:	Incyte リヌクレオチドID	ボ 配列長	選択された断片	配列断片	5'位置	3'位置
67	1623474CB1	4480	2411-3066, 1-22, 109-587, 3638-3733	429727R6 (BLADNOT01) 2158031F6 (BRAINNOT09)	335 3973	862 4480
67				3332425T6 (BRAINPET01) 6914395J1 (PITUDIR01) 6918276H1 (PLACPER06) 6128115H1 (BRAHNON05) 70758295V1 3394074H1 (LUNGNOT28) 1864803T6 (PROSNOT19) 1975441T6 (PANCYTUT02) 1975441F6 (PANCYTUT02) 70760953V1 70761097V1 70757930V1 60205344U1 6630210U1 1858593T6 (PROSNOT18) 1706443T6 (DUODNOT02) 1858593F6 (PROSNOT18) 7062677H1 (PENITMNO2) 1330249H1 (EOSINOT01) 5407812F6 (BRAINNOT01) 71427502V1 6886749J1 (BRAHIDR03) 3627391F6 (COLANNOT38) 71430251V1 1979283R6 (LUNGYTUT03)	2946 2128 218 2348 1025 1 3133 3747 3664 426 821 1651 1374 438 1044 906 686 1 1342 429 1299 632 1 1047 997	3643 2758 977 3030 1574 287 3718 4451 4246 1008 1416 2163 2060 905 1521 1513 989 468 1568 985 1887 1022 531 1612 1556
68	1706443CB1	1568	1-43			
69	1748627CB1	1887	1-649			

【表 18】

10

20

30

表 4-7

ポリスクレオチド SEQ ID NO:	Incyte リスクレオチドID	ボ 配列長	選択された断片	配列断片	5'位置	3'位置
70	1818332CB1	569	1-35	1255779F2 (MENYTUT03) 1336021T1 (COLNNOT13)	39 1	569 541
71	1822832CB1	2338	529-565, 1332-1369, 1-124, 1488-2338	1289709F6 (BRAINTUT11)	1977	2338
				1822832X352U1 (GBLATUT01)	25	655
				91975312	1	277
				SAOA01720F1	710	1312
				SAOA01416F1	1318	1894
				1452843F6 (PENITUT01)	520	1212
				SAOA01295F1	1707	2338
				SAOA00837F1	1223	1841
72	1832219CB1	481	1-21	SXAF02203V1	1	479
				1832219R6 (BRAINON01)	44	481
73	1899010CB1	1255	1-62	1899010F6 (BLADTUT06) 2174773F6 (ENDCHOT03)	341 863	824 1255
				1909527T6 (CONNTUT01)	569	1233
				1425473H1 (BEPINON01)	1	264
				1909527F6 (CONNTUT01)	34	643
74	2008768CB1	875	1-411	2008768T6 (TESTNOT03) 6025379H1 (TESTNOT11)	159 1	858 261
				563323R6 (NEUTLPT01)	472	875
75	2070984CB1	2188	1-72, 1579-1656	1273987F1 (TESTTUT02)	1769	2188
				7162982H1 (PLACNOR01)	321	914
				SBA08036D1	1307	1793
				SBA01466D1	891	1516
				6907510J1 (PITUDIR01)	112	851

10

20

30

表 4 - 8

ポリスクレオチド SEQ ID NO:	Incyte リヌクレオチド ID	ボ 配列長	選択された断片	配列断片	5'位置	3'位置
76	2193240CB1	1561	1-624	3320716H1 (PROSPT03)	1	280
				1624251F6 (BRAITUT13)	1	472
				2429918P6 (MENTUN02)	734	1174
				2429918T6 (MENTUN02)	1003	1561
				900981R1 (BRSTTUT03)	365	926
77	2235177CB1	1777	1-32	71113502V1	1140	1777
				6993448H1 (BRAQTDRO2)	1	725
				71264559V1	837	1353
				71113614V1	679	1308
78	2416227CB1	1941	1-482	6821668J1 (SINTNOR01)	194	980
				2416227T6 (HNT3AZT01)	1210	1789
				2416227F6 (HNT3AZT01)	976	1434
				854765H1 (NGANNOT01)	774	1005
				7017947H1 (KIDNNOC01)	1	604
79	2461076CB1	1616	835-861, 565-783, 1-219	7154302H1 (HEARNOC03)	1446	1841
				7039965H1 (UTRSTMR02)	110	673
				219625R6 (STOMNOT01)	948	1492
				6935657H1 (SINTTMR02)	855	1446
				219625T6 (STOMNOT01)	958	1616
80	1957517CB1	1434	1-111	2461076F6 (THYRNOT08)	484	938
				6073858H1 (UTREDIT09)	1	273
				7161288H1 (PLACNOR01)	448	1014
				1233279F1 (LUNGRET03)	1	537
				6573238H1 (COLRTUS02)	655	1348
81	866038CB1	2085	51-124	1575944H1 (LNODNOT03)	1214	1434
				6913288J1 (PITUDIR01)	272	823
				5964475H1 (BRAATNOT05)	1744	2085
				7132506H1 (BRAHTDK01)	816	1499

【表 2 0】

表 4 - 9

ポリスクレオチド Seq ID No:	Incyte リスクレオチドID	ポ 配列長	選択された断片	配列断片	5'位置	3'位置
				6054811H1 (BRAENOT04)	784	1447
				755563R1 (BRAITUT02)	215	753
				5483139H1 (FIBPFEN06)	1	285
				6438276H1 (BRAENOT02)	1420	2083
82	3869704CB1	904	1-36	705156H1 (SYNORAT04)	1	232
				71052048V1	331	904
				3901129R8 (LUNGNON03)	48	697
				3904253R9 (LUNGNON03)	217	837
83	1415179CB1	1496	1-248, 606-836	2042611R6 (HIPONON02)	729	1206
				660950R6 (BRAINOT03)	557	1155
				4713560H1 (BRAIHCT01)	1	252
				658639F1 (BRAINOT03)	839	1496
				2708523H1 (PONSATZT01)	430	732
				2967826F6 (SCORNOT04)	58	686
84	1664792CB1	2837	1-1559	70858742V1	407	959
				4797546H1 (LIVRTUT09)	1248	1524
				2699003T6 (OVARUTUT10)	2154	2825
				71224728V1	867	1504
				2542259F6 (BONRTUT01)	1859	2386
				7460645H1 (LIVRTUT01)	441	1062
				2260182R6 (UTRSNOT02)	2545	2837
				1664792T6 (BRSTNOT09)	1698	2381
				6988160H1 (BRAIFER05)	1	444
				1664792F6 (BRSTNOT09)	1257	1834
				4179737H1 (SINITUT03)	1589	1863
85	2079396CB1	1123	1-45, 993-1123	6820736H1 (SINTNOR01)	566	1063
				91401473	1	507
				874769R1 (LUNGAST01)	786	1123

【表 2 1】

表 4-10

ボリスクレオチド SEQ ID NO:	Incyte リスクレオチド ID	ボ	配列長	選択された断片	配列断片	S'位置	3'位置
85					6819702J1 (OVARDIR01)	54	794
86	5390115CB1	1549	1-270		6335288H1 (BRANDIN01) 1258145F1 (MENITUT03) 1466677F1 (PANCUTUT02) 4250616F6 (BRADDIR01) 1310308F1 (COLNFET02)	17 589 1086 1 758	509 1247 1549 633 1337
87	1403326CB1	4820	1-3502		1306452F6 (PLACNOT02) 70607520V1 4322557H1 (TLYMUNT01) 3080429F6 (BRAIUNT01) 70476331V1 70604815V1 2801448F6 (PENCNOT01) 4729710H1 (GBLADIT01) 6245574H1 (TESTNOT17) 70815905V1 5718724H1 (PANCNOT16) 6863174H1 (BRAGNON02) 6937903H1 (FTUBTUR01) 6489460H1 (MIXDUNB01) 4884473H2 (LUNLTMT01) 2820527T6 (BRSTNOT14) 5642645R8 (UTRSTMR01) 1251961F1 (LUNGFET03)	1069 291 2789 2387 1812 1 3054 1585 3553 417 1188 4389 3745 2100 2991 4094 603 2550	1588 789 3045 3023 2459 395 3590 1844 4124 793 1815 4820 4307 2703 3242 4505 1080 3112
88	7690129CB1	3599	1878-1968, 1-934, 2349-3111		1851125T6 (LUNGFET03) 7757184J1 (SPLNTUE01) 6800356J1 (COLENOR03) 6831480J1 (SINTNOR01)	2986 706 222 1821	3599 1447 920 2515

【表 2 2】

表 4-11

ボリスクレオチド Seq ID No:	Incyte リスクレオチド ID	ボ	配列長	選択された断片	配列断片	5'位置	3'位置
					6883161H1 (BRAHTR03)	580	1029
					5868845F8 (COLTDT04)	1644	2425
					71137279V1	2150	2848
					2185757F6 (PROSNOT26)	1	495
					7612578J1 (KIDCTME01)	1041	1732

【表 2 3】

10

20

30

【表 2 4】

表 5-1

ポリヌクレオチド SEQ ID NO:	lncyte プロジェクトID	代表的ライブラリ
45	210168CB1	BRAITUT02
46	5452330CB1	BRAIDIT01
47	4362432CB1	SKIRNOT01
48	5308104CB1	BRAYDIN03
49	3092736CB1	BRAITUT08
50	3580257CB1	293TF3T01
51	3634758CB1	HUVENOB01
52	4027923CB1	COLNNOT16
53	4348533CB1	LIVRON08
54	4521857CB1	SPLNNOT04
55	4722253CB1	TESTNOT03
56	4878134CB1	LUNGNON03
57	5050133CB1	FIBPFEN06
58	5630124CB1	LUNGNOT09
59	5677286CB1	PROSTUT12
60	6436791CB1	MEGBUNT01
61	1820972CB1	SPLNNOT04
62	3286805CB1	SKINDIA01
63	3506590CB1	COLDDIE01
64	003600CB1	HMC1NOT01
65	1251534CB1	THYMNOT05
66	1402311CB1	CARCINT02
67	1623474CB1	HMC1NOT01
68	1706443CB1	DUODNOT02
69	1748627CB1	FIBPFEN06
70	1818332CB1	ISL1NOT01
71	1822832CB1	BRAINOT11
72	1832219CB1	TESTNOT03
73	1899010CB1	BLADTUT06
74	2008768CB1	TESTNOT03
75	2070984CB1	PLACNOT07

表 5-2

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
76	2193240CB1	BRAITUT13
77	2235177CB1	HNT2AGT01
78	2416227CB1	LUNGNOT09
79	2461076CB1	STOMNOT01
80	1957517CB1	OVARTUT01
81	856038CB1	BRAITUT03
82	3869704CB1	LUNGNOT03
83	1415179CB1	BRAINOT03
84	1664792CB1	BRSTTUT01
85	2079396CB1	CONUTUT01
86	5390115CB1	BRAITUT03
87	1403326CB1	BRSTNOT01
88	7690129CB1	PROSTUT12

【表 2 5】

表 6-1

ライブラリ	ベクター	ライブラリの説明
293TF3T01	pINCY	ライブラリは腎臓上皮組織由来の血清創設させて形質転換した胚性細胞系(293-BEN)から単離した RNA を用いて作製した。この細胞はアデノウイルス DNA で形質転換した。
BLADTUT06	pINCY	ライブラリは、根治的膀胱切除、及び胃瘻造設術時の 58 才白人男性から取り除かれた膀胱癌組織から単離された RNA を用いて作製された。病理は左膀胱癌にグレード 3 の移行性細胞癌が見られた。残りの膀胱には顕著な膀胱炎があり、生体内原位腫に散在する微小病巣の移行性細胞癌を伴っていた。患者の病歴にはアンギナ、気腫および喫煙がある。家族歴には急性心筋梗塞、アテローム硬化性動脈疾患および II 型糖尿病がある。
BRAIDT01	pINCY	ライブラリは膀胱癌組織から単離された RNA を用いて作製した。患者の病歴には、多発性硬化症、タイプ II 病変があった。
BRAINOT03	PSPORT1	ライブラリは、26 才の白人男性の頭蓋形成術および脳腫瘍病変の切除時に、摘出された脳組織から単離した RNA を用いて作製した。関連腫瘍組織の病理は脳の右前頭頂部に於けるグレード 4 の少星細胞腫を示した。
BRAINOT11	pINCY	ライブラリは、5 才の白人男子の大半半球切除術時に右側頭葉から摘出された脳組織から単離された RNA を用いて作製された。病理は、慢性的脳腫瘍発作に一致する広範囲の多小脳回症と軽度から中等度の神経膠症(主に、軟膜下、皮質下)を示した。家族歴には子宮頸癌が含まれる。
BRAITUT02	PSPORT1	ライブラリは、58 才の白人男性の脳腫瘍病変の切除時、前頭葉から採取した脳腫瘍組織から単離した RNA を用いて作製した。病理学検査では、グレード 2 の転移性副腎腫が見られた。患者の病歴は、グレード 2 の腎細胞癌、不眠症および慢性的気道閉塞などであった。家族歴には腎臓の悪性腫瘍が含まれている。
BRAITUT03	PSPORT1	ライブラリは、17 才の白人女性の脳腫瘍病変の切除時に、左前頭葉から採取した脳腫瘍組織から単離した RNA を用いて作製した。病理はグレード 4 の繊維状の巨大および小細胞の星状細胞腫を示した。家族歴には良性高血圧および心臓血管系疾患および脳血管系疾患がある。
BRAITUT08	pINCY	ライブラリは、47 才の白人男性の脳腫瘍病変の切除時、左前頭葉から採取した脳腫瘍組織から単離した RNA を用いて作製した。病理学検査では、焦点腫瘍性放射線壊死による等級 4 の線維性星状細胞腫が見られた。患者の病歴には、脳血管障害、欠乏性貧血、高脂血症、脳腫および喫煙がある。家族の病歴には、脳血管障害、悪性前立腺腫瘍がある。
BRAITUT13	pINCY	ライブラリは、68 才の白人男性の脳腫瘍病変の切除時、左前頭葉から採取した脳腫瘍組織から単離した RNA を用いて作製した。病理は左前頭葉に髄膜腫を示した。
BRAYDIN03	pINCY	標準化したライブラリは、脳組織ライブラリの 670 万個の独立性クローンから作製した。開始 RNA は脳血管事故から死亡した 57 才の白人男子から採取された脳脊髄液と脳組織から単離した RNA から作製した。患者の病歴にはハンチントン舞蹈病および肺気腫がある。このライブラリは、極めて長時間(48 時間/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ことを除いては、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 を改作した条件を用いて 2 回にわたり標準化した。このライブラリはクローンを含む挿入断片用に選択するために直線化し、再度標準化した。
BRSTNCT01	PBLUESCRIPT	ライブラリは自動事故で死亡した 56 才の白人女性から採取した乳房組織から単離した RNA を用いて作製した。

10

20

30

【表 26】

表 6-2

ライブラリ	ベクター	ライブラリの説明
BRSTT01	PSORT1	ライブラリは 56 歳の白人女性の片側性乳房単純乳房切除術時に採取した乳房腫瘍組織から単離した RNA を用いて作製した。病理は侵襲度 4 の左乳房を大きく含む、小葉状、管状、管状を混合した乳腺癌を示した。腫瘍は乳管近くの深い真皮に見られ、関節外への伸展を伴っていた。7 箇所の中部位および低部位、また 5 箇所の高部位の腫瘍リンパ節に陽性を示した。増殖性細胞性変化はアポクリン異形成、硬化腺癌、癌形成および非異型性管過形成を特徴としていた。患者の病歴には心房性頻脈、血便、および良性乳房腫瘍が含まれていた。家族歴には、良性高血圧症、アテローム硬化型冠動脈疾患、脳血管疾患および抑鬱性疾患が含まれる。
CARCT02	PSORT1	ライブラリは全膝関節置換術時に得られたブールされた膝軟骨から単離された軟骨細胞の RNA を用いて作製した。軟骨は基礎の骨から外し、小片に切り刻み、5 ng/ml の IL-1 で 18 時間刺激された。
COLDDIE01	PCDNA2.1	この 5 プライムに属するランダムプライムされたライブラリは、28 歳白人男性の腹腔内全結腸切除および一時的回腸人工肛門時に採取された盲腸下結腸組織から単離した RNA を用いて作製した。病理は、中等度から高度の悪性増殖性大腸炎を示し、大腸の末端 23cm を含む進行性である。直腸 S 状部、直腸および直腸組織の 24cm 部分全体が重症進行性の慢性潰瘍性大腸炎に覆っていた。患者は血便、下痢および女子性貧血が見られた。患者の病歴は肝の筋緊張異常(回旋筋腹板の捻転)およびニコチン中毒が含まれる。患者は輸血にて治療された。患者の服用薬剤には、Asacol、プレドニゾンおよびコーナゾン錠が含まれる。家族歴には急性心筋梗塞、上葉肺癌、大腸癌、および祖父母に I 型糖尿病がある。
COLNNOT16	pINCY	ライブラリは 62 歳の白人男性の S 状結腸切除術と永久的人工肛門形成術時に採取した S 状結腸組織から単離した RNA を用いて作製した。
CONUT01	pINCY	ライブラリは 61 歳の女性の腹式子宮全摘出術および局部的リンパ節切除を伴った両側生卵巣卵巣摘出術時に得た S 字結腸腸間膜腫瘍組織から単離した RNA を用いて作製した。病理は S 状結腸腸間膜の 2 つの部位に転移性の等級 4 の悪性ミューラー管混合腫瘍の存在が示された。
DOODNOT02	pINCY	ライブラリは、脳外傷で死亡した 8 才白人女子の十二指腸組織から単離された RNA を用いて作製された。血清学はサイトメガロウイルス(CMV)に陽性であった。
FIBPFEN06	pINCY	標準化された前立腺間質性線維芽細胞ライブラリは前立腺線維芽細胞ライブラリからの 155 万個の独立性クローンから作製した。開始 RNA は妊娠 26 週間後死亡した男子胎児から採取した前立腺間質の線維芽細胞から作製した。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て除いては、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 を改作した条件を用いて 2 回にわたって標準化した。このライブラリは次のようにして、クローンを含む挿入断片用に選択するために直線化し、再度環状にした。プラスミド DNA は軟寒大形質転換後に標準化した前立腺間質性線維芽細胞組織ライブラリからの約 100 万個のクローンから調製した。
HMC1NOT01	PBLUESCRIPT	ライブラリは 52 歳の女性から由来する HMC-1 ヒト肥満細胞株から単離した RNA を用いて作製した。患者の病歴には肥満細胞白血病が含まれる。

【表 27】

表 6-3

ライブラリの説明		ライブラリの説明	
ライブラリ	ベクター	ライブラリ	ライブラリは、hNT2 細胞株(神経前駆細胞の特性を示したヒトの奇形癌から由来する)から単離した RNA を使用して Stratagene (STR937233)にて作製した。細胞はレチノイン酸で5週間、有糸分裂阻害剤で2週間処理して、さらに馴化培地で4週間にわたり成熟させた。
	PBLUESCRIPT		ライブラリは HUV-EC-C (ATCC CRL 1730)細胞から単離した RNA を用いて作製した。
HUVENOB01	PBLUESCRIPT	ライブラリ	ライブラリは、肺組織ライブラリの256万個の独立性クロンから作製した。RNAは58才白人男性の標準化したライブラリは、プールされた肝臓組織ライブラリから570万個の独立性クロンから作製した。開始 RNAは無酸素症で死亡した4才のラテンアメリカ系男子および無酸素症で妊娠16週で死亡した16週女子胎児から抽出され、プールされた肝臓組織から作成した。血清学は四才男子においては細胞巨ウイルス(cytomegalovirus)が陽性であった。四才児患者の病歴は喘息がある。家族歴には胎児の母において出産前用ビタミンの日常的服用および膈疝片脱出がある。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ことを除いては、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research 6(1996):791 を適応した条件を用いて2回にわたり標準化した。
	PINCY		この標準化したライブラリは、肺組織ライブラリの256万個の独立性クロンから作製した。RNAは58才白人男性の部分的肺切除術時に左葉から抽出した肺組織から作製した。関連する腫瘍組織の病理は転移性グレード3(4のうち)の骨肉腫を示した。患者の病歴は、二次性肺病、前立腺癌および急性出血性十二指腸潰瘍を含む。患者はまた、後腹膜腔に放射線療法を受けている。家族歴には前立腺癌、乳癌および急性白血病がある。標準化およびハイブリダイゼーションの条件は Soares 他, PNAS (1994) 91:9228; Swaroop 他, NAR (1991) 19:1954; および Bonaldo 他, Genome Research (1996) 6:791 から採用した。
LUNGNOT03	ESPORT1	ライブラリ	ライブラリは79才の白人男子の肺組織から単離した RNA を用いて作製した。関連する腫瘍組織の病理学検査では、グレード4の癌が見られた。患者の病歴には良性的前立腺腫瘍およびアテローム性動脈硬化症が含まれた。
	PINCY		ライブラリは妊娠23週間の白人男子胎児の肺組織から単離した RNA を用いて作製した。妊娠は小児型多嚢胎腎を超音波にて診断後中絶された。
MEGHUNT01	PINCY	ライブラリ	ライブラリは、慢性骨髄性白血病の巨核芽球発症の55才男性から得た骨髓細胞から由来する未処理の MEG-01 巨核芽球細胞から単離した RNA を用いて作製した。
	ESPORT1		ライブラリは、卵管と卵巣を切除した43歳の白人女性から採取した卵巣腫瘍組織の分離 RNA を使用して構築された。病理学検査では、左卵巣全体を含むグレード2のムチン性嚢胞性腫瘍が見られた。患者の病歴には、側副弁疾患、肺炎およびウィルスの肝臓炎があった。家族の病歴には、アテローム硬化性冠動脈疾患、糖尿病、高血圧、脳血管疾患、乳癌および子宮癌が見られた。
PLACNOT07	PINCY	ライブラリ	ライブラリは胎児性死亡および水頭症で妊娠16週目に死亡した白人胎児から採取した胎盤組織から単離した RNA を用いて作製した。血清学的検査は抗 CMV (サイトメガロウイルス)に陽性であった。
	PINCY		ライブラリは、前立腺全摘出新時に65才白人男性から取り除かれた前立腺腫瘍組織から単離された RNA を用いて作製

【表 28】

表 6－4

ライブラリ	ベクター	ライブラリの説明
		ライブラリは腺癌(グリソングレード2+2)を示した。腺癌組織は過形成もまた見られた。この患者は前立腺特異 された。病理には腺癌(PSA)の上昇も示した。
SKINDIA01	FSPORT1	抗腫瘍剤ライブラリは皮膚生検時に女性1人、男性4人から採取された癌組織から単離したRNAを用いて この増幅されたライブラリは皮膚生検時に女性1人、男性4人から採取された癌組織から単離したRNAを用いて 作製した。病理は癌組織から単離したRNAを用いて作製した。
SKIRNOT01	P1NCY	ライブラリは26才の白人女性の両側乳房形成術時に採取した皮膚組織から単離したRNAを用いて作製した。
SPLNNOT04	P1NCY	ライブラリは、脳無酸素症で死亡したオラテンメリア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。 過去の病歴および血清検査は陰性であった。
STOMNOT01	PBLUESCRIPT	ライブラリは、心臓停止で死亡した55才白人男子の胃組織から単離されたRNAを用いて作製された。
TESTNOT03	PBLUESCRIPT	ライブラリは肝臓で死亡した37才の白人男子から単離した精巣組織から単離したRNAを用いて作製した。患者の病歴 には、硬変、黄変および肝不全が含まれる。
THYMNOT05	P1NCY	ライブラリは3才のラテンメリア系男子の胸腺腫瘍および動脈管閉存の閉鎖時に採取した胸腺組織から単離したRNA を用いて作製した。患者の病歴には心臓カテーテル法および心エコーが含 まれる。過去の手術にはプレロック・トーション短絡術および肺弁閉鎖術が含まれる。患者が服用している薬剤はなか った。家族歴には良性高血圧、骨関節炎、抑鬱性疾患、および祖父母に外因性喘息が含まれる。

【表 2 9】

【表 3 0】

表 7 - 1

プログラム	説明	引用文献	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、特定の塩基をマスクするプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finderは、アミノ酸または核酸配列の比較、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はアミノ酸および核酸配列の配列類似性検索に有用であり、BLAST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には最小5つの機能(fasta、tfasta、tfastaおよびsearch)がある。	Pearson, W.R.およびD.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; Smith, T.F. および M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E 値=1.0E-6 構築されたESTs: fasta J 同 一性=95%以上 一致した長さ=200塩基以上 fasta F 値=1.0E-8 以下 完全長配列: fasta スコア=100以上
BLIMPS	配列をBLOCKS、PRINIS、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グアブプリント領域を検索するBLocks IMPROVED サーチャャー。	Henikoff, S. および J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. および S. Henikoff (1996) Methods Enzymol. 266:88-105; Attwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	確率値= 1.0E-3以下
HMMER	PFAMのようなタンパク質ファミリーコンセンサス配列の隠れマルコフモデル(HMM)に基づいたデータベースに対して問合せ配列を検索するアルゴリズム。	Krogh, A. 他(1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. 他 (1988) Nucleic Acids Res. 26:320-322; Durbin, R. 他 (1998) Our World View. in a Nutsheil, 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

30

20

10

表 7-2

プログラム	説明	引用文献	パラメータ閾値
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.	標準化された質の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のトレースを調べる塩基読み出しアルゴリズム。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的なインプリメンテーションに基づくSWATやCrossMatchを含む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. および M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T. F. および M. S. Waterman (1981) J. Mol. Biol. 147:195-197; Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=56以上
Conscd	A graphical tool for viewing and editing Phrap assemblies.	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付けマトリックス解析プログラム。	Nelson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J.M. および S. Audic (1997) CABIOS 12:431-439.	スコア=3.5以上
TMAP	加重マトリックスを用いて蛋白配列での膜貫通セグメントの描写および配向を決定するプログラム。	Persson, B. および P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. および P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	隠れMarkov モデル(HMM)を使って蛋白配列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。	Sommadossi, 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頁	
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, M51-69頁, Genetics Computer Group, Madison, WI	

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
27 December 2001 (27.12.2001)

PCT

(10) International Publication Number
WO 01/98353 A2

(51) International Patent Classification: C07K 14/705

(21) International Application Number: PCT/US01/19862

(22) International Filing Date: 20 June 2001 (20.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/212,890	20 June 2000 (20.06.2000)	US
60/213,166	23 June 2000 (23.06.2000)	US
60/214,601	27 June 2000 (27.06.2000)	US
60/222,372	31 July 2000 (31.07.2000)	US
60/231,155	8 September 2000 (08.09.2000)	US
60/232,889	15 September 2000 (15.09.2000)	US

94552 (US), **BATRA**, Sajeev [US/US], 355 El Camino Real #709, Sunnyvale, CA 94087 (US), **BURFORD**, Neil [GB/US], 105 Wildwood Circle, Durham, CT 06422 (US), **YAO**, Monique, G. [US/US], 111 Frederick Court, Mountain View, CA 94035 (US), **NGUYEN**, Daniel, R. [US/US], 1403 Ridgewood Drive, San Jose, CA 95118 (US), **LI**, Dyong, Anna, M. [US/US], 233 Coy Drive, San Jose, CA 95128 (US), **WALLA**, Narinder, K. [US/US], 890 Davis Street #205, San Leandro, CA 94577 (US), **GANDHI**, Anceena, R. [US/US], 837 Roble Avenue #1, Menlo Park, CA 94025 (US), **AU-YOUNG**, Janice [US/US], 233 Golden Eagle Lane, Brisbane, CA 94005 (US), **PATTERSON**, Chandra [US/US], 490 Sherwood Way #1, Menlo Park, CA 94025 (US).

(74) Agents: HANLEY-COX, Diana et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (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): **HOLLMAN**, Jennifer, L. [US/US], 230 Monroe Drive #17, Mountain View, CA 94040 (US), **TANG**, Y., Tom [US/US], 4230 Rancick Court, San Jose, CA 95119 (US), **YTF**, Henry [US/US], 826 Lois Avenue, Sunnyvale, CA 94087 (US), **ELLIOTT**, Vicki, S. [US/US], 3770 Polon Place Way, San Jose, CA 95121 (US), **TRIBOULEY**, Catherine, M. [FR/US], 1121 Tennessee Street #5, San Francisco, CA 94107 (US), **LEE**, Ernestine, A. [US/US], 621 Kains Street, Albany, CA 94706 (US), **RAMKUMAR**, Jayalaxmi [IN/US], 34359 Maybird Circle, Fremont, CA 94555 (US), **LAL**, Preeti [IN/US], P.O. Box 5142, Santa Clara, CA 95056 (US), **XU**, Yuning [US/US], 1739 Walnut Drive, Mountain View, CA 94040 (US), **WARREN**, Bridget, A. [US/US], 10130 Parwood Drive #2, Cupertino, CA 95014 (US), **HAFILLA**, April, J., A. [US/US], 2227 Calle de Primavera, Santa Clara, CA 95051 (US), **BAUGUIN**, Mariah, R. [US/US], 11214 Santiago Road, San Leandro, CA 94577 (US), **AZIMZAI**, Yalda [US/US], 5518 Boulder Canyon Drive, Castro Valley, CA

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

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

Published:

without international search report and to be republished upon receipt of that report
— with sequence listing part of description published separately in electronic form and available upon request from the International Bureau

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 01/98353 A2

WO 01/58353

PCT/US01/19862

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, antineoplastic/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, vasomodulators, 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 "shotgun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various

WO 01/58353

PCT/US01/19862

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

5 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

10 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

15 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 spondin B; and cell adhesion receptors such as cell adhesion molecules (CAMs), cadherins, and integrins. (Reviewed

20 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,

25 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

30 (Theopold, U. et al. (1996) *J. Biol. Chem.* 271:12708-12715).

Tuftelins are one of four different enamel matrix proteins that have been identified so far. The other three known enamel matrix proteins are the amelogenins, enamelin and ameloblastin. Assembly of the enamel extracellular matrix from these component proteins is believed to be critical in producing a matrix competent to undergo mineral replacement. (Paine C.T. et al. (1998) *Connect*

35 *Tissue Res.*38:257-267). Tuftelin mRNA has been found to be expressed in human ameloblastoma

WO 01/58353

PCT/US01/19862

tumor, a non-mineralized odontogenic tumor (Deutsch D. et al. (1998) Connect Tissue Res. 39:177-184).

Olfactomedin-related proteins are extracellular matrix, secreted glycoproteins with conserved C-terminal motifs. They are expressed in a wide variety of tissues and in broad range of species, from *Caenorhabditis elegans* to *Homo sapiens*. Olfactomedin-related proteins comprise a gene family with 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 (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.H. et al., (1994) J. Neurosci. Res. 38:468-478).

Mac-2 binding protein is a 90-kD serum protein (90K) and another secreted glycoprotein, isolated from both the human breast carcinoma cell line SK-BR-3, and human breast milk. It 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 (Kotbs, K. et 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 members and are found in vertebrates, invertebrates, and even certain viruses. All semaphorins contain the sema domain which is approximately 500 amino acids in length. Neuropilin, a semaphorin receptor has been shown to promote neurite outgrowth *in vitro*. The extracellular region of neuropilins consists of three different domains: CUB, discoidin, and MAM domains. The CUB and the MAM motifs of neuropilin have been suggested as having roles in protein-protein interactions and are suggested to be involved in the binding of semaphorins through the sema and the C-terminal domains (reviewed in Raper, J.A. (2000) Curr. Opin. Neurobiol. 10:83-94). Plexins are neuronal cell surface molecules that mediate cell adhesion via a homophilic binding mechanism in the

WO 01/58353

PCT/US01/19862

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
 5 ancestors of classical semaphorins and binding partners for semaphorins (Winberg, M.J. 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
 10 (Plouzek C.A. and Chon J.Y., *Endocrinology* 129:950-958) Different subpopulations of PSG have been found to be produced by the trophoblasts of the human placenta, and the amnionic, and chorionic membranes (Plouzek C.A. et al. (1993) *Placenta* 14:277-285).

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.
 15 Autocrine motility factor receptor (AMFR) expression has been found to be associated with tumor progression in thymoma (Ohta Y. et al. (2000) *Int. J. Oncol.* 17:259-264). AMFR is a cell surface glycoprotein of molecular weight 78KDa.

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
 20 compositions and mechanisms of action, hormones can be grouped into two categories. One category includes small lipophilic hormones that diffuse through the plasma membrane of target cells, bind to cytosolic or nuclear receptors, and form a complex that alters gene expression. Examples of these molecules include retinoic acid, thyroxine, and the cholesterol-derived steroid hormones such as progesterone, estrogen, testosterone, cortisol, and aldosterone. The second category includes
 25 hydrophilic hormones that function by binding to cell surface receptors that transduce signals across the plasma membrane. Examples of such hormones include amino acid derivatives such as catecholamines (epinephrine, norepinephrine) and histamine, and peptide hormones such as glucagon, insulin, gastrin, secretin, cholecystokinin, adrenocorticotropic hormone, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, and vasopressin. (See, for example, Lodish et al.
 30 (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 hormone includes smaller peptides with distinct biological activities: alpha-melanotropin (alpha-
 35 MSH) and corticotropin-like intermediate lobe peptide (CLIP) are formed from ACTH; gamma-

WO 01/58353

PCT/US01/19862

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-onset obesity, adrenal insufficiency, and red hair pigmentation (Chretien, M. et al., (1979) *Canad. J. Biochem.* 57:1111-1121, Krato, H. et al., (1998) *Nature Genet.* 19:155-157, Online Mendelian Inheritance in Man, OMIM. Johns Hopkins University, Baltimore, MD. OMIM Number: 176830; August 1, 2000. World Wide Web URL: www.ncbi.nlm.nih.gov/omim/).

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

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

The Slit protein, first identified in *Drosophila*, is critical in central nervous system midline formation and potentially in nervous tissue histogenesis and axonal pathfinding. Itoh et al. have identified mammalian homologues of the slit gene (human Slit-1, Slit-2, Slit-3 and rat Slit-1). The encoded proteins are putative secreted proteins containing L1-G-like motifs and leucine-rich repeats, both are conserved protein-protein interaction domains. Slit-1, -2, and -3 mRNAs are expressed in the brain, spinal cord, and thyroid, respectively (Itoh, A. et al., (1998) *Brain Res. Mol. Brain Res.*

WO 01/58353

PCT/US01/19862

62:175-186). The Slit family of proteins are indicated to be functional ligands of glypican-1 in nervous tissue and suggests that their interactions may be critical in certain stages during central nervous system histogenesis (Liang, Y. et al., (1999) J. Biol. Chem. 274:17885-17892).

Neuropeptides and vasomediators (NP/VM) comprise a large family of endogenous signaling molecules. Included in this family are neuropeptides and neuropeptide hormones such as bombesin, neuropeptide Y, neurotensin, neuromedin N, melanocortins, opioids, galanin, somatostatin, tachykinins, urotensin II and related peptides involved in smooth muscle stimulation, vasopressin, vasoactive intestinal peptide, and circulatory system-borne signaling molecules such as angiotensin, complement, calcitonin, endothelins, formyl-methionyl peptides, glucagon, cholecystokinin and gastrin. NP/VMs can transduce signals directly, modulate the activity or release of other neurotransmitters and hormones, and act as catalytic enzymes in cascades. The effects of NP/VMs range from extremely brief to long-lasting. (Reviewed in Martin, C.R. et al. (1985) Endocrine Physiology, Oxford University Press, New York, NY, pp. 57-62.)

NP/VMs are involved in numerous neurological and cardiovascular disorders. For example, neuropeptide Y is involved in hypertension, congestive heart failure, affective disorders, and appetite regulation. Somatostatin inhibits secretion of growth hormone and prolactin in the anterior pituitary, as well as inhibiting secretion in intestine, pancreatic acinar cells, and pancreatic beta-cells. A reduction in somatostatin levels has been reported in Alzheimer's disease and Parkinson's disease. Vasopressin acts in the kidney to increase water and sodium absorption, and in higher concentrations stimulates contraction of vascular smooth muscle, platelet activation, and glycogen breakdown in the liver. Vasopressin and its analogues are used clinically to treat diabetes insipidus. Endothelin and angiotensin are involved in hypertension, and drugs, such as captopril, which reduce plasma levels of angiotensin, are used to reduce blood pressure (Watson, S. and S. Arkinwall (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 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. 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.

WO 01/58353

PCT/US01/19862

- (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 (H. Toh (1991) *Protein Seq. Data Anal.* 4:111-117 and Iwai, N. et al., (1994) *Hypertension* 23:375-380).
- 10 Other proteins that contain signal peptides include enzymes involved in the glycosylation of proteins in transit through the secretory pathway. Mucin-type O-linked glycosylation is a dominant form of protein glycosylation. Initiation of mucin-type glycosylation occurs by the addition of the monosaccharide N-acetylgalactosamine to the hydroxyl group of serine and threonine amino acids (GalNAc-1-O-Ser/Thr). GalNAc O-glycosylation is more prominent on high molecular weight
- 15 secretory glycoproteins such as mucins, but is also found on a variety of glycoproteins (White, T. et. al., *J. Biol. Chem.* (1995) 270:24156-24165). Additionally, serine/threonine-rich tandem repeats are a characteristic of human mucin core proteins. The tandem repeat region also contains numerous antigenic determinants as recognized by the monoclonal antibodies HMFG-1, HMFG-1, and SM-3. Glycosylation sites within the tandem repeat region were found to be differentially glycosylated
- 20 depending on the organ from which Muc1 was isolated. The finding of variable glycosylation activity may be critical to further understanding of the molecular basis of cancer-associated epitopes which map to the Muc1 tandem repeat (Gendler, S.J. et al. (1990) *J. Biol. Chem.* 265:15286-15293).
- 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
- 25 parasitic infections. A key feature of the immune system is its ability to distinguish foreign molecules, 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
- 30 comprised of antiparallel β sheets joined by a disulfide bond in an arrangement called the Ig fold. Members of the Ig superfamily include T-cell receptors, major histocompatibility (MHC) proteins, antibodies, and immune cell-specific surface markers such as the "cluster of differentiation" or CD antigens. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shotgun" techniques. These techniques have resulted in the production of hundreds of mAbs
- 35 directed against unknown cell surface leukocytic antigens. These antigens have been grouped into

WO 01/58353

PCT/US01/19862

"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.)

- 10 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 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.
- 15 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.
- 20 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, or immunoglobulins, 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.
- 30

- 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
- 35

WO 01/58353

PCT/US01/19862

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
 5 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, supra, pp. 1206-1213 and 1216-1217.)

Both H-chains and L-chains contain repeated Ig domains. For example, a typical H-chain contains four Ig domains, three of which occur within the constant region and one of which occurs
 10 within the variable region and contributes to the formation of 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
 15 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
 20 rate of somatic mutation within these segments further contribute to the generation of a diverse antibody population.

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
 25 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
 30 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
 35 activity. For example, in *Drosophila*, the CyP NinaA is required for correct localization of

WO 01/58353

PCT/US01/19862

rhodopsins, while a mammalian CyP (Cyp40) is part of the Hsp90/Hsc70 complex that binds steroid receptors. The mammalian CypA has been shown to bind the gag protein from human immunodeficiency virus 1 (HIV-1), an interaction that can be inhibited by cyclosporin. Since cyclosporin has potent anti-HIV-1 activity, CypA may play an essential function in HIV-1 replication.

- 5 Finally, Cyp40 has been shown to bind and inactivate the transcription factor c-Myb, an effect that is reversed by cyclosporin. This effect implicates CyPs in the regulation of transcription, transformation, and differentiation (Bergsma, D.J. et al (1991) J. Biol. Chem. 266:23204 - 23214; Hunter, T. (1998) Cell 92: 141-143; and Levenson, J.D. and Ness, S.A. (1998) Mol. Cell. 1:203-211).

- Gamma-carboxyglutamic acid (Gla) proteins rich in proline (PRGPs) are members of a family of vitamin K-dependent single-pass integral membrane proteins. These proteins are characterized by 10 of extracellular amino terminal domain of approximately 45 amino acids rich in Gla. The intracellular carboxyl terminal region contains one or two copies of the sequence PPXY, a motif present in a variety of proteins involved in such diverse cellular functions as signal transduction, cell cycle progression, and protein turnover (Kulman, J.D. et al., (2001) Proc. Natl. Acad. Sci. U.S.A. 98:1370-1375). The process of post-translational modification of glutamic residues to form Gla is Vitamin K-dependent carboxylation. Proteins which contain Gla include plasma proteins involved in blood coagulation. These proteins are prothrombin, proteins C, S, and Z, and coagulation factors VII, IX, and X. Osteocalcin (bone-Gla protein, BGP) and matrix Gla-protein (MGP) also contain Gla (Friedman, P.A., and C.T. Przysiecki (1987) Int. J. Biochem. 19:1-7; C. Vermeer (1990) Biochem. J. 266:625-636).

- The *Drosophila* sp. gene *crossveinless 2* is characterized as having a putative signal or transmembrane sequence, and a partial Von Willebrand Factor D domain similar to those domains known to regulate the formation of intramolecular and intermolecular bonds and five cysteine-rich domains, known to bind BMP-like (bone morphogenetic proteins) ligands. These features suggest 25 that *crossveinless 2* may act extracellularly or in the secretory pathway to directly potentiate ligand signaling and hence, involvement in the BMP-like signaling pathway known to play a role in vein specification (Conley, C.A. et al., (2000) Development 127:3947-3959). The dorsal-ventral patterning in both vertebrate and *Drosophila* embryos requires a conserved system of extracellular proteins to generate a positional informational gradient.

- 30 Another protein that contains a signal peptide is encoded by the seizure-related gene, SEZ-6, a brain specific cDNA whose expression is increased by the convulsant drug pentylenetetrazole. The SEZ-6 protein is expressed in the cerebrum and cerebellum. SEZ-6 contains five short consensus repeats (SCR, or sushi domains) and two CUB (complement C1r/s-like repeat) domains in addition to a signal peptide and a single transmembrane domain (Shimizu-Nishikawa, K. et al. (1995) Biochem. Biophys. Res. Commun. 216:382-389).

WO 01/58353

PCT/US01/19862

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," "SECP-25," "SECP-26," "SECP-27," "SECP-28," "SECP-29," "SECP-30," "SECP-31," "SECP-32," "SECP-33," "SECP-34," "SECP-35," "SECP-36," "SECP-37," "SECP-38," "SECP-39," "SECP-40," "SECP-41," "SECP-42," "SECP-43," and "SECP-44." 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:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-44.

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:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-44. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:45-88.

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

WO 01/58353

PCT/US01/19862

of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. 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:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44.

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:45-88, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

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

WO 01/58353

PCT/US01/19862

consisting of SEQ ID NO:45-88, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional SHCP, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino

WO 01/58353

PCT/US01/19862

acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional 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 sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. 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

WO 01/58353

PCT/US01/19862

activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44. 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 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 sequence selected from the group consisting of SEQ ID NO:45-88, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

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:45-88, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, 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 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:45-88, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting

WO 01/58353

PCT/US01/19862

of j)-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 toxicity of the test compound.

5

BRIEF DESCRIPTION OF THE TABLES

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

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability score for the match between each polypeptide and its GenBank homolog is also shown.

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

Table 4 lists the cDNA 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.

25

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.

35

WO 01/58353

PCT/US01/19862

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

DEFINITIONS

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

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

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

"Altered" nucleic acid sequences encoding SECP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as SECP or a polypeptide with at least one functional characteristic of SECP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding SECP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding SECP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent SECP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of SECP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged

WO 01/58353

PCT/US01/19862

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.

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

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

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of SECP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small 15 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.

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. 20 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, 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, 25 and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures 30 on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as 35 phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified

WO 01/58353

PCT/US01/19862

sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring

5 nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

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

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

15 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding SECP or fragments of SECP may be

20 employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated

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

30 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

35 as conservative amino acid substitutions.

WO 01/58353

PCT/US01/19862

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

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

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

The term "derivative" refers to a chemically modified polynucleotide or polypeptide.

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

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

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

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

WO 01/58353

PCT/US01/19862

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:45-88 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:45-88, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:45-88 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:45-88 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:45-88 and the region of SEQ ID NO:45-88 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-44 is encoded by a fragment of SEQ ID NO:45-88. A fragment of SEQ ID NO:1-44 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-44. For example, a fragment of SEQ ID NO:1-44 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-44. The precise length of a fragment of SEQ ID NO:1-44 and the region of SEQ ID NO:1-44 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

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

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

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

Percent identity between polynucleotide sequences may be determined using the default

WO 01/58353

PCT/US01/19862

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

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bj2.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

WO 01/58353

PCT/US01/19862

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, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

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

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

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

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

WO 01/98353

PCT/US01/19862

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

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

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

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

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

WO 01/58353

PCT/US01/19862

stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A

5 hybridization complex may be formed in solution (e.g., C₀t or R₀t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide
10 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.

15 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.

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

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

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

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
30 antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

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

WO 01/58353

PCT/US01/19862

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

5 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.

10 "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

15 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

20 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

25 example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that

30 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

35 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer

WO 01/58353

PCT/US01/19862

selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

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

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

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

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid,

WO 01/98353

PCT/US01/19862

amino acid, or antibody. Reporter molecules include radioisotopes; 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, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or
10 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
15 structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope 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" refers to the collective pattern of gene expression by a particular cell
30 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 sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based
35 on the type of host cell being transformed and may include, but is not limited to, bacteriophage or

WO 01/98353

PCT/US01/19862

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 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.

5 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 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
10 a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection,
15 transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

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

WO 01/58353

PCT/US01/19862

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.

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

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and

WO 01/58353

PCT/US01/19862

motifs. In particular, the locations of signal peptides (as indicated by "Signal_peptide" or "Signal_cleavage") are shown. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, tables 2 and 3 summarize the properties of each polypeptide of the invention, and these properties establish that the claimed polypeptides are secreted proteins. For example, SEQ ID NO:1 is 51% identical to human UDP-Ga1NAc:polypeptide N-acetylgalactosaminyltransferase (GenBank ID g971461) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.5e-141, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains a signal peptide and a transmembrane domain as determined by hidden Markov model (HMM)-based methods. (See Table 3.) Likewise, SPScan analysis also indicates the presence of an N-terminal signal peptide in SEQ ID NO:1. Taken together, the evidence shows that SEQ ID NO:1 is present in the secretory pathway as an N-acetylgalactosaminyl transferase.

For example, SEQ ID NO:2 is 90% identical to mouse seizure-related gene product 6 type 2 precursor (GenBank ID g1139548) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.9, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:2 also contains five sushi domains and two CUB domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) In addition, SEQ ID NO:2 contains a signal peptide and a single transmembrane domain, as identified by HMMER analysis.

For example, SEQ ID NO:3 is 43% identical to Gallus gallus lysozyme (GenBank ID g4467410) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 5.2e-40, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:3 also contains a G-lysozyme signature domain as determined by searching for statistically significant matches in the BLIMPS analysis of the PRINTS database of conserved protein motifs. (See Table 3.) Data from the PFAM, PRODOM and DOMO databases provide further corroborative evidence that SEQ ID NO:3 is a lysozyme.

For example, SEQ ID NO:17 has a signal peptide, as determined by SPScan and hidden Markov model (HMM) based analyses. SEQ ID NO:17 is 86% identical to human immunoglobulin lambda light chain (GenBank ID g33702) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 2.2e-106, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:17 also contains an immunoglobulin domain as determined by searching for statistically significant matches in the HMM-based PFAM database of conserved protein family domains. (See Table 3.) Data from

WO 01/58353

PCT/US01/19862

BLIMPS, MOTIFS, and PROFILES SCAN analyses provide further corroborative evidence that SEQ ID NO:17 is a secreted immunoglobulin. The available evidence shows that SEQ ID NO:19 is also a secreted immunoglobulin.

For example, SEQ ID NO:38 shows 95% identity to human immunoglobulin lambda light chain (GenBank ID g33718) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $5.2e-114$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:38 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, MOTIFS, and PROFILES SCAN analyses provide further corroborative evidence that SEQ ID NO:38 is a secreted protein, and more specifically an immunoglobulin. SEQ ID NO:4-16, SEQ ID NO:18-37, and SEQ ID NO:39-44 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-44 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO.) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:45-88 or that distinguish between SEQ ID NO:45-88 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences in column 5 relative to their respective full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 6735891H1 is the identification number of an Incyte cDNA sequence, and LIVRTUT13 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 71013085V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g1496797) which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. The Genscan-predicted coding sequences

WO 01/98353

PCT/US01/19862

may have been edited prior to assembly. (See Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and GenScan-predicted exons brought together by an "exon stitching" algorithm. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and GenScan-predicted exons brought together by an "exon-stretching" algorithm. (See Example V.) In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but 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:45-88, which encodes SECP. The polynucleotide sequences of SEQ ID NO:45-88, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding SECP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding SECP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:45-88 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:45-88. 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.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding SECP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide

WO 01/98353

PCT/US01/19862

sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring SECP, and all such variations are to be considered as being specifically disclosed.

5 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
10 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.

15 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.

20 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:45-88 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
25 "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
30 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
35 sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system

WO 01/98353

PCT/US01/19862

(Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) *Short Protocols in Molecular Biology*, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) *Molecular Biology and Biotechnology*, Wiley VCH, New York NY, pp. 856-853.)

- 5 The nucleic acid sequences encoding SECP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.)
- 10 Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme
- 15 digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-3060.) Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of
- 20 about 68°C to 72°C.

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

- Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the
- 35 emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate

WO 01/98353

PCT/US01/19862

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

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. 20 Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of SECP, such as its biological or enzymatic activity or its ability to bind to other 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 25 properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of 30 homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding 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 35 Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.)

WO 01/98353

PCT/US01/19862

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, W.H. Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis
 5 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.

The peptide may be substantially purified by preparative high performance liquid
 10 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.)

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
 15 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 may also be used to achieve more efficient translation of sequences encoding SECP. Such signals
 20 include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding SECP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be
 25 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.)

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

35 A variety of expression vector/host systems may be utilized to contain and express sequences

WO 01/98353

PCT/US01/19862

encoding SECP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); 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 Hecke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Sonja (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding SECP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding SECP can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPT11 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 Hecke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of SECP are needed, e.g. for the production of antibodies, vectors which direct high level expression of SECP may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

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

WO 01/98353

PCT/US01/19862

Bio/Technology 12:181-184.)

Plant systems may also be used for expression of SECP. Transcription of sequences encoding SECP may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology 10 (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding SECP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain 15 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 20 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 25 of SECP in cell lines is preferred. For example, sequences encoding SECP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance 30 to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine 35 phosphoribosyltransferase genes, for use in *tk*⁻ and *aprt*⁻ cells, respectively. (See, e.g., Wigler, M. et

WO 01/58353

PCT/US01/19862

- 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 methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *put* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g.,
- 5 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., *trpH* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate
- 10 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.)

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

15 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 promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

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

25 Immunological methods for detecting and measuring the expression of SECP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on SECP is preferred, but a

30 competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) *Serological Methods, a Laboratory Manual*, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) *Current Protocols in Immunology*, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) *Immunochemical Protocols*, Humana Press, Totowa NJ.)

35 A wide variety of labels and conjugation techniques are known by those skilled in the art and

WO 01/58353

PCT/US01/19862

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, 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 ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

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

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted 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, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing 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

WO 01/98353

PCT/US01/19862

cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, c-myc, and hemagglutinin (HA) enable immunoprecipitation of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the SECP encoding 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.

10 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.

15 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 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.

30 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

WO 01/98353

PCT/US01/19862

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 Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) *Science* 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. et al. (1997) *Nucleic Acids Res.* 25:4323-4330). 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

WO 01/98353

PCT/US01/19862

(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
 5 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).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists
 10 between regions of SHCP and secreted proteins. In addition, the expression of SECP is closely associated with reproductive, endocrine, immune system, gastrointestinal, fibroblastic, lung, brain and neurological tissue. 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
 15 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
 20 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,
 25 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-
 30 candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxicity, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or
 35 pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's

WO 01/98353

PCT/US01/19862

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

5 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

10 disease, congenital heart disease, complications of cardiac transplantation, arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery; congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary

15 hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasma pneumoniae, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis,

20 hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung

25 transplantation; 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

30 empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Sträussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebellar hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central

35 nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic

WO 01/98353

PCT/US01/19862

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),

- 5 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 seizures; and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia,
- 10 genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary 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.

- 15 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
- 20 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.

- 25 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
- 30 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 SHCP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of SHCP including, but not limited to, those described above.

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

WO 01/98353

PCT/US01/19862

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.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, 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, polyamines, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

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

Monoclonal antibodies to SECP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1983) *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.* 4:109-120.)

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

WO 01/98353

PCT/US01/19862

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 idiotype composition, may be
 5 generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA
 10 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for SECP may also be generated. For example, such fragments include, but are not limited to, F(ab)₂ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab)₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and
 15 easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such
 20 immunoassays typically involve the measurement of complex formation between SECP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering SECP epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay
 25 techniques may be used to assess the affinity of antibodies for SECP. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of SECP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple SECP epitopes, represents the average affinity, or avidity, of the antibodies for
 30 SECP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular SECP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{13} 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
 35 procedures which ultimately require dissociation of SECP, preferably in active form, from the

WO 01/98353

PCT/US01/19862

antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to
 5 determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of SECP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g.,
 10 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
 15 SECP. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding SECP. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered
 20 intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood*
 25 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding SECP may be used for
 30 somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency
 35 (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475),

WO 01/98353

PCT/US01/19862

cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Sonja (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA, 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in SECP expression or regulation causes disease, the expression of SECP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in SECP are treated by constructing mammalian expression vectors encoding SECP and introducing these vectors by mechanical means into SECP-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J.-L. and H. R  capon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of SECP include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX 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-RFX plasmid (Invitrogen)); the ecodysone-inducible promoter (available in the plasmids PVGRXR and PIND; invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. *supra*); or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding SECP from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID

WO 01/98353

PCT/US01/19862

TRANSEFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, B. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to 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 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 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 Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ 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; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, J. (1997) *Blood* 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding SECP to cells which have one or more genetic abnormalities with respect to the expression of SECP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) *Transplantation* 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinuzzi, P.A. et al. (1999)

WO 01/98353

PCT/US01/19862

Annu. Rev. Nutr. 19:511-544 and Verma, J.M. and N. Sornia (1997) *Nature* 389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding SECP to target cells which have one or more genetic abnormalities with respect to the expression of SECP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing SHCP 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 (Liu, X. et al. (1999) *Exp. Eye Res.* 69:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding SECP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1993) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for SECP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of SECP-coding RNAs and the synthesis of high levels of SECP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SDN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will

WO 01/98353

PCT/US01/19862

allow the introduction of SECP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding SECP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding SECP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible

WO 01/98353

PCT/US01/19862

modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, 5 guanosine, 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 compound which is effective in altering expression of a polynucleotide encoding SECP. Compounds 10 which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of 15 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 20 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 25 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 30 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 35 a test compound indicates that the test compound is effective in altering the expression of the

WO 01/98353

PCT/US01/19862

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:B15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13).

5 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 (Brucce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Brucce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

10 Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat.

15 Biotechnol. 15:462-465.)

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
20 which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of SECP, antibodies to SECP, and mimetics, agonists, antagonists, or inhibitors of SECP.

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

Compositions for pulmonary administration may be prepared in liquid or dry powder form.
30 These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration
35

WO 01/58353

PCT/US01/19862

without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

5 Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising SECP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, SECP or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to
10 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, or pigs. An animal model may also be used to determine the appropriate concentration
15 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 ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by
20 standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are
25 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
30 subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week,
35 or biweekly depending on the half-life and clearance rate of the particular formulation.

WO 01/58353

PCT/US01/19862

Normal dosage amounts may vary from about 0.1 μ g to 100,000 μ g, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

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

WO 01/98353

PCT/US01/19862

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:45-88 or from genomic sequences including promoters, enhancers, and introns of the SECP gene.

Means for producing specific hybridization probes for DNAs encoding SECP include the cloning of polynucleotide sequences encoding SECP or SECP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to 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 ^{32}P or ^{35}S , 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), myofibroblastosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, a cancer of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxicity, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cardiovascular disorder such as congestive heart failure, ischemic heart disease, angina pectoris,

WO 01/98353

PCT/US01/19862

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,

5 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; congenital lung anomalies, atelectasis, pulmonary congestion and edema,

10 pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative

15 interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease,

20 and complications of lung transplantation; 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

25 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 syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebellar hemangioblastomatosis, encephalotrigeminal syndrome, mental

30 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,

35 and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia,

WO 01/98353

PCT/US01/19862

diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia and seizures; 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. 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 multi-format ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered SECP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding SECP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding SECP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding SECP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of SECP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding SECP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the

WO 01/58353

PCT/US01/19862

patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding SECP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding SECP, or a fragment of a polynucleotide complementary to the polynucleotide encoding SECP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding SECP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding SECP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplicons 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-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).

Methods which may also be used to quantify the expression of SECP include radiolabeling or

WO 01/98353

PCT/US01/19862

biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplue, C. et al. (1999) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, SHCP, fragments of SECP, or antibodies specific for SECP may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhauer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, 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.

WO 01/98353

PCT/US01/19862

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed
5 molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain
10 expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of
15 gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in
20 toxicological screening using toxicant signatures to include all expressed gene sequences.

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

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
30 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
35 separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are

WO 01/98353

PCT/US01/19862

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 (Lucking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendoza, 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, N.J. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing

WO 01/98353

PCT/US01/19862

the amino acid residues of the individual proteins and comparing those 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
5 incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

10 Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are
15 well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding SECP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be
20 preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1
25 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:343-355; 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).
30 (See, for example, Lander, H.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
35 physical map and a specific disorder, or a predisposition to a specific disorder, may help define the

WO 01/98353

PCT/US01/19862

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 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 preferred specific

WO 01/58353

PCT/US01/19862

embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/214,601, U.S. Ser. No. 60/212,890, U.S. Ser. No. 60/222,372, U.S. Ser. No. 60/213,466, U.S. Ser. No. 60/231,435, and U.S. Ser. No. 60/232,889, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

10 Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIzol (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or
15 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,
20 Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURH 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
25 vector system (Stratagene) or SUPERScript 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-
30 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., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto
35 CA), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells

WO 01/58353

PCT/US01/19862

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 (Ede Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.H.A.I. 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

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 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 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.

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

WO 01/58353

PCT/US01/19862

databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.)

- 5 The queries were performed using programs based on BLAST, FASTA, BLIMPS, and UMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or GenScan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and
- 10 Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the
- 15 GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the
- 20 CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used,

25 the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

- 30 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 SBQ ID NO:45-88. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

- 35 Putative secreted proteins were initially identified by running the GenScan gene identification

WO 01/58353

PCT/US01/19862

program against public genomic sequence databases (e.g., gbpr1 and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) *J. Mol. Biol.* 268:78-94, and Burge, C. and S. Karlin (1998) *Curr. Opin. Struct. Biol.* 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a PASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode secreted proteins, the encoded polypeptides were analyzed by querying against PFAM models for secreted proteins. Potential secreted proteins were also identified by homology to Incyte cDNA sequences that had been annotated as secreted proteins. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpr1 public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide 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

"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.

WO 01/58353

PCT/US01/19862

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 *gbpt* public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from *genpept*. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either 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.

20 VI. Chromosomal Mapping of SECP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:45-88 were compared with sequences from the Incyte LIFESHQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:45-88 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 (SIIGC), Whitehead Institute for Genomic Research (WIGR), and Génethon 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.

30 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 Génethon which provide boundaries for radiation

WO 01/58353

PCT/US01/19862

hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap '99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

5 In this manner, SEQ ID NO:48 was mapped to chromosome 15 within the interval from 72.3 to 77.4 centiMorgans.

In this manner, SEQ ID NO:54 was mapped to chromosome 20 within the interval from 6.20 to 9.40 centiMorgans. SEQ ID NO:61 was mapped to chromosome 22 within the interval from 0.00 to 19.50 centiMorgans.

10 In this manner, SEQ ID NO:82 was mapped to chromosome 22 within the interval from 0.0 to 19.5 centiMorgans. SEQ ID NO:85 was mapped to chromosome 12 within the interval from 84.7 to 92.5 centiMorgans and from 137.5 to 145.7 centiMorgans. More than one map location is reported for SEQ ID NO:85, indicating that sequences having different map locations were assembled into a single cluster. This situation occurs, for example, when sequences having strong similarity, but not
15 complete identity, are assembled into a single cluster.

In this manner, SEQ ID NO:66 was mapped to chromosome 16 within the interval from 65.60 to 72.60 centiMorgans. In this manner, SEQ ID NO:67 was mapped to chromosome 11 within the interval from 59.50 to 65.00 centiMorgans. In this manner, SEQ ID NO:69 was mapped to chromosome 6 within the interval from 132.70 to 144.40 centiMorgans.

20 VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel (1995) *supra*, ch. 4 and 16.)

25 Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFEBSEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$30 \quad \frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum}(\text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}))}$$

The product score takes into account both the degree of similarity between two sequences and the
35 length of the sequence match. The product score is a normalized value between 0 and 100, and is

WO 01/98353

PCT/US01/19862

calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding SECP are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hematologic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding SECP. cDNA sequences and cDNA library/tissue information are found in the 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 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.

WO 01/98353

PCT/US01/19862

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA

WO 01/58353

PCT/US01/19862

recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

- 5 In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Labelling and Use of Individual Hybridization Probes

- Hybridization probes derived from SEQ ID NO:45-88 are employed to screen cDNAs,
10 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
15 (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPTHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

- 20 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
25 compared.

X. Microarrays

- The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra),
30 mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Skena (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
35 those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g.,

WO 01/58353

PCT/US01/19862

Schemm, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

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

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 µg/µl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/µl RNase inhibitor, 500 µM dATP, 500 µM dGTP, 500 µM dTTP, 40 µM dCTP, 40 µM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 µl 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 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 µl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 µg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia

WO 01/58353

PCT/US01/19862

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

- 5 Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

- Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 µl of the array element DNA, at an average
10 concentration of 100 ng/µl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

- Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). 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
15 buffered saline (PBS) (Pepco, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

- Hybridization reactions contain 9 µl of sample mixture consisting of 0.2 µg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample
20 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 µl 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
25 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
30 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.

WO 01/98353

PCT/US01/19862

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
 5 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
 10 the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the
 15 two fluorophores and adding identical amounts of each to the hybridization mixture.

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

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

XI. Complementary Polynucleotides

Sequences complementary to the SECP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring SECP. Although use of oligonucleotides
 30 comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of SECP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is
 35 designed to prevent ribosomal binding to the SECP-encoding transcript.

WO 01/58353

PCT/US01/19862

XII. Expression of SECP

Expression and purification of SECP is achieved using bacterial or virus-based expression systems. For expression of SECP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express SECP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of SECP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding SECP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, SECP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from SECP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10 and 16). Purified SECP obtained by these methods can be used directly in the assays shown in Examples XVI and XVII, where applicable.

XIII. 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

WO 01/58353

PCT/US01/19862

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 analysis or microarray techniques.

XIV. 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 rabbits 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

WO 01/58353

PCT/US01/19862

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
 5 antipeptide and anti-SECP activity by, for example, binding the peptide or SECP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring SECP Using Specific Antibodies

Naturally occurring or recombinant SECP is substantially purified by immunoaffinity
 10 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
 15 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.

XVI. Identification of Molecules Which Interact with SECP

SECP, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent.
 20 (See, e.g., Bolton A.B. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled SECP, washed, and any wells with labeled SECP complex are assayed. Data obtained using different concentrations of SECP are used to calculate values for the number, affinity, and association of SECP with the
 25 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)
 30 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).

XVII. Demonstration of SECP Activity

An assay for the determination of SECP activity consists of an enzyme reaction mixture
 35 consisting of 25 mM Tris-HCl (pH 7.4), 0.25% Triton X-100, 5 mM MgCl₂, 5 mM CDP-choline, 5

WO 01/58353

PCT/US01/19862

mM 2-mercaptoethanol, 0.05 mM UDP-[¹⁴C]GalNAc (4,000 cpm/nmol), 250 μM peptide, and varying amounts of SECP in a final volume of 100 μL. The reaction mixture is incubated for 10 min. at 37°C followed by Dowex 1 ion exchange (formic acid form) chromatography. Eluted peptide-containing fractions are subjected to scintillation counting. The amount of [¹⁴C]GalNAc present in the peptide-containing fractions is proportional to SECP activity. Confirmation of substrate and SECP source can be evaluated by C-18 chromatography (C2C18 3.2 Smart System, Pharmacia Biotech Inc.) to ensure peptide stability and that incorporated [¹⁴C]GalNAc is associated with the peptide (Sørensen, T. et al. (1995) J. Biol. Chem. 270:24166-24173).

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 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 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 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 the amount of SECP in transit through the secretory pathway.

WO 01/58353

PCT/US01/19862

In another alternative, SECP recognizes and precipitates antigen from serum. This activity can be measured by the quantitative precipitin reaction. (Golub, E. S. et al. (1987) Immunology: A Synthesis, Sinauer Associates, Sunderland, MA, pages 113-115.) SECP is isotopically labeled using methods known in the art. Various serum concentrations are added to constant amounts of labeled

5 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 precipitation curve is obtained, in

10 which the amount of precipitable SECP-antigen complex initially increases proportionately with increasing serum concentration, peaks at the equivalence point, and then decreases proportionately with further increases in serum concentration. Thus, the amount of precipitable SECP-antigen complex is a measure of SECP activity which is characterized by sensitivity to both limiting and excess quantities of antigen.

Alternatively, an assay for SECP activity measures the expression of SECP on the cell

15 surface. cDNA encoding SECP is transfected into a non-leukocytic cell line. Cell surface proteins are labeled with biotin (de la Fuente, M.A. et al. (1997) Blood 90:2398-2405). Immunoprecipitations are performed using SECP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of SECP expressed on the cell surface.

20 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

25 embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

WO 01/98353

PCT/IS01/19862

Table 1

Project ID	Incyte	Polypeptide	Incyte	Polynucleotide	Incyte	Polynucleotide
		SEQ ID NO:		SEQ ID NO:		SEQ ID NO:
210386	1		210386CD1	45		210386CD1
436230	2		436230CD1	46		436230CD1
436232	3		436232CD1	47		436232CD1
5308104	4		5308104CD1	48		5308104CD1
3092736	5		3092736CD1	49		3092736CD1
3580257	6		3580257CD1	50		3580257CD1
3534758	7		3534758CD1	51		3534758CD1
4027923	8		4027923CD1	52		4027923CD1
4348533	9		4348533CD1	53		4348533CD1
4521857	10		4521857CD1	54		4521857CD1
4722253	11		4722253CD1	55		4722253CD1
4878134	12		4878134CD1	56		4878134CD1
5050133	13		5050133CD1	57		5050133CD1
5630134	14		5630134CD1	58		5630134CD1
5677286	15		5677286CD1	59		5677286CD1
6435754	16		6435754CD1	60		6435754CD1
1828282	17		1828282CD1	61		1828282CD1
3203550	18		3203550CD1	62		3203550CD1
0035500	19		0035500CD1	63		0035500CD1
1351534	20		1351534CD1	64		1351534CD1
1402211	21		1402211CD1	65		1402211CD1
1623474	22		1623474CD1	66		1623474CD1
1706443	23		1706443CD1	67		1706443CD1
1748627	24		1748627CD1	68		1748627CD1
1818332	25		1818332CD1	69		1818332CD1
1828282	26		1828282CD1	70		1828282CD1
1832219	27		1832219CD1	71		1832219CD1
1895010	28		1895010CD1	72		1895010CD1
2008768	29		2008768CD1	73		2008768CD1
2070984	30		2070984CD1	74		2070984CD1
2132400	31		2132400CD1	75		2132400CD1
2132400	32		2132400CD1	76		2132400CD1
2132400	33		2132400CD1	77		2132400CD1
2414272	34		2414272CD1	78		2414272CD1
24651076	35		24651076CD1	79		24651076CD1
1957517	36		1957517CD1	80		1957517CD1
866038	37		866038CD1	81		866038CD1
3869704	38		3869704CD1	82		3869704CD1
1415179	39		1415179CD1	83		1415179CD1

WO 01/98353

PCT/US01/19862

Table 1 (cont.)

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
1684752	40	1684752CD1	84	1684752CE1
5390129	41	5390129CD1	85	5390129CE1
1403326	42	1403326CD1	86	1403326CE1
1403326	43	1403326CD1	87	1403326CE1
7690129	44	7690129CD1	88	7690129CE1

WO 01/98353

PCT/US01/19862

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID NO:	GenBank ID NO:	Probability Score	GenBank Homolog
1	2401688CD1	9971461	1.50E-141	UDP-GalNAc:polypeptide N-acetylglucosaminyl transferase (bovine) [Watanabe et al. J. Biol. Chem. [1995] 270(41):24156-65]
2	5452330CD1	94139548	0	Seizure-related gene product 6 Type 2 precursor [Mus musculus] (Shimizu-Mishikawa, K. et al. [1995] Biochem. Biophys. Res. Commun. 216:382-389)
3	4352432CD1	94467410	5.20E-40	Lysozyme [Gallus gallus] (Takano, T. & Graf, T. [1992] Oncogene 7:1377-1384)
4	5308104CD1	93878261	2.10E-92	Similarity to S. pombe BMM/BD05 (G. S. S. [1995] J. Biol. Chem. 270:1620-1629)
16	6435791CD1	913274582	5.00E-39	Thyous macrophage-related protein (Mus musculus)
17	1820972CD1	933702	2.20E-106	Immunoglobulin lambda light chain (Homo sapiens)
18	3286805CD1	9431420	1.50E-283	Macrophage specific protein NP51 (Mus musculus) (Spillsbury, K. et al. [1995] Blood 85:1620-1629)
19	3505596CD1	9577056	1.00E-211	C gamma 3 (Homo sapiens)
29	1899016CD1	913384378	8.00E-43	Putative phosphate translocator [Oryza sativa]
36	1957517CD1	91572802	2.90E-65	Enterococcus faecalis TMA8
37	866038CD1	9849238	1.90E-30	Scallop shell protein 1 (Caenorhabditis elegans)
38	3869704CD1	933718	5.20E-114	Immunoglobulin lambda light chain (Homo sapiens)
43	1403326CD1	93983152	8.10E-56	Schlafen3 lymphoid growth regulatory protein (Mus musculus) (Schwarz, D.A. et al. [1998] Immunity 9:657-668)
44	7650129CD1	95715117	3.10E-219	MR1 (Homo sapiens) Melastatin/TAP related protein found in Beckwith-Wiedemann syndrome chromosomal region 1p15.5 (Prawitt, D. et al. [2000] Hum. Mol. Genet. 9:203-216)

Table 3

[illegible]

88

[illegible]

WO 01/98353

PCT/US01/19862

Table 3 (cont.)

SEQ ID	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Parameters
10	4521857CD1	130	S107 W75		Signal cleavage: M1-A38 Transmembrane domain: G20-Y10	SPSCAN HMMER
11	4722254CD1	279	S171 S230 S73 S77 T107 T243 T268	N191 N266 N71	Signal cleavage: M1-A42	SPSCAN
12	4878134CD1	458	S15 S229 S279 S321 S340 S381 S439 T127 T53	N198 N259 N319	Transmembrane domain: L22-L41 Fgd: R118-D120	HMMER MOTIFS
13	5050133CD1	173	S130 S50		Signal cleavage: M1-A31	SPSCAN
14	5630124CD1	335	S295 S302 S384 S478 S481 T404 T180 T225 T243 T252 T275 T292 Y332		Signal cleavage: M1-A39 Signal cleavage: M1-G36	HMMER SPSCAN
15	5677286CD1	71	T42		Signal peptide: M1-A34	HMMER
16	6136791CD1	148	S143 S16 T18 S140 S206 S219 S74	N31	Signal cleavage: M1-A56 Transmembrane domain: L409-F176	SPSCAN HMMER
17	1820873CD1	231			Signal cleavage: M1-S20 Signal cleavage: M1-G16 Go IMMUNOGLOBULIN IG: HISTOCOMPATIBILITY; MAJOR PROTEIN R3994911-118: Y119-C230 MHC FRAMEWORK DOMAIN DMO397/S24319/1-128; MHC class II B-cell chain associated S4520 protein precursor PD174509: L23-V108 Immunoglobulins and MHC protein signature HL00296: T150-S172, Y210-P227 Immunoglobulins and MHC protein signature I9_mhc.prf: K190-S231 Immunoglobulin domain ig: G34-V108, A146-V214 I9_MHC: Y210-F216	SPSCAN BLAST_DOMO BLAST_DOMO BLAST_PRODOM BLIPS_BLOCKS PROFILESSCAN HMMER_PFM
18	3286605CD1	716	S179 S231 S266 S381 S446 S453 S491 T347 T353 T307 T440 T447 T613 T679 T707 S19 T72 Y67 Y78	M155 M255 M375 N272 M375	Signal peptide: M1-E29 Transmembrane domain: S653-1676 Signal cleavage: M1-A17	MOTIFS HMMER HMMER SPSCAN

Table 3 (cont.)

SEQ ID	Inocyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Notes	
119	3504590CD1	519	S105 S144 S339 S356 S396 S426 S75 S82 T334 S771 T503 Y113 Y368	N367	Signal cleavage: M1-C19 Signal cleavage: M1-C19 Signal peptide: M1-C19 MHC HLA-DR DOMAIN DMO1060 P01862 [1-329; 1142-E275, R285-G518 IG GAMMA3 CHAIN C REGION HEAVY DISEASE PROTEIN HPC IMMUNOGLOBULIN GLYCOPROTEIN P0028815: E241-6309 Immunoglobulins and MHC protein signature BL00290: S436-Q459, F495-S512 Immunoglobulins and MHC protein signature Immunoglobulin domain: Ig Ig alpha Y238-Y272, S376-S395, F432-V459 Ig beta Y238-Y272, F459-F501 Signal peptide: M6-L23 Signal cleavage: M1-A28 Transmembrane domain: L12-N30 Leucine zipper motif: L12-L33 Signal peptide: M43-M67 Transmembrane domain: A250-2267 Signal peptide: M345-H366	SPSCAN HMMER HMMER HMMER	
120	003600CD1	172	T73 W71 T50 S128			MOTIFS HMMER SPSCAN	
121	1261534CD1	314				HMMER MOTIFS HMMER	
122	1402321LTD1	542	S430 S131 S437 S186 T273 S371 S395 T417 T426 S454 T334 S44 S453 S379 S269	N2 N359 N408 N409 N424 N529		HMMER	
123	1623474CD1	715	T334 S339 S376 S380 S388 T436 T475 T524 S543 S585 S586 S647 T659 S704 T709 S5 T108 T222 T279 S172 S390 S395 S406 S429 S445 S463 S503 S505 S590	N328 N325 N61 N63 N67 N68 N69 N70 N71 N72 N73 N74 N75 N76 N77 N78 N79 N80 N81 N82 N83 N84 N85 N86 N87 N88 N89 N90 N91 N92 N93 N94 N95 N96 N97 N98 N99 N100 N101 N102 N103 N104 N105 N106 N107 N108 N109 N110 N111 N112 N113 N114 N115 N116 N117 N118 N119 N120 N121 N122 N123 N124 N125 N126 N127 N128 N129 N130 N131 N132 N133 N134 N135 N136 N137 N138 N139 N140 N141 N142 N143 N144 N145 N146 N147 N148 N149 N150 N151 N152 N153 N154 N155 N156 N157 N158 N159 N160 N161 N162 N163 N164 N165 N166 N167 N168 N169 N170 N171 N172 N173 N174 N175 N176 N177 N178 N179 N180 N181 N182 N183 N184 N185 N186 N187 N188 N189 N190 N191 N192 N193 N194 N195 N196 N197 N198 N199 N200 N201 N202 N203 N204 N205 N206 N207 N208 N209 N210 N211 N212 N213 N214 N215 N216 N217 N218 N219 N220 N221 N222 N223 N224 N225 N226 N227 N228 N229 N230 N231 N232 N233 N234 N235 N236 N237 N238 N239 N240 N241 N242 N243 N244 N245 N246 N247 N248 N249 N250 N251 N252 N253 N254 N255 N256 N257 N258 N259 N260 N261 N262 N263 N264 N265 N266 N267 N268 N269 N270 N271 N272 N273 N274 N275 N276 N277 N278 N279 N280 N281 N282 N283 N284 N285 N286 N287 N288 N289 N290 N291 N292 N293 N294 N295 N296 N297 N298 N299 N300 N301 N302 N303 N304 N305 N306 N307 N308 N309 N310 N311 N312 N313 N314 N315 N316 N317 N318 N319 N320 N321 N322 N323 N324 N325 N326 N327 N328 N329 N330 N331 N332 N333 N334 N335 N336 N337 N338 N339 N340 N341 N342 N343 N344 N345 N346 N347 N348 N349 N350 N351 N352 N353 N354 N355 N356 N357 N358 N359 N360 N361 N362 N363 N364 N365 N366 N367 N368 N369 N370 N371 N372 N373 N374 N375 N376 N377 N378 N379 N380 N381 N382 N383 N384 N385 N386 N387 N388 N389 N390 N391 N392 N393 N394 N395 N396 N397 N398 N399 N400 N401 N402 N403 N404 N405 N406 N407 N408 N409 N410 N411 N412 N413 N414 N415 N416 N417 N418 N419 N420 N421 N422 N423 N424 N425 N426 N427 N428 N429 N430 N431 N432 N433 N434 N435 N436 N437 N438 N439 N440 N441 N442 N443 N444 N445 N446 N447 N448 N449 N450 N451 N452 N453 N454 N455 N456 N457 N458 N459 N460 N461 N462 N463 N464 N465 N466 N467 N468 N469 N470 N471 N472 N473 N474 N475 N476 N477 N478 N479 N480 N481 N482 N483 N484 N485 N486 N487 N488 N489 N490 N491 N492 N493 N494 N495 N496 N497 N498 N499 N500 N501 N502 N503 N504 N505 N506 N507 N508 N509 N510 N511 N512 N513 N514 N515 N516 N517 N518 N519 N520 N521 N522 N523 N524 N525 N526 N527 N528 N529 N530 N531 N532 N533 N534 N535 N536 N537 N538 N539 N540 N541 N542 N543 N544 N545 N546 N547 N548 N549 N550 N551 N552 N553 N554 N555 N556 N557 N558 N559 N560 N561 N562 N563 N564 N565 N566 N567 N568 N569 N570 N571 N572 N573 N574 N575 N576 N577 N578 N579 N580 N581 N582 N583 N584 N585 N586 N587 N588 N589 N590 N591 N592 N593 N594 N595 N596 N597 N598 N599 N600 N601 N602 N603 N604 N605 N606 N607 N608 N609 N610 N611 N612 N613 N614 N615 N616 N617 N618 N619 N620 N621 N622 N623 N624 N625 N626 N627 N628 N629 N630 N631 N632 N633 N634 N635 N636 N637 N638 N639 N640 N641 N642 N643 N644 N645 N646 N647 N648 N649 N650 N651 N652 N653 N654 N655 N656 N657 N658 N659 N660 N661 N662 N663 N664 N665 N666 N667 N668 N669 N670 N671 N672 N673 N674 N675 N676 N677 N678 N679 N680 N681 N682 N683 N684 N685 N686 N687 N688 N689 N690 N691 N692 N693 N694 N695 N696 N697 N698 N699 N700 N701 N702 N703 N704 N705 N706 N707 N708 N709 N710 N711 N712 N713 N714 N715 N716 N717 N718 N719 N720 N721 N722 N723 N724 N725 N726 N727 N728 N729 N730 N731 N732 N733 N734 N735 N736 N737 N738 N739 N740 N741 N742 N743 N744 N745 N746 N747 N748 N749 N750 N751 N752 N753 N754 N755 N756 N757 N758 N759 N760 N761 N762 N763 N764 N765 N766 N767 N768 N769 N770 N771 N772 N773 N774 N775 N776 N777 N778 N779 N780 N781 N782 N783 N784 N785 N786 N787 N788 N789 N790 N791 N792 N793 N794 N795 N796 N797 N798 N799 N800 N801 N802 N803 N804 N805 N806 N807 N808 N809 N810 N811 N812 N813 N814 N815 N816 N817 N818 N819 N820 N821 N822 N823 N824 N825 N826 N827 N828 N829 N830 N831 N832 N833 N834 N835 N836 N837 N838 N839 N840 N841 N842 N843 N844 N845 N846 N847 N848 N849 N850		Protein signature: P317-G373 Signal peptide: N187-V211 Transmembrane domain: I49-F67	MOTIFS HMMER HMMER

WO 01/98353

PCT/US01/19862

Table 3 (cont.)

Seq ID	Protein	Anticodon	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
24	1706443CDL	469	S178 T70 S103 S158 T283 T337 S364 T77 S163 S179 T182 S292 S316 S159 T436 S452 S466		Signal peptide: M1-G24 Signal cleavage: M1-G24	HOMER SPSCAN
25	1748627CDL	274	T9 T90 T237 S241 S248 S82 S100 S136 S191	N254 N270	Signal cleavage: M1-A59	SPSCAN
26	1818132CDL	154	S120 S136 T41 S106 T76 S98 S138		Signal cleavage: M1-A26	SPSCAN
27	1822632CDL	102	T13 T19	N16	Signal peptide: M34-P57 Signal motif: R31-D33	HOMER MOTIFS
28	1832219CDL	113			Signal cleavage: M1-G29	SPSCAN
29	1899010CDL	313	S127 S145 S300 S35 S49 T64 S78 S117 T77	N43 N92 N97 N98 N218 N294	Signal peptide: M194-G211 Transmembrane domain: H11-T35, F191-V171, W219-V237	HOMER HOMER
30	2008768CDL	195			Signal peptide: M121-A139	HOMER
31	2070384CDL	350			Transmembrane domain: I95-N116, N122-I145 Signal cleavage: M1-A66	HOMER SPSCAN
32	2193240CDL	360	S227 S220 S221 S7 S38 T138 S318	N159 N207 N218 N142	Transmembrane domain: V40-G61, N84-C102, T193-V191 Signal peptide: M101-S121	HOMER
33	2235177CDL	559	S301 S412 S520 T11 T27 S29 S42 T76 T156 S165	N70 N171 N357 N325 N417	Signal peptide: M191-A209	HOMER
34	2416227CDL	198	S252 T277 T303 S336 T462 T120 T121 S292 S322 S356 S47 T414 S156 S167 S177	N98 N58 N75 N92	Signal peptide: M1-S18 Signal cleavage: M1-S18 Transmembrane domain: T113-T133	HOMER SPSCAN HOMER
35	2451076CDL	73	T40 S25 T41		Signal peptide: M1-G21 Signal cleavage: M1-V19	HOMER SPSCAN

Table 3 (cont.)

[illegible]

WO 01/98353

PCT/US01/19862

Table 3 (cont.)

SEQ ID NO.	Unclonable Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and References
43	A00326CD1	901	S128 S13 S139 S219 S269 S383 S511 S531 S538 S603 S641 S703 S801 S805 S853 S858 T154 T230 T25 T256 T344 T352 T354 T453 T502 T650 T658 T778 T779 T821 T824 T825 T826 T827 T828 T829 T830 T831 T832 T833 T834 T835 T836 T837 T838 T839 T840 T841 T842 T843 T844 T845 T846 T847 T848 T849 T850 T851 T852 T853 T854 T855 T856 T857 T858 T859 T860 T861 T862 T863 T864 T865 T866 T867 T868 T869 T870 T871 T872 T873 T874 T875 T876 T877 T878 T879 T880 T881 T882 T883 T884 T885 T886 T887 T888 T889 T890 T891 T892 T893 T894 T895 T896 T897 T898 T899 T900 T901 T902 T903 T904 T905 T906 T907 T908 T909 T910 T911 T912 T913 T914 T915 T916 T917 T918 T919 T920 T921 T922 T923 T924 T925 T926 T927 T928 T929 T930 T931 T932 T933 T934 T935 T936 T937 T938 T939 T940 T941 T942 T943 T944 T945 T946 T947 T948 T949 T950 T951 T952 T953 T954 T955 T956 T957 T958 T959 T960 T961 T962 T963 T964 T965 T966 T967 T968 T969 T970 T971 T972 T973 T974 T975 T976 T977 T978 T979 T980 T981 T982 T983 T984 T985 T986 T987 T988 T989 T990 T991 T992 T993 T994 T995 T996 T997 T998 T999		E-loop Arg_Gly_A: G599-T606	MOTIFS
44	7690139CD1	1040	S191 S254 S367 S399 S579 S679 S699 S971 S978 T112 T140 T162 T503 T515 T544 T729 T93	N116 N54 N818	Leucine zipper: L695-L716 EGF: R40-D42 R341-D343 Transmembrane domain: V806-F823, M553-A773, W544-V862 PROTEIN CHROMOSOME TRANSMERANE ELASTATIN ELAST PRODOM C05C12.3 T01H8.5 I P54D1.5 IV PD151509: V730-R1018 PD18035: R394-R405 PD23792: S282-S296	MOTIFS HMMER

WO 01/98353

PCT/US01/19862

Table 4

Polynucleotide Seq. ID NO.	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
45	210188CB1	2508	71-123	575589111 (LIVRUT01) 562018041 (LIVRUT01) 582448681 (BRAHMO04) 370453281 (UTRUT03) 359104661 (29375701) 501854711 (HNT24NN03) 512431111 (BRAHMO05) 770048931 (KIDP1D801) 347096876 (BRA101T01)	892 1957 1988 192 1 1327 1053 409 1862	370 229 2508 314 304 203 1632 938 2432
46	5452330CB1	4034	1493-1673, 1-1081, 2638-2908, 3129-3535	598285578 (BRA1PER05) 477509111 (BRAHMO01) 540406796 (BRAHMO01) 72930878 (BRA1PER06) 520743281 (LIVRUT01) 540404766 (BRAHMO01) 711548911 (BRAHMO01) 593056811 (BRA1PER05) 72930878 (BRA1PER06) 757959411 (BRA1PER01) 729133878 (BRA1PER06) 436243276 (SK1RMO01) 536243279 (SK1RMO01) 710130851 (SK1RMO01) 580963571 (SK1RMO01) 504450301 (OVARUT01) 55076986 (PROSRO06) 59047681 (COLER03) 710141507 (LIVRUT01) 582448681 (BRAHMO04) 5063027661 (TUE31DT02) 5063027661 (TUE31DT02) 5063027661 (TUE31DT02) 561176476 (COLNUT06)	1 3798 5471 368 303 2931 2298 480 1078 2446 1717 228 1689 1 214 1989 4820 956 219 347 695 1153	427 4034 428 369 363 345 2745 1092 1790 2962 2352 654 355 2273 532 765 2283 1814 1810 2308 1072 113 1587
47	4352432CB1	845	1-44, 685-845	729133878 (BRA1PER06)	1717	2352
48	5308104CB1	2300	1-807, 2192-2300	710130851 (SK1RMO01)	1689	2273
49	3092736CB1	1587	1-180	5063027661 (TUE31DT02)	347	1072

WO 01/98353

PCT/US01/19862

Table 4 (Cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
49				262391F6 (ADREVT06)	1031	1512
				5CGA12762V1	1	524
50	3580257CB1	669	1-24	3580257F6 (295TF4701)	133	669
				5107219H1 (PROG70S191)	1	240
				51496797	1	435
51	3634758CB1	1463	1-51	4719037H1 (BRAHECT02)	1177	1432
				82AF05002V1	1	521
				82AF05483V1	379	468
				3243182H1 (BRAINGT19)	1331	1463
				2129861H1 (OVARCT04)	1035	1333
52	4027921CB1	1686	1-204, 1666-1686	584505521V1	584	1130
				3524361F6 (OVARCT02)	673	1173
				2561351H1 (OVARCT03)	559	1173
				564136H1 (SCORNT01)	1428	1688
				3585158H1 (295TF4701)	325	539
				1281432T6 (COLMNC0716)	1051	1684
				577296721 (BRAHNC01)	75	607
53	4348533CB1	2497	1556-1848, 1-150, 3371-2497, 762-909	6933091H1 (SINTTMR03)	1146	1901
				91617775	1	405
				2890155F6 (LUNGFE04)	1	483
				5781002V1 (OVARCT01)	153	503
				2622314H1 (OVARCT02)	2139	2850
				172311F6 (PROG70S191)	564	1160
				5945311H1 (LUNGFE04)	1132	1798
54	4521857CB1	1783	1-733, 805-890	3003172H1 (LUNGCT06)	900	1194
				4521857F6 (LUNGCT01)	1	537
				825638V1 (PROG70S191)	1100	1762
				857689V1 (INGANN01)	1194	1777
				3644845F6 (LUNGCT04)	450	903
				36241756 (PROG70S191)	1227	1783
55	4722253CB1	1461	1-499	4722253H1 (COLMNC0716)	913	1204
				7018504H1 (KIDNCT01)	1	566
				2455753F6 (ENDACT01)	577	1109

WO 01/98353

PCT/US01/19862

Table 4 (Cont.)

Polynucleotide Seq. ID NO:	Incyra Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
55				93053012	975	1461
				30824556 (HEARF02)	1292	1461
56	4878134CB1	2116	1-1071	339623581 (BRADIT01)	1865	2116
				4501019F6 (BRADIT02)	544	1119
				58A01857D1	467	1044
				3521453T6 (LUNGNO03)	1181	1713
				5021921B1 (OVARNO03)	1745	2041
				3766951H1 (BRST0224)	1553	1851
				4501019T6 (BRADIT02)	1050	1704
57	5050135CB1	702	1-28, 651-702	7067475V1	1	594
				91804638	321	702
				5871022H1 (BRACNO02)	1	830
				372950F6 (SKCNO03)	159	666
58	5630124CB1	2613	1-975	6821360T1 (STIMNO01)	520	1293
				6431013H1 (LUNGNO02)	1175	1881
				6655455H1 (BRADIT09)	1	643
				3678611T6 (SPLEN011)	2075	2590
				1358001T6 (LUNGNO09)	1940	2586
				481430R7 (LIVERCT01)	854	1381
				2252822R6 (OVARNO01)	2111	3613
				481430R7 (LIVERCT01)	1371	2013
59	5677286CB1	1778	1736-1778, 1-143, 672-767	70613827V1	1195	1777
				7053934H2 (BRACNO02)	502	1280
				5340571H2 (BRADIT01)	669	2324
				3820871R6 (BRST0224)	1	1776
60	6436791CB1	1234	1-132	3312851R6 (BRST0224)	1421	276
				3510031R6 (CONCNO01)	556	520
				3943483R6 (SCORNO04)	764	1234
61	1820972CB1	863	1-228, 843-863	60144357B1	227	833
				70636975V1	253	863
				1820972B1 (GBLATUT01)	1	267
62	3286805CB1	2521	1-135, 1165-2294	5030319F7 (COLCIT01)	1300	1973
				7168560H1 (MCLRNO01)	575	1020
				6953075H1 (SKINDIAC1)	1	674

WO 01/98353

PCT/US01/19862

Table 4 (Cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
62				326805F6 (HEA00T05)	1706	2266
				545603H1 (PLAC0801)	764	1319
				710500D91	1949	2531
63	3506390CB1	1765	1-798	71409670V1	1051	1755
				7710638H1 (TES0T0E02)	524	1080
				70515763V1	1	556
64	003600CB1	1264	1-699	7733848H2 (COLD00E01)	543	1266
				2748119H1 (HY00T09)	195	442
				2377116V6 (HY01A2T01)	714	1264
				003600R6 (HNC0T01)	509	1042
				003108H1 (HNC0T01)	269	458
				4532216H1 (BAS0T01)	398	525
				231156H1 (KID0T019)	1	325
65	1251534CB1	3415	1-322, 3125-3228, 3482-3415, 834-1643	551771H6 (BRA0T03)	2338	2367
				4697183P6 (BRA0T01)	235	863
				487800H1 (HT02A0T01)	1633	1906
				474028H1 (HY00R02)	1210	1470
				171727V6 (UC00T02)	2865	3415
				7761153H1 (HY00R02)	1	708
				6855869H1 (BRA0T08)	1969	2594
				5443677H1 (BRA0T02)	1814	2453
				3291485F6 (POR0T01)	798	1410
				540433H1 (BRA0T01)	1669	1969
				1251534H1 (JUNG0T03)	1448	1861
66	140221CB1	2289	1707-2289	165008F6 (PRO0T09)	1336	2512
				422477V6 (BRA0T01)	1636	2449
				322477H1 (PRO0T09)	1636	2449
				313603R6 (KID0T03)	1022	512
				348955F7 (AD00T05)	564	1084
				538704H1 (LIV00T03)	1031	1264
				926101H1 (PRO0T06)	1332	1928
				439747H6 (BRA0T01)	335	862
67	1623474CB1	4480	2411-3066, 1-22, 109-587, 3638-3733	2158031P6 (BRA0T09)	3973	4480

WO 01/98353

PCT/US01/19862

Table 4 (Cont.)

Polynucleotide SEQ ID NO:	Tribyte polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
67				333242576 (BRA12701)	2945	3643
				691439501 (PI2UDIR01)	2128	2758
				691627681 (PIACPF06)	213	977
				612811581 (BRABR005)	2349	3030
				707582951	1025	1574
				339407481 (JUNGNOT28)	1	287
				186480386 (PROSMOT19)	3133	3718
				137244176 (PARCUT02)	3747	4452
				137244176 (PARCUT02)	2664	3436
				137244176 (PARCUT02)	2664	3436
				707510911	891	1418
				707579301	1651	2153
				603053441	1374	2060
68	170644381	1568	1-43	663021011 (PROSMOT18)	438	905
				186859376 (PROSMOT18)	1044	1521
				170644376 (DIADNOT02)	506	1511
				186859376 (PROSMOT18)	656	1511
				706267781 (PENITMO02)	1	468
				139024981 (EOSTNOT01)	1342	1568
69	174852701	1887	1-649	540781288 (BRAMNOT1)	439	985
				714275021	1299	1887
				688674901 (BRAMTDR03)	632	1022
				362739176 (COLNOT38)	1	531
				714302511	1047	1612
				197928386 (JUNGNOT03)	997	1556
70	181833201	559	1-35	128577952 (PENITMO03)	39	569
				133602471 (COLNOT13)	1	541
71	182283201	2338	529-565, 1332-1369, 1488-2338	128970976 (BRAMNOT11)	1977	2338
				18228323201 (GSLAUT01)	25	655
				61975312	1	277
				SAGA017201	710	1312
				SAGA014161	1318	1894
				145284376 (EAMITUT01)	520	1212
				SAGA012951	1707	2338
				SAGA008371	1223	1841

Table 4 (Cont.)

Polynucleotide Seq ID NO:	Incyele Polynucleotide ID	Sequence Length	Selected Fragment(s) 1-21	Sequence Fragments	5' Position	3' Position
72	1832419CBI	881	1-21	5X202203V1 18322219R6 (BRAINOM01)	1	479
73	1839910CBI	1255	1-62	18399010P6 (BLADTUM03) 214773F6 (CONNTU01) 19095277G6 (CONNTU01) 1425473H1 (BSPINOM01) 19095277G6 (CONNTU01) 200876B7G6 (TESTWOT03)	44 341 569 1 35	481 863 1255 1233 264 623
74	200876B7CBI	875	1-411	200876B7G6 (TESTWOT03) 6055379H1 (TESTWOT11) 563232R6 (BUTLPT01) 1271987F1 (TESTWOT02)	1 472 1 1769	858 875 2188
75	2070984CBI	2188	1-72, 1579-1666	715293H1 (BLACDOR01) SBA08315D1 SBA01468D1 696751A1 (PTUDIR01) 3320716H1 (PROSEPT01) 1624251P6 (BRAFUT01) 2459918R6 (METHUN02) 2459918T6 (METHUN02) 900981R1 (BRSPTU03) 7111502V1 6993449H1 (BRAFUT02) 7126525V1 6821261V1 (SINTH0R01) 2455227G6 (BNT3AZT01) 2456227G6 (BNT3AZT01) 854765H1 (NSANN001) 7017947H1 (KIDNOC01) 7154302H1 (HEARW0R03) 7059965H1 (TESTW0R02)	321 1107 891 112 280 1 734 1003 365 1140 1 17 194 1210 976 774 1 1446 110	914 1593 1516 81 81 280 472 1174 1561 926 1777 743 679 1208 902 1389 1436 1005 504 1841 673
76	2193240CBI	1561	1-624	2193240CBI	1	472
77	2215177CBI	1777	1-32	2215177CBI	1	472
78	2456227CBI	1861	1-433 518-1018	2456227CBI	1	433
79	2461076CBI	1616	835-861, 565-783, 1-219	2461076CBI	1	219

WO 01/98353

PCT/US01/19862

Table 4 (Cont.)

Polynucleotide SEQ ID NO:	Inventor Polynucleotide ID	Sequence Length	Selected segment(s)	Sequence Fragments	5' Position	3' Position
80	1957517CB1	1434	1-111	7161288H1 (FLACNOT01)	448	1014
				1233279F1 (LUNGFEU01)	1	537
				6573238H1 (COLLAFUS02)	695	1348
81	865038CB1	2085	51-124	1575844H1 (LUNGNOT03)	1214	1434
				6913289F1 (PIRUDIR01)	272	823
				5964475H1 (BRAHNOT05)	1744	2085
				7132506F1 (BRAHNOT01)	816	1499
				6034821H1 (BRAHNOT04)	784	1447
				735563F1 (BRAHNOT02)	215	753
				5493459H1 (PIRFEU06)	1	283
82	3897704CB1	904	1-36	6424276H1 (SICRNOT02)	1420	2033
				708156H1 (STORAT04)	1	204
				1588168H1 (SICRNOT03)	33	202
				3901238H1 (LUNGNOT03)	48	897
83	1415179CB1	1496	1-248 606-836	3904253H1 (LUNGNOT03)	217	837
				2042611H1 (HIFONOT02)	729	1206
				560950H1 (BRALNOT03)	557	1155
				4713360H1 (BRAHNOT01)	1	252
				558639F1 (BRAHNOT03)	839	1496
				2708523H1 (SORSAT01)	430	732
				2967826F1 (SORSNOT04)	58	686
84	1664792CB1	2437	1-1559	70858742V1	407	959
				4787546H1 (LIVRNOT09)	1349	1524
				3699003T1 (OVARNOT01)	2154	2825
				7122472H1	867	1504
				2542339F1 (BONRNOT01)	1859	2386
				746043H1 (LIVRNOT03)	441	1062
				2204842H1 (PERSNOT02)	2845	3237
				1588168H1 (SICRNOT02)	1678	449
				1664792H1 (PERSNOT02)	1678	449
				1664792H1 (PERSNOT02)	1678	449
				4179737H1 (SINTNOT03)	1257	1834
85	2079396CB1	1123	1-45 993-1123	6620736H1 (SINTNOT01)	1569	1863
				91401473	1	1063
				874769F1 (LUNGAST01)	786	507
				5819702J1 (OVARNOT01)	54	1143
85				6335288H1 (BRANDINO1)	17	794
						509

WO 01/98353

PCT/US01/19862

Table 4 (Cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
86	539013CB1	1549	1-270	1258145F1 (RENITUR01)	589	1247
				1466677F1 (PARCTUR02)	1086	1549
				4256616F6 (BRADITUR01)	1	633
				1340406F1 (COLNFER02)	756	1337
87	1403326CB1	4820	1-3502	1306452F6 (PLACOTUR02)	1069	1588
				70607520V1	291	789
				4323557E1 (LYMNUR01)	2789	3045
				3090429F6 (BRADITUR01)	2387	3023
				7047633V1	1812	459
				1064858F1	1064	1588
				2801438F6 (RENCHOT01)	3054	3590
				473910H3 (GRADITUR01)	1585	1844
				5245574H1 (TESTUR01)	3253	4123
				70815905V1	417	793
				5718724H1 (PARCHOT16)	1188	1815
				6863174H1 (BRACHOT02)	4389	4820
				5937303H1 (PTCHOTUR01)	3745	4307
				6489460H1 (KINDON01)	2100	2703
				4684473H3 (LUNLMTUR01)	2991	3242
				2820527F6 (BRSTOT16)	4094	4505
38	7690129CB1	3599	1978-1968, 1-934, 2349-3111	5643645E8 (PTSTMT01)	603	1080
				1251961F1 (LUNGFE03)	2250	3112
				4651125F6 (LUNGFE03)	2855	3599
				5600156F1 (COLNFER03)	292	320
				5631480F3 (STMTUR03)	1821	2515
				5631161H1 (BRADITUR03)	580	1029
				5656845F8 (COLTIT04)	1644	2425
				71137279V1	2150	2848
				2185757F6 (PROSDT26)	1	495
				7612578J1 (KINDON01)	1041	1732

WO 01/98353

PCT/US01/19862

Table 5

Seq. ID No.	Polynucleotide	Incyte	Representative library
45		Project ID	BRAUT02
46		210168CBL	BRAUT01
47		5452330CBL	SKINOT01
48		4362432CBL	BRAUT03
49		5308104CBL	BRAUT08
50		3082736CBL	29377301
51		3580257CBL	HUVER001
52		3634758CBL	COLANN016
53		4027923CBL	LIVRAN008
54		4148533CBL	SPLANT004
55		4521857CBL	TESTN0T03
56		4722253CBL	LUNGRN003
57		4878134CBL	913PFRN06
58		5050133CBL	LUNGNT009
59		5630124CBL	PROSTUT12
60		567286CBL	MGBUNT01
61		636791CBL	SPLANT004
62		1320972CBL	SKINDIA01
63		2066005CBL	COLDIE01
64		3506590CBL	HMCINOT01
65		0026000CBL	THYMO005
66		1233334CBL	CASCET02
67		1402211CBL	SKINOT01
68		1623474CBL	DOONOT02
69		1706443CBL	938PFRN06
70		1748627CBL	TESTN0T01
71		1818332CBL	BRAINOT11
72		1822832CBL	TESTN0T03
73		1832219CBL	BLADUT006
74		1899010CBL	TESTN0T03
75		2008768CBL	PLACHT007
76		2070984CBL	BRAUT013
77		2193240CBL	BRAUT013
		2235177CBL	BRAUT013

WO 01/98353

PCT/US01/19862

Table 5 (Cont.)

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
78	2416227CB1	LUNGNDT09
79	2461076CB1	STOMADT01
80	1957517CB1	OVARTUT01
81	866038CB1	BRAITUT03
82	3869704CB1	LUNGNDT03
83	1415179CB1	SPALNDT03
84	1664792CB1	BASTUT01
85	2078196CB1	CONUTUT01
86	6380115CB1	BRAITUT03
87	1403226CB1	BRASTUT01
88	7690129CB1	BRASUTUT12

WO 01/98353

PCT/US01/19862

Table 6

Library	Vector	Library Description
BRAITUT01	PINCY	Library was constructed using RNA isolated from a serum-starved transformed embryonal cell line (293-EBNA) derived from kidney epithelial tissue. The cells were transformed with adenovirus 5 DNA.
BRAITUT06	PINCY	Library was constructed using RNA isolated from bladder tumor tissue removed from the bladder of a 58-year-old Caucasian male during a radical cystectomy. Pathology indicated extensive polyploid carcinoma in the left lateral bladder wall. The remaining bladder showed marked cystitis with scattered microscopic foci of transitional cell carcinoma in situ. Patient history included angina, emphysema and tobacco use. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
BRAITUT01	PINCY	Library was constructed using RNA isolated from diseased brain tissue. Patient history included multiple sclerosis, type II lesion.
BRAITUT03	REPORT1	Library was constructed using RNA isolated from brain tissue removed from a 26-year-old Caucasian male during cranioplasty and excision of a cerebral meningial lesion. Pathology indicated a grade 4 oligodendroglioma in the right fronto-parietal part of the brain.
BRAITUT01	PINCY	Library was constructed using RNA isolated from brain tissue removed from the right temporal lobe of a 5-year-old Caucasian male during a hemispherectomy. Pathology indicated extensive polymicrogyria and mild to moderate gliosis (predominantly subpial and subcortical), consistent with chronic seizure disorder. Family history included a cervical neoplasm.
BRAITUT02	REPORT1	Library was constructed using RNA isolated from brain tumor tissue removed from the frontal lobe of a 58-year-old Caucasian male during excision of a cerebral meningial lesion. Pathology indicated a grade 2 metastatic adenocarcinoma. Patient history included angina, emphysema, and tobacco use. Family history included a malignant neoplasm of the kidney.
BRAITUT03	REPORT1	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 17-year-old Caucasian female during excision of a cerebral meningial lesion. Pathology indicated a grade 4 fibrillary glioma and small-cell astrocytoma. Family history included benign hypertension and cerebrovascular disease.
BRAITUT03	PINCY	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 47-year-old Caucasian male during excision of cerebral meningial tissue. Pathology indicated grade 4 fibrillary astrocytoma with focal tumoral necrosis. Patient history included cerebrovascular disease, deficiency anemia, and angina. Family history included cerebrovascular disease and a malignant prostate neoplasm.
BRAITUT03	PINCY	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 68-year-old Caucasian male during excision of a cerebral meningial lesion. Pathology indicated a meningioma in the left frontal lobe.

Table 6 (Cont.)

Library	Vector	Library Description
BRAIN003	pUC19	This normalized library was constructed from 6.7 million independent clones from a brain tumor cDNA library. The library was constructed from total RNA isolated from hypodermic tissue removed from a 57-year-old Caucasian male who died of a cerebrovascular accident. Patient history included Huntington's disease and emphysema. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:3228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours-round) reannealing hybridization was used. The library was linearized and recircularized to select for insert containing clones.
BRSN001	PBUBSCRIPT	Library was constructed using RNA isolated from the breast tissue of a 56-year-old female. The library was constructed using RNase-free conditions.
BRSIT001	PSPORT1	Library was constructed using RNA isolated from breast tumor tissue removed from a 55-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated invasive grade 4 mammary adenocarcinoma of mixed lobular and ductal type, extensively involving the left breast. The tumor was identified in the deep dermis near the lactiferous ducts with extracapsular extension. Seven mild and low and five high axillary lymph nodes were positive for tumor. Proliferative fibrocytic changes were characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia. Immunohistochemical studies showed immunoreactivity for estrogen receptor and a benign breast neoplasm. Family history included benign hypertension, arterosclerotic coronary artery disease, cerebrovascular disease, and depressive disorder.
CARC002	PSPORT1	Library was constructed using RNA from chondrocytes that were isolated from pooled knee cartilage obtained during total knee joint replacement. The cartilage was removed from the underlying bone, chopped into smaller pieces, and stimulated with 5 ng/ml IL-1 for 24 hours.
COLD001	PCRWA-1	This two phase random primed library was constructed using RNA isolated from diseased descending colon tissue removed from a 38-year-old Caucasian male during a total intra-abdominal colectomy and temporary ileostomy. Pathology indicated chronic ulcerative colitis, moderate to severe, actively involving the distal 23 cm of colon. The entire 34 cm segment of rectosigmoid, rectum, and rectal tissue was involved with chronic ulcerative colitis, severely active. The patient presented with blood in the stool, diarrhea, and deficiency anemia. Patient history included shoulder dystonia and rheumatoid arthritis. The patient underwent colectomy and proctocolectomy with ileoanal pouch construction. Postoperative complications included ascariasis, pneumonia, and corticosteroid resistance. Patient included acute myocardial infarction, upper lobe lung cancer, colon cancer, family history diabetes in the grandparents(s).
COLM001S	pUC19	Library was constructed using RNA isolated from sigmoid colon tissue removed from a 62-year-old Caucasian male during a sigmoidectomy and permanent colostomy.
CONUT001	pUC19	Library was constructed using RNA isolated from sigmoid mesenteric tumor tissue obtained from a 62-year-old female during a total abdominal hysterectomy and bilateral oophorectomy. The patient had a long history of endometriosis with a mucinous cystic grade 4 malignant mixed müllerian tumor present in the sigmoid mesentery at two sites.

WO 01/98353

PCT/US01/19862

Table 6 (Cont.)

Library	Vector	Library Description
DUODU02	pINCY	Library was constructed using RNA isolated from duodenal tissue of a 3-year-old Caucasian female, who died from head trauma. Serology was positive for cytomegalovirus (CMV).
ETBPEN06	pINCY	The normalized prostate stromal fibroblast tissue libraries were constructed from 156 million independent clones. The library was constructed from a male status who died after 26 weeks' gestation. The libraries were normalized in two roundusing conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldi et al., Genome Research (1996) 6:791, except that a significantly longer (48-hours/round) reannealing hybridization was used. The library was then linearized and recircularized to select for insert containing clones as follows: plasmid DNA was prepped from approximately 1 million clones from the normalized prostate stromal fibroblast tissue libraries following soft agar transformation.
HMCIN001	PMIDESCRIP	Library was constructed using RNA isolated from the HMC-1 human hair cell line derived from a deaf-blind patient. The library included mast cell leukemia.
HWZAG01	PMIDESCRIP	Library was constructed at Stratagene (SW937233), using RNA isolated from the HWZ cell line derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor. Cells were treated with retinoic acid for 5 weeks and with mitotic inhibitors for two weeks and allowed to mature for an additional 4 weeks in conditioned medium.
HUVEN001	PMIDESCRIP	Library was constructed using RNA isolated from HUV-EC-C (ATCC CRL 1730) cells.
ISLIN001	pINCY	Library was constructed using RNA isolated from a pooled collection of pancreatic islet cells.

Table 6 (Cont.)

Library	Vector	Library Description
LIVNOR008	pINCY	This normalized library was constructed from 5.7 million independent clones from a pooled liver tissue library. Starting RNA was made from pooled liver tissue removed from a 4-year-old diabetic male who died from toxemia and a 16 week female fetus who died from a congenital infection. The library was constructed using RNA isolated from the chromosomes of the 4-year-old patient. Family history included asthma in the 4-year-old. Family history included taking daily prenatal vitamins and maternal valve prolapse in the mother of the fetus. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:3228 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
LUNGNOR01	PSFOR1	This normalized library was constructed from 2.56 million independent clones from a lung tissue library. RNA was made from lung tissue removed from the left lobe a 58-year-old Caucasian male during a segmental lung resection. Pathology for the associated tumor tissue indicated a metastatic grade 3 (4) osteosarcoma. Patient history included soft tissue sarcoma of the right arm, a lung cancer, and a second osteosarcoma of the right tibia with hemorrhage. Patient also received radiation therapy to the retroperitoneum. Family history included prostate cancer, breast cancer, and acute leukemia. The normalization and hybridization conditions were adapted from Soares et al., PNAS (1994) 91:3228; Swarcop et al., NAR (1991) 19:1954; and Bonaldo et al., Genome Research (1996) 6:791.
LUNGNOT03	PSFOR1	Library was constructed using RNA isolated from lung tissue of a 79-year-old Caucasian male. Pathology for the associated tumor tissue indicated grade 4 carcinoma. Patient history included a benign prostate neoplasm and atherosclerosis.
LUNGNOT09	pINCY	Library was constructed using RNA isolated from the lung tissue of a 23-week-old Caucasian male during pregnancy was terminated following a diagnosis by ultrasound of a fetal lung lesion.
REGNUN01	pINCY	Library was constructed using RNA isolated from an untreated H9-01 megakaryoblast cell line, derived from bone marrow cells obtained from a 55-year-old male in megakaryoblastic crisis of chronic myelogenous leukemia.
OVANUT01	PSFOR1	Library was constructed using RNA isolated from ovarian tumor tissue removed from a 43-year-old Caucasian female during removal of the fallopian tubes and ovaries. Pathology indicated grade 2 mucinous cystadenocarcinoma involving the entire left ovary. Patient history included maternal valve disorder, pneumonia, and viral hepatitis. Family history included atherosclerotic coronary artery disease, pancreatic cancer, stress reaction, cardiovascular disease, breast cancer, and ovarian cancer.
FLACN07	pINCY	Library was constructed using RNA isolated from fetal tissue removed from a Caucasian fetus who died after 16 weeks' gestation from fetal demise and hydrocephalus. Serology was positive for anti-CMV (cytomegalovirus).

WO 01/98353

PCT/US01/19862

Table 6 (Cont.)

Library	Vector	Library Description
PROSTUT12	pINCY	Library was constructed using RNA isolated from prostate tumor tissue removed from a 65-year-old Caucasian male during a radical prostatectomy. Pathology indicated an adenocarcinoma (Gleason grade 4+2). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Tissue was removed from the prostate gland and the PSA level was elevated. Tissue was removed from the prostate gland and the PSA level was elevated. Tissue was removed from the prostate gland and the PSA level was elevated.
SKINDTA01	pSPORT1	Library was constructed using RNA isolated from skin tissue removed from a 65-year-old Caucasian male during a skin biopsy. Pathology indicated a tuberculoid and lepromatous leprosy.
SKINW001	pINCY	Library was constructed using RNA isolated from skin tissue removed from the breast of a 26-year-old Caucasian female during bilateral reduction mammoplasty.
SPLENOT04	pINCY	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.
STOMW001	PELVESCRIP1	Library was constructed using RNA isolated from the stomach tissue of a 55-year-old Caucasian male, who died from cardiopulmonary arrest.
TESTW003	PELVESCRIP1	Library was constructed using RNA isolated from testicular tissue removed from a 37-year-old Caucasian male, who died from liver disease. Patient history included AIDS.
THYMW005	pINCY	Library was constructed using RNA isolated from thymus tissue removed from a 3-year-old Hispanic male during a thymectomy and closure of a patent ductus arteriosus. The patient presented with severe pulmonary stenosis and cyanosis. Patient history included a cardiac catheterization and echocardiogram. Previous surgeries included Blalock-Taussig shunt and pulmonary valvotomy. The patient was not taking any medications. Family history included benign hypertension, osteoarthritis, depressive disorder, and extrinsic asthma in the grandparents.

WO 01/98353

PCT/US01/19862

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA.; Peracel 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, fasta, fasta, fasta, and search.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 185:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta.E value= 1.0E-6 Assembled ESTs: fasta.Identity= 95% or greater and Match Length=200 bases or greater; fasta.E value=1.0E-8 or less Full Length sequences: fasta.score=100 or greater
BLIMPS	A BLIck 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.C. Henikoff (1991) Nucleic Acids Res. 19:556-557; Henikoff, J.C. 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
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.J., et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater

WO 01/98353

PCT/US01/19862

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Grisham, M. et al. (1988) <i>CABIOS</i> 4:61-66; Grisham, M. et al. (1989) <i>Methods Enzymol.</i> 183:146-159; Barasch, A. et al. (1997) <i>Nucleic Acids Res.</i> 25:217-221.	Normalized quality scores; GCG-specified "HIGH" value for that particular Profile motif. Generally, scores > 4.2.1.
Phred	A base-calling algorithm that examines automated sequence 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 Phred 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 = 20 or greater; Match length = 56 or greater
Conrad	A graphical tool for viewing and editing Phrap assemblies.	Gordon, 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.	Nilsson, H. et al. (1997) <i>Protein Engineering</i> 10:146; Claverie, J.M. and S. Alric (1997) <i>CABIOS</i> 12:431-439.	Score > 0.5 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> 5:393-371.	
TMEMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Souhammar, E.L. et al. (1998) <i>Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.</i>	
Modis	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) <i>Nucleic Acids Res.</i> 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 01/98353

PCT/US01/19862

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:1-44,
 - b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-44,
 - c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, and
 - 10 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-44,
- 15 3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
- 20 5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:45-88.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
- 25 7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
- 30 9. A method for 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

WO 01/98353

PCT/US01/19862

b) recovering the polypeptide so expressed.

10. An isolated antibody which specifically binds to a polypeptide of claim 1.

5 11. 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:45-88,

b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:45-88,

10 c) a polynucleotide complementary to a polynucleotide of a),

d) a polynucleotide complementary to a polynucleotide of b), and

e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a
15 polynucleotide of claim 11.

13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides
20 comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

25

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

30 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.

WO 01/98353

PCT/US01/19862

16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected
5 from the group consisting of SEQ ID NO:1-44.

18. 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 16.

19. A method for 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.

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

21. A method for treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment a composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- 25
- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
 - b) detecting antagonist activity in the sample.

23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional SECP, comprising administering to a patient in need of such treatment a composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

WO 01/98353

PCT/US01/19862

- 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.

5

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said 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,
- 10 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
- 15 compound that modulates the activity of the polypeptide of claim 1.

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

- 20 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.

25

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

- a) treating a biological sample containing nucleic acids with the test compound;
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific
- 30 hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;
- c) quantifying the amount of hybridization complex; and
- d) comparing the amount of hybridization complex in the treated biological sample with the
- 35 amount of hybridization complex in an untreated biological sample, wherein a difference in the

WO 01/98353

PCT/US01/19862

amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

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

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

10

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

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

15

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

20

32. 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 31.

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

25

34. 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 33.

30

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

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 1-44, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

WO 01/98353

PCT/US01/19862

h) isolating antibodies from said animal; and
c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44.

5

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

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

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

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

15 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

20 e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44.

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

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

25

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

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

43. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44 in a sample, comprising the steps of:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific

WO 01/98353

PCT/US01/19862

binding of the antibody and the polypeptide; and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44 in the sample.

5

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44 from a sample, the method comprising:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

10

b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-44.

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

15

46. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

47. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

48. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

20

49. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

50. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

25

51. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

52. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

53. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

30

54. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

55. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

WO 01/98353

PCT/US01/19862

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
- 5 58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.
60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.
- 10 61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.
62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.
63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.
- 15 64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.
65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.
- 20 66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.
67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.
- 25 68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.
69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.
70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.
- 30 71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.
72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.

WO 01/58353

PCT/US01/19862

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.
74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.
- 5 75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.
76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.
77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.
- 10 78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.
79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.
- 15 80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.
81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:37.
82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.
- 20 83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:39.
84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:40.
- 25 85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:41.
86. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:42.
87. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:43.
- 30 88. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:44.
89. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:45.

WO 01/98353

PCT/US01/19862

90. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:46.

5 91. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:47.

92. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:48.

10 93. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:49.

94. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:50.

95. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:51.

20 96. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:52.

97. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:53.

25 98. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:54.

99. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:55.

30 100. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:56.

WO 01/98353

PCT/US01/19862

101. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:57.
102. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
5 NO:58.
103. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:59.
104. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
10 NO:60.
105. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:61.
- 15 106. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:62.
107. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
20 NO:63.
108. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:64.
- 25 109. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:65.
110. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:66.
- 30 111. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:67.
112. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID

WO 01/98353

PCT/US01/19862

NO:68.

113. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:69.

5

114. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:70.

115. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
10 NO:71.

116. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:72.

117. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
15 NO:73.

118. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:74.

20

119. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:75.

120. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
25 NO:76.

121. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:77.

122. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
30 NO:78.

123. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:79.

WO 01/98353

PCT/US01/19862

124. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:80.

5 125. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:81.

126. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:82.

10 127. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:83.

128. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
15 NO:84.

129. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:85.

20 130. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:86.

131. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:87.

25 132. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:88.

WO 01/58353

PCT/US01/19862

<110> TRCYTE GENOMICS, INC.
 HILLMAN, Jennifer L.
 TANG, Y. Tom
 YUE, Henry
 ELLIOTT, Vicki S.
 TRISOULEY, Catherine M.
 LEE, Ernestine A.
 RAMKUMAR, Jayalaxmi
 LAL, Preeti
 XU, Yuming
 WARREN, Bridget A.
 HAPALA, April J. A.
 DAUGHN, Mariah R.
 ALAMZAI, Yalda
 BATHA, Sajeev
 BIRFORD, Neil
 YAO, Monique G.
 NGUYEN, Daniel B.
 LU, Dyung Aina M.
 WALIA, Nazinder K.
 GANDHI, Ameena R.
 AU-YOUNG, Janice
 PATTERSON, Chandra

<120> SECRETED PROTEINS

<130> PI-0133 PCT

<140> To Be Assigned
 <141> Herewith

<150> 60/212,890; 60/213,466; 60/214,601; 60/222,372; 60/231,435; 60/232,889
 <151> 2000-05-20; 2000-06-23; 2000-06-27; 2000-07-31; 2000-08-08; 2000-09-15

<160> 88

<170> PEARL Program

<210> 1

<211> 552

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2101688CD1

<400> 1

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

WO 01/98353

PCT/US01/19862

```

110      115      120
Asn Glu Ala Arg Ser Thr Leu Leu Arg Thr Ile Arg Ser Val Leu
125      130      135
Asn Arg Thr Pro Thr His Leu Ile Arg Glu Ile Ile Leu Val Asp
140      145      150
Asp Phe Ser Asn Asp Pro Asp Asp Cys Lys Gln Leu Ile Lys Leu
155      160      165
Pro Lys Val Lys Cys Leu Arg Asn Asn Glu Arg Gln Gly Leu Val
170      175      180
Arg Ser Arg Ile Arg Gly Ala Asp Ile Ala Gln Gly Thr Thr Leu
185      190      195
Thr Phe Leu Asp Ser His Cys Glu Val Asn Arg Asp Trp Leu Gln
200      205      210
Pro Leu Leu His Arg Val Lys Glu Asp Tyr Thr Arg Val Val Cys
215      220      225
Pro Val Ile Asp Ile Ile Asn Leu Asp Thr Phe Thr Tyr Ile Glu
230      235      240
Ser Ala Ser Glu Leu Arg Gly Gly Phe Asp Trp Ser Leu His Phe
245      250      255
Gln Trp Glu Gln Leu Ser Pro Glu Gln Lys Ala Arg Arg Leu Asp
260      265      270
Pro Thr Glu Pro Ile Arg Thr Pro Ile Ile Ala Gly Gly Leu Phe
275      280      285
Val Ile Asp Lys Ala Trp Phe Asp Tyr Leu Gly Lys Tyr Asp Met
290      295      300
Asp Met Asp Ile Trp Gly Gly Glu Asn Phe Glu Ile Ser Phe Arg
305      310      315
Val Trp Met Cys Gly Gly Ser Leu Glu Ile Val Pro Cys Ser Arg
320      325      330
Val Gly His Val Phe Arg Lys Lys His Pro Tyr Val Phe Pro Asp
335      340      345
Gly Asn Ala Asn Thr Tyr Ile Lys Asn Thr Lys Arg Thr Ala Glu
350      355      360
Val Trp Met Asp Glu Tyr Lys Gln Tyr Tyr Tyr Ala Ala Arg Pro
365      370      375
Phe Ala Leu Glu Arg Pro Phe Gly Asn Val Glu Ser Arg Leu Asp
380      385      390
Leu Arg Lys Asn Leu Arg Cys Gln Ser Phe Lys Trp Tyr Leu Glu
395      400      405
Asn Ile Tyr Pro Glu Leu Ser Ile Pro Lys Glu Ser Ser Ile Gln
410      415      420
Lys Gly Asn Ile Arg Gln Arg Gln Lys Cys Leu Glu Ser Gln Arg
425      430      435
Gln Asn Asn Gln Glu Thr Pro Asn Leu Lys Leu Ser Pro Cys Ala
440      445      450
Lys Val Lys Gly Glu Asp Ala Lys Ser Gln Val Trp Ala Phe Thr
455      460      465
Tyr Thr Gln Lys Ile Leu Gln Glu Glu Leu Cys Leu Ser Val Ile
470      475      480
Thr Leu Phe Pro Gly Ala Pro Val Val Leu Val Leu Cys Lys Asn
485      490      495
Gly Asp Asp Arg Gln Gln Trp Thr Lys Thr Gly Ser His Ile Glu
500      505      510
His Ile Ala Ser His Leu Cys Leu Asp Thr Asp Met Phe Gly Asp
515      520      525
Gly Thr Glu Asn Gly Lys Glu Ile Val Val Asn Pro Cys Glu Ser
530      535      540
Ser Leu Met Ser Gln His Trp Asp Met Val Ser Ser
545      550

```

<210> 2
 <211> 994
 <212> PRT

WO 01/58353

PCT/US01/19862

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5452330C01

<400> 2

```

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

```

WO 01/58353

PCT/US01/19862

Arg Asn Ala Thr Thr Gly Arg Ile Val Ser Pro Gly Phe Pro Gly
 425 430 435
 Asn Tyr Ser Asn Asn Leu Thr Cys His Trp Leu Leu Glu Ala Pro
 440 445 450
 Glu Gly Gln Arg Leu His Leu His Phe Glu Lys Val Ser Leu Ala
 455 460 465
 Glu Asp Asp Asp Arg Leu Ile Ile Arg Asn Gly Asp Asn Val Glu
 470 475 480
 Ala Pro Pro Val Tyr Asp Ser Tyr Glu Val Glu Tyr Leu Pro Ile
 485 490 495
 Glu Gly Leu Leu Ser Ser Gly Lys His Phe Phe Val Glu Leu Ser
 500 505 510
 Thr Asp Ser Ser Gly Ala Ala Ala Gly Met Ala Leu Arg Tyr Glu
 515 520 525
 Ala Phe Gln Gln Gly His Cys Tyr Glu Pro Phe Val Lys Tyr Gly
 530 535 540
 Asn Phe Ser Ser Ser Thr Pro Thr Tyr Pro Val Gly Thr Thr Val
 545 550 555
 Glu Phe Ser Cys Asp Pro Gly Tyr Thr Leu Glu Gln Gly Ser Ile
 560 565 570
 Ile Ile Glu Cys Val Asp Pro His Asp Pro Gln Trp Asn Glu Thr
 575 580 585
 Glu Pro Ala Cys Arg Ala Val Cys Ser Gly Glu Ile Thr Asp Ser
 590 595 600
 Ala Gly Val Val Leu Ser Pro Asn Trp Pro Glu Pro Tyr Gly Arg
 605 610 615
 Gly Gln Asp Cys Ile Trp Gly Val His Val Glu Glu Asp Lys Arg
 620 625 630
 Ile Met Leu Asp Ile Arg Val Leu Arg Ile Gly Pro Gly Asp Val
 635 640 645
 Leu Thr Phe Tyr Asp Gly Asp Asp Leu Thr Ala Arg Val Leu Gly
 650 655 660
 Gln Tyr Ser Gly Pro Arg Ser His Phe Lys Leu Phe Thr Ser Met
 665 670 675
 Ala Asp Val Thr Ile Gln Phe Gln Ser Asp Pro Gly Thr Ser Val
 680 685 690
 Leu Gly Tyr Gln Gln Gly Phe Val Ile His Phe Phe Glu Val Pro
 695 700 705
 Arg Asn Asp Thr Cys Pro Glu Leu Pro Glu Ile Pro Asn Gly Trp
 710 715 720
 Lys Ser Pro Ser Gln Pro Glu Leu Val His Gly Thr Val Val Thr
 725 730 735
 Tyr Gln Cys Tyr Pro Gly Tyr Gln Val Val Gly Ser Ser Val Leu
 740 745 750
 Met Cys Gln Trp Asp Leu Thr Trp Ser Glu Asp Leu Pro Ser Cys
 755 760 765
 Gln Arg Val Thr Ser Cys His Asp Pro Gly Asp Val Glu His Ser
 770 775 780
 Arg Arg Leu Ile Ser Ser Pro Lys Phe Pro Val Gly Ala Thr Val
 785 790 795
 Gln Tyr Ile Cys Asp Gln Gly Phe Val Leu Met Gly Ser Ser Ile
 800 805 810
 Leu Thr Cys His Asp Arg Gln Ala Gly Ser Pro Lys Trp Ser Asp
 815 820 825
 Arg Ala Pro Lys Cys Leu Leu Glu Gln Leu Lys Pro Cys His Gly
 830 835 840
 Leu Ser Ala Pro Glu Asn Gly Ala Arg Ser Pro Glu Lys Gln Leu
 845 850 855
 His Pro Ala Gly Ala Thr Ile His Phe Ser Cys Ala Pro Gly Tyr
 860 865 870
 Val Leu Lys Gly Gln Ala Ser Ile Lys Cys Val Pro Gly His Pro
 875 880 885
 Ser His Trp Ser Asp Pro Pro Pro Ile Cys Arg Ala Ala Ser Leu

WO 01/58353

PCT/US01/19862

```

      890      895      900
Asp Gly Phe Tyr Asn Ser Arg Ser Leu Asp Val Ala Lys Ala Pro
      905      910      915
Ala Ala Ser Ser Thr Leu Asp Ala Ala His Ile Ala Ala Ala Ile
      920      925      930
Phe Leu Pro Leu Val Ala Met Val Leu Leu Val Gly Gly Val Tyr
      935      940      945
Phe Tyr Phe Ser Arg Leu Gln Gly Lys Ser Ser Leu Gln Leu Pro
      950      955      960
Arg Pro Arg Pro Arg Pro Tyr Asn Arg Ile Thr Ile Glu Ser Ala
      965      970      975
Phe Asp Asn Pro Thr Tyr Glu Thr Gly Ser Leu Ser Phe Ala Gly
      980      985      990
Asp Glu Arg Ile

```

```

<210> 3
<211> 212
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc.feature
<223> Incyte ID No: 4362432CD1

```

```

<400> 3
Met Leu Ser Ser Val Val Phe Trp Gly Leu Ile Ala Leu Ile Gly
 1      5      10      15
Thr Ser Arg Gly Ser Tyr Pro Phe Ser His Ser Met Lys Pro His
 20      25      30
Leu His Pro Arg Leu Tyr His Gly Cys Tyr Gly Asp Ile Met Thr
 35      40      45
Met Lys Thr Ser Gly Ala Thr Cys Asp Ala Asn Ser Val Met Asn
 50      55      60
Cys Gly Ile Arg Gly Ser Glu Met Phe Ala Glu Met Asp Leu Arg
 65      70      75
Ala Ile Lys Pro Tyr Gln Thr Leu Ile Lys Glu Val Gly Gln Arg
 80      85      90
His Cys Val Asp Pro Ala Val Ile Ala Ala Ile Ile Ser Arg Glu
 95      100      105
Ser His Gly Gly Ser Val Leu Gln Asp Gly Trp Asp His Arg Gly
110      115      120
Leu Lys Phe Gly Leu Met Gln Leu Asp Lys Gln Thr Tyr His Pro
125      130      135
Val Gly Ala Trp Asp Ser Lys Glu His Leu Ser Gln Ala Thr Gly
140      145      150
Ile Leu Thr Glu Arg Ile Lys Ala Ile Gln Lys Lys Phe Pro Thr
155      160      165
Trp Ser Val Ala Glu His Leu Lys Gly Gly Leu Ser Ala Phe Lys
170      175      180
Ser Gly Ile Glu Ala Ile Ala Thr Pro Ser Asp Ile Asp Asn Asp
185      190      195
Phe Val Asn Asp Ile Ile Ala Arg Ala Lys Phe Tyr Lys Arg Gln
200      205      210
Ser Phe

```

```

<210> 4
<211> 308
<212> PRT
<213> Homo sapiens
<220>

```


WO 01/58353

PCT/US01/19862

<221> misc_feature

<223> Incyte ID No: 5308104CD1

<400> 4

```

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

```

<210> 5

<211> 328

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3092736CD1

<400> 5

```

Met Trp Cys Ser Phe Leu Ala Pro Val Ser Ser Ser Cys Leu Cys
1      5      10      15
Trp Val Trp Ala Cys Trp Gly Gln Gly His Cys Cys Gln Arg Gly
20     25     30
Thr Asp Phe Leu Met Val Leu Pro Lys Val Asn Val Gly Asp Thr
35     40     45

```

WO 01/58353

PCT/US01/19862

```

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

```

<210> 6

<211> 69

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3580257Cbl

<400> 6

```

Met Ala Met Ala Val Asp Val Ala Ala Ser Ala Asp Gly Val Leu
1 5 10 15
Ala Val Ala Met Glu Ala Thr Asp Met Ala Leu Ala Leu Glu Ala
20 25 30
Thr Asp Met Ala Leu Ala Leu Glu Ala Thr Asp Met Ala Leu Ala
35 40 45
Leu Glu Ala Met Asp Met Ala Ala Ala Ala His Arg Thr Met Glu
50 55 60
Asp Thr Asp Ser Leu Ala Phe Ile Lys
65

```

<210> 7

<211> 156

<212> PRT

WO 01/58353

PCT/US01/19862

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3634758CDi

<400> 7

```

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

```

<210> 6

<211> 463

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4027923CDi

<400> 8

```

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

```

WO 01/58353

PCT/US01/19862

```

170      175      180
Tyr Leu Lys Asp Trp His Leu Cys Arg Asp Phe Pro Val Glu Asp
185      190      195
Val Phe Thr Leu Pro Val Tyr Phe Ser Ser Asp Trp Leu Asn Glu
200      205      210
Phe Trp Asp Ala Leu Asp Val Asp Asp Tyr Arg Phe Val Tyr Ala
215      220      225
Gly Pro Ala Gly Ser Trp Ser Pro Phe His Ala Asp Ile Phe Arg
230      235      240
Ser Phe Ser Trp Ser Val Asn Val Cys Gly Arg Lys Lys Trp Leu
245      250      255
Leu Phe Pro Pro Gly Gln Glu Glu Ala Leu Arg Asp Arg His Gly
260      265      270
Asn Leu Pro Tyr Asp Val Thr Ser Pro Ala Leu Cys Asp Thr His
275      280      285
Leu His Pro Arg Asn Gln Leu Ala Gly Pro Pro Leu Glu Ile Thr
290      295      300
Gln Glu Ala Gly Glu Met Val Phe Val Pro Ser Gly Trp His His
305      310      315
Gln Val His Asn Leu Asp Asp Thr Ile Ser Ile Asn His Asn Trp
320      325      330
Val Asn Gly Phe Asn Leu Ala Asn Met Trp Arg Phe Leu Gln Gln
335      340      345
Glu Leu Cys Ala Val Gln Glu Glu Val Ser Glu Trp Arg Asp Ser
350      355      360
Met Pro Asp Trp His His His Cys Gln Val Ile Met Arg Ser Cys
365      370      375
Ser Gly Ile Asn Phe Glu Glu Phe Tyr His Phe Leu Lys Val Ile
380      385      390
Ala Glu Lys Arg Leu Leu Val Leu Arg Glu Ala Ala Ala Glu Asp
395      400      405
Gly Ala Gly Leu Gly Phe Glu Gln Ala Ala Phe Asp Val Gly Arg
410      415      420
Ile Thr Glu Val Leu Ala Ser Leu Val Ala His Pro Asp Phe Glu
425      430      435
Arg Val Asp Thr Ser Ala Phe Ser Pro Gln Pro Lys Glu Leu Leu
440      445      450
Gln Gln Leu Arg Glu Ala Val Asp Ala Ala Ala Ala Pro
455      460

```

<210> 9
<211> 648
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4348533CD1

<400> 9
Met Glu Lys Ala Arg Arg Gly Gly Asp Gly Val Pro Arg Gly Pro
1 5 10 15
Val Leu His Ile Val Val Val Gly Phe His His Lys Lys Gly Cys
20 25 30
Gln Val Glu Phe Ser Tyr Pro Pro Leu Ile Pro Gly Asp Gly His
35 40 45
Asp Ser His Thr Leu Pro Glu Glu Trp Lys Tyr Leu Pro Phe Leu
50 55 60
Ala Leu Pro Asp Gly Ala His Asn Tyr Gln Glu Asp Thr Val Phe
65 70 75
Phe His Leu Pro Pro Arg Asn Gly Asn Gly Ala Thr Val Phe Gly
80 85 90
Ile Ser Cys Tyr Arg Gln Ile Glu Ala Lys Ala Leu Lys Val Arg

WO 01/58353

PCT/US01/19862

	95		100		105
Gln	Ala	Asp	Ile	Thr	Arg
Leu	Ser	Lys	Leu	Pro	Leu
Leu	Ile	Thr	His	Ala	Tyr
Ser	Ile	Leu	Lys	Glu	Leu
Gly	Ala	Ser	Leu	Glu	Ser
Arg	Asp	Leu	Val	Leu	His
Lys	Leu	Ile	Leu	Leu	Glu
Val	Asn	Lys	Leu	Val	Gly
Pro	Gly	Met	Ile	Glu	His
Pro	Arg	Lys	Ser	Met	Ser
Pro	Cys	Ala	Asp	Asp	Phe
His	Thr	Asn	Leu	Gly	Thr
Gly	Glu	Asp	Ala	Ala	Met
Glu	Asp	Ser	Ser	Lys	Gly
Leu	Lys	Pro	Pro	Ser	Arg
Trp	Glu	Thr	Leu	Asp	Pro
Glu	Arg	Glu	Gln	Leu	Gly
Asp	Ser	Val	Pro	Ser	Glu
Ala	Asn	Thr	Gly	Glu	Val
Leu	Glu	Glu	Asp	Gln	Tyr
Gly	Tyr	Leu	Cys	Leu	Pro
Leu	Ser	Asp	Val	Thr	Val
Ile	Leu	Phe	Arg	Gln	Lys
Val	Glu	Glu	Ala	Leu	Ile
Leu	Leu	Asn	Pro	Thr	Thr
Val	Arg	His	Val	Thr	Glu
Thr	Gly	Trp	Glu	Gly	Asp
Val	Tyr	Ile	His	Ala	Leu
Glu	Lys	Ile	Leu	Ser	Asp
Lys	Asn	Thr	His	Asn	Tyr
Ala	Leu	Ala	Glu	Ile	Asn

WO 01/98353

PCT/US01/19862

```

Ser Val Ser Asp Met Lys Leu Arg Phe Ser His Ser Val Gln Asn
575 580 585
Ser Glu Arg Gly Lys Lys Ile Gly Asn Val Met Val Thr Thr Ser
590 595 600
Arg Asn Val Val Gln Thr Gly Lys Ala Val Gly Gln Ser Val Gly
605 610 615
Gly Ala Phe Ser Ser Ala Lys Thr Ala Met Ser Ser Trp Leu Ser
620 625 630
Thr Phe Thr Thr Ser Thr Ser Gln Ser Leu Thr Glu Pro Pro Asp
635 640 645
Glu Lys Pro

```

```

<210> 10
<211> 130
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4521857CD1

```

```

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

```

```

<210> 11
<211> 279
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4722253CD1

```

```

<400> 11
Met Gly Arg Gly Leu Arg Trp Trp Gly Gly Arg Gly Arg Arg His
1 5 10 15
Gly Gln Ala Pro Glu Trp Gly Pro Leu Val Gly Ala Arg Leu Lys
20 25 30
Gly Val Ala Arg Ala Ala Ser Leu Val Gly Arg Arg Arg Ala Gly
35 40 45
Thr Gly Met Ala Leu Leu Leu Cys Leu Val Cys Leu Thr Ala Ala
50 55 60
Leu Ala His Gly Cys Leu His Cys His Ser Asn Phe Ser Lys Lys
65 70 75
Phe Ser Phe Tyr Arg His His Val Asn Phe Lys Ser Trp Trp Val

```

WO 01/98353

PCT/US01/19862

```

      80      85      90
Gly Asp Ile Pro Val Ser Gly Ala Leu Leu Thr Asp Trp Ser Asp
  95      100      105
Asp Thr Met Lys Glu Leu His Leu Ala Ile Pro Ala Lys Ile Thr
  110      115      120
Arg Glu Lys Leu Asp Gln Val Ala Thr Ala Val Tyr Gln Met Met
  125      130      135
Asp Gln Leu Tyr Gln Gly Lys Met Tyr Phe Pro Gly Tyr Phe Pro
  140      145      150
Asn Glu Leu Arg Asn Ile Phe Arg Glu Gln Val His Leu Ile Gln
  155      160      165
Asn Ala Ile Ile Glu Ser Arg Ile Asp Cys Gln His Arg Cys Gly
  170      175      180
Ile Phe Gln Tyr Glu Thr Ile Ser Cys Asn Asn Cys Thr Asp Ser
  185      190      195
His Val Ala Cys Phe Gly Tyr Asn Cys Glu Ser Ser Ala Gln Trp
  200      205      210
Lys Ser Ala Val Gln Gly Leu Leu Asn Tyr Ile Asn Asn Trp His
  215      220      225
Lys Gln Asp Thr Ser Met Arg Pro Arg Ser Ser Ala Phe Ser Trp
  230      235      240
Pro Gly Thr His Arg Ala Thr Pro Ala Phe Leu Val Ser Pro Ala
  245      250      255
Leu Arg Cys Leu Glu Pro Pro His Leu Ala Asn Leu Thr Leu Glu
  260      265      270
Asp Ala Ala Glu Cys Leu Lys Gln His
  275

<210> 12
<211> 458
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Trcyle ID No: 4878134CD1

<400> 12
Met Pro Thr Ile Leu Trp Leu Met Asp Trp Ser Asp Met Asn Ser
  1      5      10      15
Asn Leu Asp Leu Leu Ala Leu Leu Gly Leu Gly Ile Ser Ser Phe
  20      25      30
Val Leu Ile Thr Gly Cys Ala Asn Met Leu Leu Met Ala Ala Leu
  35      40      45
Trp Gly Leu Tyr Met Ser Leu Val Asn Val Gly His Val Trp Tyr
  50      55      60
Ser Phe Gly Trp Glu Ser Gln Leu Leu Glu Thr Gly Phe Leu Gly
  65      70      75
Ile Phe Leu Cys Pro Leu Trp Thr Leu Ser Arg Leu Pro Gln His
  80      85      90
Thr Pro Thr Ser Arg Ile Val Leu Trp Gly Phe Arg Trp Leu Ile
  95      100      105
Phe Arg Ile Met Leu Gly Ala Gly Leu Ile Lys Ile Arg Gly Asp
  110      115      120
Arg Cys Trp Arg Asp Leu Thr Cys Met Asp Phe His Tyr Glu Thr
  125      130      135
Gln Pro Met Pro Asn Pro Val Ala Tyr Tyr Leu His His Ser Pro
  140      145      150
Trp Trp Phe His Arg Phe Glu Thr Leu Ser Asn His Phe Ile Glu
  155      160      165
Leu Leu Val Pro Phe Phe Leu Phe Leu Gly Arg Arg Ala Cys Ile
  170      175      180
Ile His Gly Val Leu Gln Ile Leu Phe Gln Ala Val Leu Ile Val

```

WO 01/98353

PCT/US01/19862

```

185      190      195
Ser Gly Asn Leu Ser Phe Leu Asn Trp Leu Thr Met Val Pro Ser
200      205      210
Leu Ala Cys Phe Asp Asp Ala Thr Leu Gly Phe Leu Phe Pro Ser
215      220      225
Gly Pro Gly Ser Leu Lys Asp Arg Val Leu Gln Met Gln Arg Asp
230      235      240
Ile Arg Gly Ala Arg Pro Glu Pro Arg Phe Gly Ser Val Val Arg
245      250      255
Arg Ala Ala Asn Val Ser Leu Gly Val Leu Leu Ala Trp Leu Ser
260      265      270
Val Pro Val Val Leu Asn Leu Leu Ser Ser Arg Gln Val Met Asn
275      280      285
Thr His Phe Asn Ser Leu His Ile Val Asn Thr Tyr Gly Ala Phe
290      295      300
Gly Ser Ile Thr Lys Glu Arg Ala Glu Val Ile Leu Gln Gly Thr
305      310      315
Ala Ser Ser Asn Ala Ser Ala Pro Asp Ala Met Trp Glu Asp Tyr
320      325      330
Glu Phe Lys Cys Lys Pro Gly Asp Pro Ser Arg Arg Pro Cys Leu
335      340      345
Ile Ser Pro Tyr His Tyr Arg Leu Asp Trp Leu Met Trp Phe Ala
350      355      360
Ala Phe Gln Thr Tyr Glu His Asn Asp Trp Ile Ile His Leu Ala
365      370      375
Gly Lys Leu Leu Ala Ser Asp Ala Glu Ala Leu Ser Leu Leu Ala
380      385      390
His Asn Pro Phe Ala Gly Arg Pro Pro Pro Arg Trp Val Arg Gly
395      400      405
Glu His Tyr Arg Tyr Lys Phe Ser Arg Pro Gly Gly Arg His Ala
410      415      420
Ala Glu Gly Lys Trp Trp Val Arg Lys Arg Ile Gly Ala Tyr Phe
425      430      435
Pro Pro Leu Ser Leu Glu Glu Leu Arg Pro Tyr Phe Arg Asp Arg
440      445      450
Gly Trp Pro Leu Pro Gly Pro Leu
455

```

<210> 13

<211> 1/3

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5050133CD1

<400> 13

```

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

```


WO 01/98353

PCT/US01/19862

```

110      115      120
Arg Trp Leu Gly Ser Ala Gly Leu Pro Ser Phe Arg Asn Lys Ser
125      130      135
Pro Ala Pro Val Pro Val Thr Tyr Glu Leu Pro Thr Leu Tyr Arg
140      145      150
Thr Glu Asp Tyr Phe Pro Val Asp Ala Gly Glu Ala Gln His His
155      160      165
Pro Arg Thr Cys Pro Arg Pro Leu
170

```

```

<210> 14
<211> 335
<212> EST
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5630124C01

```

```

<400> 14
Met Gly Ala Ser Ser Ser Ala Leu Ala Arg Leu Gly Leu Pro
1      5      10      15
Ala Arg Pro Trp Pro Arg Trp Leu Gly Val Ala Ala Leu Gly Leu
20      25      30
Ala Ala Val Ala Leu Gly Thr Val Ala Trp Arg Arg Ala Trp Pro
35      40      45
Arg Arg Arg Arg Arg Leu Gln Gln Val Gly Thr Val Ala Lys Leu
50      55      60
Trp Ile Tyr Pro Val Lys Ser Cys Lys Gly Val Pro Val Ser Glu
65      70      75
Ala Glu Cys Thr Ala Met Gly Leu Arg Ser Gly Asn Leu Arg Asp
80      85      90
Arg Phe Trp Leu Val Ile Lys Glu Asp Gly His Met Val Thr Ala
95      100      105
Arg Gln Glu Pro Arg Leu Val Leu Ile Ser Ile Ile Tyr Glu Asn
110      115      120
Asn Cys Leu Ile Phe Arg Ala Pro Asp Met Asp Gln Leu Val Leu
125      130      135
Pro Ser Lys Gln Pro Ser Ser Asn Lys Leu His Asn Cys Arg Ile
140      145      150
Phe Gly Leu Asp Ile Lys Gly Arg Asp Cys Gly Asn Glu Ala Ala
155      160      165
Lys Trp Phe Thr Asn Phe Leu Lys Thr Glu Ala Tyr Arg Leu Val
170      175      180
Gln Phe Glu Thr Asn Met Lys Gly Arg Thr Ser Arg Lys Leu Leu
185      190      195
Pro Thr Leu Asp Gln Asn Phe Gln Val Ala Tyr Pro Asp Tyr Cys
200      205      210
Pro Leu Leu Ile Met Thr Asp Ala Ser Leu Val Asp Leu Asn Thr
215      220      225
Arg Met Glu Lys Lys Met Lys Met Glu Asn Phe Arg Pro Asn Ile
230      235      240
Val Val Thr Gly Cys Asp Ala Phe Glu Glu Asp Thr Trp Asp Glu
245      250      255
Leu Leu Ile Gly Ser Val Glu Val Lys Lys Val Met Ala Cys Pro
260      265      270
Arg Cys Ile Leu Thr Thr Val Asp Pro Asp Thr Gly Val Ile Asp
275      280      285
Arg Lys Gln Pro Leu Asp Thr Leu Lys Ser Tyr Arg Leu Cys Asp
290      295      300
Pro Ser Glu Arg Glu Leu Tyr Lys Leu Ser Pro Leu Phe Gly Ile
305      310      315
Tyr Tyr Ser Val Glu Lys Ile Gly Ser Leu Arg Val Gly Asp Pro

```

WO 01/98353

PCT/US01/19862

Val Tyr Arg Met Val 320
335

<210> 15
<211> 71
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5677286CD1

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

<210> 16
<211> 148
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6436791CD1

<400> 16
Met Leu Pro Arg Gly Leu Lys Met Ala Pro Arg Gly Lys Arg Leu
1 5 10 15
Ser Ser Thr Pro Leu Glu Ile Leu Phe Phe Leu Asn Gly Trp Tyr
20 25 30
Asn Ala Thr Tyr Phe Leu Leu Glu Leu Phe Ile Phe Leu Tyr Lys
35 40 45
Gly Val Leu Leu Pro Tyr Pro Thr Ala Asn Leu Val Leu Asp Val
50 55 60
Val Met Leu Leu Leu Tyr Leu Gly Ile Glu Val Ile Arg Leu Phe
65 70 75
Phe Gly Thr Lys Gly Asn Leu Cys Gln Arg Lys Met Pro Leu Ser
80 85 90
Ile Ser Val Ala Leu Thr Phe Pro Ser Ala Met Met Ala Ser Tyr
95 100 105
Tyr Leu Leu Leu Gln Thr Tyr Val Leu Arg Leu Glu Ala Ile Met
110 115 120
Asn Gly Ile Leu Leu Phe Phe Cys Gly Ser Glu Leu Leu Leu Glu
125 130 135
Val Leu Thr Leu Ala Ala Phe Ser Ser Met Asp Thr Ile
140 145

<210> 17
<211> 231
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature

WO 01/98353

PCT/US01/19862

<223> Incyte ID No: 1820972CD1

<400> 17

```

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

```

<210> 18

<211> 716

<212> PRT

<213> Homo sapiens

<220>

<221> misc_features

<223> Incyte ID No: 3285805CD1

<400> 18

```

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

```

WO 01/98353

PCT/US01/19862

125	130	135
Met Lys Thr Leu	Gln Val Lys Asp Gln	Ala Ile Thr Thr Arg Val
140	145	150
Gln Val Arg Asn	Leu Val Tyr Thr Val	Lys Ile Asn Pro Thr Leu
155	160	165
Glu Leu Ser Ser	Gly Phe Arg Lys Glu	Leu Leu Asp Ile Ser Asp
170	175	180
Arg Leu Glu Asn	Asn Gln Thr Arg Met	Ala Thr Tyr Leu Ala Glu
185	190	195
Leu Leu Val Leu	Asn Tyr Gly Thr His	Val Thr Thr Ser Val Asp
200	205	210
Ala Gly Ala Ala	Leu Ile Gln Glu Asp	His Leu Arg Ala Ser Phe
215	220	225
Leu Gln Asp Ser	Gln Ser Ser Arg Ser	Ala Val Thr Ala Ser Ala
230	235	240
Gly Leu Ala Phe	Gln Asn Thr Val Asn	Phe Lys Phe Glu Glu Asn
245	250	255
Tyr Thr Ser Gln	Asn Val Leu Thr Lys	Ser Tyr Leu Ser Asn Arg
260	265	270
Thr Asn Ser Arg	Val Gln Ser Ile Gly	Gly Val Pro Phe Tyr Pro
275	280	285
Gly Ile Thr Leu	Gln Ala Trp Gln Gln	Gly Ile Thr Asn His Leu
290	295	300
Val Ala Ile Asp	Arg Ser Gly Leu Pro	Leu His Phe Phe Ile Asn
305	310	315
Pro Asn Met Leu	Pro Asp Leu Pro Gly	Pro Leu Val Lys Lys Val
320	325	330
Ser Lys Thr Val	Gln Thr Ala Val Lys	Arg Tyr Tyr Thr Phe Asn
335	340	345
Thr Tyr Pro Gly	Cys Thr Asp Leu Asn	Ser Pro Asn Phe Asn Phe
350	355	360
Gln Ala Asn Thr	Asp Asp Gly Ser Cys	Glu Gly Lys Met Thr Asn
365	370	375
Phe Ser Phe Gly	Gly Val Tyr Gln Glu	Cys Thr Gln Leu Ser Gly
380	385	390
Asn Arg Asp Val	Leu Leu Cys Gln Lys	Leu Glu Gln Lys Asn Pro
395	400	405
Leu Thr Gly Asp	Phe Ser Cys Pro Ser	Gly Tyr Ser Pro Val His
410	415	420
Leu Leu Ser Gln	Ile His Glu Glu Gly	Tyr Asn His Leu Glu Cys
425	430	435
His Arg Lys Cys	Thr Leu Leu Val Phe	Cys Lys Thr Val Cys Gln
440	445	450
Asp Val Phe Gln	Val Ala Lys Ala Glu	Phe Arg Ala Phe Trp Cys
455	460	465
Val Ala Ser Ser	Gln Val Pro Glu Asn	Ser Gly Leu Leu Phe Gly
470	475	480
Gly Leu Phe Ser	Ser Lys Ser Ile Asn	Pro Met Thr Asn Ala Gln
485	490	495
Ser Cys Pro Ala	Gly Tyr Phe Pro Leu	Arg Leu Phe Glu Asn Leu
500	505	510
Lys Val Cys Val	Ser Gln Asp Tyr Glu	Leu Gly Ser Arg Phe Ala
515	520	525
Val Pro Phe Gly	Gly Phe Phe Ser Cys	Thr Val Gly Asn Pro Leu
530	535	540
Val Asp Pro Ala	Ile Ser Arg Asp Leu	Gly Ala Pro Ser Leu Lys
545	550	555
Lys Cys Pro Gly	Gly Phe Ser Gln His	Pro Ala Leu Ile Ser Asp
560	565	570
Gly Cys Gln Val	Ser Tyr Cys Val Lys	Ser Gly Leu Phe Thr Gly
575	580	585
Gly Ser Leu Pro	Pro Ala Arg Leu Pro	Pro Phe Thr Arg Pro Pro
590	595	600

WO 01/98353

PCT/US01/19862

```

Leu Met Ser Gln Ala Ala Thr Asn Thr Val Ile Val Thr Asn Ser
605 610 615
Glu Asn Ala Arg Ser Trp Ile Lys Asp Ser Gln Thr His Gln Trp
620 625 630
Arg Leu Gly Glu Pro Ile Glu Leu Arg Arg Ala Met Asn Val Ile
635 640 645
His Gly Asp Gly Gly Gly Leu Ser Gly Gly Ala Ala Ala Gly Val
650 655 660
Thr Val Gly Val Thr Thr Ile Leu Ala Val Val Ile Thr Leu Ala
665 670 675
Ile Tyr Gly Thr Arg Lys Phe Lys Lys Lys Ala Tyr Gln Ala Ile
680 685 690
Glu Glu Arg Gln Ser Leu Val Pro Gly Thr Ala Ala Thr Gly Asp
695 700 705
Thr Thr Tyr Gln Gln Gly Gln Ser Pro Ala
710 715

```

<210> 19

<211> 519

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3506590CD1

<400> 19

```

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

```

WO 01/98353

PCT/US01/19862

Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg
 275 280 285
 Cys Pro Glu Pro Lys Ser Cys Asp Thr Pro Pro Pro Cys Pro Arg
 290 295 300
 Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe
 305 310 315
 Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu
 320 325 330
 Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val
 335 340 345
 Gln Phe Lys Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
 350 355 360
 Thr Lys Leu Arg Glu Gln Tyr Asn Ser Thr Phe Arg Val Val
 365 370 375
 Ser Val Leu Thr Val Leu His Gln Asp Trp Leu Asn Gly Lys Glu
 380 385 390
 Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala Pro Ile Glu
 395 400 405
 Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro Gln Val
 410 415 420
 Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val
 425 430 435
 Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
 440 445 450
 Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Asn Thr
 455 460 465
 Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser
 470 475 480
 Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Ile Phe
 485 490 495
 Ser Cys Ser Val Met His Glu Ala Leu His Asn Arg Tyr Thr Gln
 500 505 510
 Lys Ser Leu Ser Leu Ser Pro Gly Lys
 515

<210> 20

<211> 172

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 003600CD1

<400> 20

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

WO 01/98353

PCT/US01/19862

Iys Glu Ile Thr Asp Tyr Val Asn Val Asn Pro Glu Arg His Lys
 140 145 150
 Pro Ser Phe Trp Tyr Phe Val Asn Pro Ala Leu Ser Glu Pro Ala
 155 160 165
 Glu Tyr Asp Gln Val Ala Met
 170

<210> 21
 <211> 314
 <212> ERT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1251534CD1

<400> 21
 Met Gly Leu Leu Asp Ser Glu Pro Gly Ser Val Leu Asn Val Val
 1 5 10 15
 Ser Thr Ala Leu Asn Asp Thr Val Glu Phe Tyr Arg Trp Thr Trp
 20 25 30
 Ser Ile Ala Asp Lys Arg Val Glu Asn Trp Pro Leu Met Gln Ser
 35 40 45
 Pro Trp Pro Thr Leu Ser Ile Ser Thr Leu Tyr Leu Leu Phe Val
 50 55 60
 Trp Leu Gly Pro Lys Trp Met Lys Asp Arg Glu Pro Phe Gln Met
 65 70 75
 Arg Leu Val Leu Ile Ile Tyr Asn Phe Gly Met Val Leu Leu Asn
 80 85 90
 Leu Phe Ile Phe Arg Glu Leu Phe Met Gly Ser Tyr Asn Ala Gly
 95 100 105
 Tyr Ser Tyr Ile Cys Gln Ser Val Asp Tyr Ser Asn Asn Val His
 110 115 120
 Glu Val Arg Ile Ala Ala Leu Trp Trp Tyr Phe Val Ser Lys
 125 130 135
 Gly Val Glu Tyr Leu Asp Thr Val Phe Phe Ile Leu Arg Lys Lys
 140 145 150
 Asn Asn Gln Val Ser Phe Leu His Val Tyr His His Cys Thr Met
 155 160 165
 Phe Thr Leu Trp Trp Ile Gly Ile Cys Trp Val Ala Gly Gly Gln
 170 175 180
 Ala Phe Phe Gly Ala Gln Leu Asn Ser Phe Ile His Val Ile Met
 185 190 195
 Tyr Ser Tyr Tyr Gly Leu Thr Ala Phe Gly Pro Trp Ile Gln Lys
 200 205 210
 Tyr Leu Trp Trp Lys Arg Tyr Leu Thr Met Leu Gln Leu Ile Gln
 215 220 225
 Phe His Val Thr Ile Gly His Thr Ala Leu Ser Leu Tyr Thr Asp
 230 235 240
 Cys Pro Phe Pro Lys Trp Met His Trp Ala Leu Ile Ala Tyr
 245 250 255
 Ile Ser Phe Ile Phe Leu Phe Leu Asn Phe Tyr Ile Arg Thr Tyr
 260 265 270
 Lys Glu Pro Lys Lys Pro Lys Ala Gly Lys Thr Ala Met Asn Gly
 275 280 285
 Ile Ser Ala Asn Gly Val Ser Lys Ser Glu Lys Gln Leu Met Ile
 290 295 300
 Glu Asn Gly Lys Lys Gln Lys Asn Gly Lys Ala Lys Gly Asp
 305 310

<210> 22
 <211> 542
 <212> ERT

WO 01/98353

PCT/US01/19862

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1402211CD1

<400> 22

```

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

```


WO 01/98353

PCT/US01/19862

```

Pro Glu Glu Asn Glu Thr Lys Lys Glu Ser Ser Val Pro Met Cys
425 430 435
Gln Gly Glu Leu Arg His Trp Lys Thr Gly His Tyr Thr Leu Ile
440 445 450
His Asp His Ser Lys Ala Glu Phe Ala Leu Asp Leu Ile Leu Tyr
455 460 465
Cys Gly Cys Glu Gly Thr Glu Pro Glu Tyr Gly Gly Phe Thr Ser
470 475 480
Tyr Ile Ala Lys Gly Glu Asp Glu Glu Leu Leu Thr Val Asn Pro
485 490 495
Glu Ser Asn Ser Leu Ala Leu Val Tyr Arg Asp Arg Glu Thr Leu
500 505 510
Lys Phe Val Lys His Ile Asn His Arg Ser Leu Glu Gln Lys Lys
515 520 525
Thr Phe Pro Asn Arg Thr Gly Phe Trp Asp Phe Ser Phe Ile Tyr
530 535 540
Tyr Glu

```

<210> 23

<211> 715

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1623474CD1

<400> 23

```

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

```

22/71

WO 01/98353

PCT/US01/19862

Pro Lys Lys Glu Lys Thr Ile Val Ile Arg Pro Pro Phe Leu Arg
 260 265 270
 Pro Glu Val Ser Asp Gly Gln Ile Thr Val Lys Ile Met Asp Asn
 275 280 285
 Gly Ile Gln Gly Glu Leu Arg Arg Thr Lys Ser Lys Gly Ser Leu
 290 295 300
 Glu Ile Thr Glu Ser Gln Ser Ala Asp Ala Glu Pro Pro Pro
 305 310 315
 Pro Lys Pro Asp Leu Ser Arg Tyr Thr Gly Leu Arg Thr His Leu
 320 325 330
 Gly Leu Ala Thr Asn Glu Asp Ser Ser Leu Leu Ala Lys Asp Ser
 335 340 345
 Pro Pro Thr Pro Thr Met Tyr Lys Tyr Arg Pro Gly Tyr Ser Ser
 350 355 360
 Ser Ser Thr Ser Ala Ala Met Pro His Ser Ser Ser Ala Lys Leu
 365 370 375
 Ser Arg Gly Asp Ser Leu Lys Glu Pro Thr Ser Ile Ala Glu Ser
 380 385 390
 Ser Arg His Pro Ser Tyr Arg Ser Glu Pro Ser Leu Glu Pro Glu
 395 400 405
 Ser Phe Arg Ser Thr Phe Gly Lys Ser Phe His Phe Asp Pro
 410 415 420
 Leu Ser Ser Gly Ser Arg Ser Ser Ser Leu Lys Ser Ala Gln Gly
 425 430 435
 Thr Gly Phe Glu Leu Gly Gln Leu Gln Ser Ile Arg Ser Glu Gly
 440 445 450
 Thr Thr Ser Thr Ser Tyr Lys Ser Leu Ala Asn Gln Thr Arg Asn
 455 460 465
 Gly Ser Leu Ser Tyr Asp Ser Leu Leu Thr Pro Ser Asp Ser Pro
 470 475 480
 Asp Phe Glu Ser Val Gln Ala Gly Pro Glu Pro Asp Pro Pro Leu
 485 490 495
 Gly Tyr Thr Ser Pro Phe Leu Ser Ala Arg Leu Ala Gln Gln Arg
 500 505 510
 Glu Ala Glu Arg His Pro Arg Leu Val Pro Thr Gly Pro Thr His
 515 520 525
 Arg Glu Pro Ser Pro Val Arg Tyr Asp Asn Leu Ser Arg His Ile
 530 535 540
 Val Ala Ser Leu Gln Glu Arg Gln Lys Leu Leu Arg Gln Ser Pro
 545 550 555
 Pro Leu Pro Gly Arg Glu Glu Glu Pro Gly Leu Gly Asp Ser Gly
 560 565 570
 Ile Gln Ser Thr Pro Gly Ser Gly His Ala Pro Arg Thr Ser Ser
 575 580 585
 Ser Ser Asp Asp Ser Lys Arg Ser Pro Leu Gly Lys Thr Pro Leu
 590 595 600
 Gly Arg Pro Ala Val Pro Arg Phe Gly Lys Pro Asp Gly Leu Arg
 605 610 615
 Gly Arg Gly Val Gly Ser Pro Glu Pro Gly Pro Thr Ala Pro Tyr
 620 625 630
 Leu Gly Arg Ser Met Ser Tyr Ser Ser Gln Lys Ala Gln Pro Gly
 635 640 645
 Val Ser Glu Thr Glu Glu Val Ala Leu Gln Pro Leu Leu Thr Pro
 650 655 660
 Lys Asp Glu Val Gln Leu Lys Thr Thr Tyr Ser Lys Ser Asn Gly
 665 670 675
 Gln Pro Lys Ser Leu Gly Ser Ala Ser Pro Gly Pro Gly Glu Pro
 680 685 690
 Pro Leu Ser Ser Pro Thr Arg Gly Gly Val Lys Lys Val Ser Gly
 695 700 705
 Val Gly Gly Thr Thr Tyr Glu Ile Ser Val
 710 715

WO 01/58353

PCT/US01/19862

<210> 24
 <211> 469
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1706443CD1

<400> 24
 Met Gly Arg Val Arg Arg Ile Tyr Pro Gln Leu Leu Leu Ala Leu
 1 5 10 15
 Leu Ile Gln Val His Tyr His Ile Gly Leu Asn Leu Pro Gly Cys
 20 25 30
 Val Ala Pro Pro Lys Asp Thr Lys Lys Gly Ala Gln Pro Ser Pro
 35 40 45
 Phe Val Pro Val Arg Trp Val Val Lys Val Val Lys Thr Leu Leu
 50 55 60
 Leu Arg Met Gly Cys Ser Tyr Glu Thr Thr Phe Leu Glu Asp Gln
 65 70 75
 Gly Gly Trp Glu Leu Met Gln Gln Val Glu Ser His His Arg Gly
 80 85 90
 Val Ala Leu Leu Ala Arg Ala Met Val Gln Tyr Ser Cys Gln Glu
 95 100 105
 Leu Cys Arg Ile Leu Tyr Leu Leu Ile Pro Leu Leu Glu Arg Gly
 110 115 120
 Asp Glu Lys His Arg Ile Thr Ala Thr Ala Phe Phe Val Glu Leu
 125 130 135
 Leu Gln Met Glu Gln Val Arg Arg Ile Pro Glu Glu Tyr Ser Leu
 140 145 150
 Gly Arg Met Ala Glu Gly Leu Ser His His Asp Pro Ile Met Lys
 155 160 165
 Val Leu Ser Ile Arg Gly Leu Val Ile Leu Ala Arg Arg Ser Glu
 170 175 180
 Lys Thr Ala Lys Val Lys Ala Leu Leu Pro Ser Met Val Lys Gly
 185 190 195
 Leu Lys Asn Met Asp Gly Met Leu Val Val Glu Ala Val His Asn
 200 205 210
 Leu Lys Ala Val Phe Lys Gly Arg Asp Gln Lys Leu Met Asp Ser
 215 220 225
 Ala Val Tyr Val Glu Met Leu Gln Ile Leu Leu Pro His Phe Ser
 230 235 240
 Asp Ala Arg Glu Asp Val Arg Ser Ser Cys Ile Asn Leu Tyr Gly
 245 250 255
 Lys Val Val Gln Lys Leu Arg Ala Pro Arg Thr Gln Ala Met Glu
 260 265 270
 Glu Gln Leu Val Ser Thr Leu Val Pro Leu Leu Thr Met Gln
 275 280 285
 Glu Gly Asn Ser Lys Val Ser Gln Lys Cys Val Lys Thr Leu Leu
 290 295 300
 Arg Cys Ser Tyr Phe Met Ala Trp Glu Leu Pro Lys Arg Ala Tyr
 305 310 315
 Ser Arg Lys Pro Trp Asp Asn Gln Gln Gln Thr Val Ala Lys Ile
 320 325 330
 Cys Lys Cys Leu Val Asn Thr His Arg Asp Ser Ala Phe Ile Phe
 335 340 345
 Leu Ser Gln Ser Leu Glu Tyr Ala Lys Asn Ser Arg Ala Ser Leu
 350 355 360
 Arg Lys Cys Ser Val Met Phe Ile Gly Ser Leu Val Pro Cys Met
 365 370 375
 Glu Ser Ile Met Thr Glu Asp Arg Leu Asn Glu Val Lys Ala Ala
 380 385 390
 Leu Asp Asn Leu Arg His Asp Pro Glu Ala Ser Val Cys Ile Tyr

WO 01/58353

PCT/US01/19862

```

395      400      405
Ala Ala Gln Val Gln Asp His Ile Leu Ala Ser Cys Trp Gln Asn
410      415      420
Ser Trp Leu Pro His Gly Asn Ser Trp Val Cys Tyr Ser Ala Thr
425      430      435
Thr His Arg Trp Ser Pro Ser Cys Glu Asn Leu Pro Thr Ser His
440      445      450
Gln Arg Arg Ser Trp Ile Met Gln Ala Leu Gly Ser Trp Lys Met
455      460      465
Ser Leu Lys Lys

```

```

<210> 25
<211> 274
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 174862/CD1

```

```

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

```

```

<210> 26
<211> 154

```

WO 01/98353

PCT/US01/19862

<212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1818332CDL

<400> 26
 Met Ala Gly Pro Val Lys Asp Arg Glu Ala Phe Gln Arg Leu Asn
 1 5 10 15
 Phe Leu Tyr Gln Ala Ala His Cys Val Leu Ala Gln Asp Pro Glu
 20 25 30
 Asn Gln Ala Leu Ala Arg Phe Tyr Cys Tyr Thr Glu Arg Thr Ile
 35 40 45
 Ala Lys Arg Leu Val Leu Arg Arg Asp Pro Ser Val Lys Arg Thr
 50 55 60
 Leu Cys Arg Gly Cys Ser Ser Leu Leu Val Pro Gly Leu Thr Cys
 65 70 75
 Thr Gln Arg Gln Arg Arg Cys Arg Gly Gln Arg Trp Thr Val Gln
 80 85 90
 Thr Cys Leu Thr Cys Gln Arg Ser Gln Arg Phe Leu Asn Asp Pro
 95 100 105
 Gly His Leu Leu Trp Gly Asp Arg Pro Glu Ala Gln Leu Gly Ser
 110 115 120
 Gln Ala Asp Ser Lys Pro Leu Gln Pro Leu Pro Asn Thr Ala His
 125 130 135
 Ser Ile Ser Asp Arg Leu Pro Glu Glu Lys Met Gln Thr Gln Gly
 140 145 150
 Ser Ser Asn Gln

<210> 27
 <211> 102
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1822632CDL

<400> 27
 Met Lys Phe Asp Trp Val Met Gly Leu Arg Ser Ile Thr Leu Lys
 1 5 10 15
 Asn Ser Ser Thr Gly Arg Gly Asp Gly Pro Lys Gln His Leu Gln
 20 25 30
 Ala Asp Pro Met Leu Ile Ile Arg Ala Arg Thr Leu Ser Leu Ser
 35 40 45
 Val Ser Leu Ser Val Ser Pro Leu Gly Leu Thr Pro His Trp Thr
 50 55 60
 Pro Leu His Pro Cys Pro Ser His Asn Thr Ala Ala Val Ser Ser
 65 70 75
 Ala Cys Leu Trp Glu Ser Pro Leu Phe Ser Ser Val Phe Phe Ser
 80 85 90
 Ser Cys Pro Ile Thr Pro Cys Thr Ser Pro Phe Pro

<210> 28
 <211> 113
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature

WO 01/98353

PCT/US01/19862

<223> Incyte ID No: 1832219CD1

<400> 28

```

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

```

<210> 29

<211> 313

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1899010CD1

<400> 29

```

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

```

WO 01/58353

PCT/US01/19862

```

      245      250      255
Phe Gly His Phe Lys Phe Cys Ile Thr Leu Phe Gly Gly Tyr Val
      260      265      270
Leu Phe Lys Asp Pro Leu Ser Ile Asn Gln Ala Leu Gly Ile Leu
      275      280      285
Cys Thr Leu Phe Gly Ile Leu Ala Tyr Thr His Phe Lys Leu Ser
      290      295      300
Glu Gln Glu Gly Ser Arg Ser Lys Leu Ala Gln Arg Pro
      305      310

```

<210> 30

<211> 195

<212> PRT

<213> Homo, sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2008758CD1

<400> 30

```

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

```

<210> 31

<211> 350

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2070994CD1

<400> 31

```

Met Asn Leu Arg Gly Leu Phe Gln Asp Phe Asn Pro Ser Lys Phe
  1      5      10      15
Leu Ile Tyr Ala Cys Leu Leu Leu Phe Ser Val Leu Leu Ala Leu
  20      25      30
Arg Leu Asp Gly Ile Ile Gln Trp Ser Tyr Trp Ala Val Phe Ala

```

WO 01/98353

PCT/US01/19862

```

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

```

<210> 32

<211> 360

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2193240CD3

<400> 32

```

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

```


WO 01/98353

PCT/US01/19862

```

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

```

<210> 33

<211> 559

<212> FRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2235177CD1

<400> 33

```

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

```

30/71

WO 01/98353

PCT/US01/19862

Ser Gly Ala Lys Ser Cys Phe Cys Ile Leu Ser Tyr Leu Pro Trp
 95 100 105
 Phe Glu Val Phe Tyr Lys Leu Leu Asn Ile Leu Ala Asp Tyr Thr
 110 115 120
 Thr Lys Arg Gln Gln Asn Gln Trp Asn Glu Leu Leu Glu Thr Leu
 125 130 135
 His Lys Leu Pro Ile Pro Asp Pro Gly Val Ser Val His Leu Ser
 140 145 150
 Val His Ser Tyr Phe Thr Val Pro Asp Thr Arg Glu Leu Pro Ser
 155 160 165
 Ile Pro Glu Asn Arg Asn Leu Thr Glu Tyr Phe Val Ala Val Asp
 170 175 180
 Val Asn Asn Met Leu His Leu Tyr Ala Ser Met Leu Tyr Glu Arg
 185 190 195
 Arg Ile Leu Ile Ile Cys Ser Lys Leu Ser Thr Leu Thr Ala Cys
 200 205 210
 Ile His Gly Ser Ala Ala Met Leu Tyr Pro Met Tyr Trp Gln His
 215 220 225
 Val Tyr Ile Pro Val Leu Pro Pro His Leu Leu Asp Tyr Cys Cys
 230 235 240
 Ala Pro Met Pro Tyr Leu Ile Gly Ile His Leu Ser Leu Met Glu
 245 250 255
 Lys Val Arg Asn Met Ala Leu Asp Asp Val Val Ile Leu Asn Val
 260 265 270
 Asp Thr Asn Thr Leu Glu Thr Pro Phe Asp Asp Leu Gln Ser Leu
 275 280 285
 Pro Asn Asp Val Ile Ser Ser Leu Lys Asn Arg Leu Lys Lys Val
 290 295 300
 Ser Thr Thr Thr Gly Asp Gly Val Ala Arg Ala Phe Leu Lys Ala
 305 310 315
 Gln Ala Ala Phe Phe Gly Ser Tyr Arg Asn Ala Leu Lys Ile Glu
 320 325 330
 Pro Glu Glu Pro Ile Thr Phe Cys Glu Glu Ala Phe Val Ser His
 335 340 345
 Tyr Arg Ser Gly Ala Met Arg Gln Phe Leu Gln Asn Ala Thr Gln
 350 355 360
 Leu Gln Leu Phe Lys Gln Phe Ile Asp Gly Arg Leu Asp Leu Leu
 365 370 375
 Asn Ser Gly Glu Gly Phe Ser Asp Val Phe Glu Glu Glu Ile Asn
 380 385 390
 Met Gly Glu Tyr Ala Gly Ser Asp Lys Leu Tyr His Gln Trp Leu
 395 400 405
 Ser Thr Val Arg Lys Gly Ser Gly Ala Ile Leu Asn Thr Val Lys
 410 415 420
 Thr Lys Ala Asn Pro Ala Met Lys Thr Val Tyr Lys Phe Ala Lys
 425 430 435
 Asp His Ala Lys Met Gly Ile Lys Glu Val Lys Asn Arg Leu Lys
 440 445 450
 Gln Lys Asp Ile Ala Glu Asn Gly Cys Ala Pro Thr Pro Glu Glu
 455 460 465
 Gln Leu Pro Lys Thr Ala Pro Ser Pro Leu Val Glu Ala Lys Asp
 470 475 480
 Pro Lys Leu Arg Glu Asp Arg Arg Pro Ile Thr Val His Phe Gly
 485 490 495
 Gln Val Arg Pro Pro Arg Pro His Val Val Lys Arg Pro Lys Ser
 500 505 510
 Asn Ile Ala Val Glu Gly Arg Arg Thr Ser Val Pro Ser Pro Glu
 515 520 525
 Gln Asn Thr Ile Ala Thr Pro Ala Thr Leu His Ile Leu Gln Lys
 530 535 540
 Ser Ile Thr His Phe Ala Ala Lys Phe Pro Thr Arg Gly Trp Thr
 545 550 555
 Ser Ser Ser His

WO 01/58353

PCT/US01/19862

<210> 34
 <211> 198
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> misc_feature
 <223> Incyte ID No: 2416237CD1

 <400> 34
 Met Ala Leu Arg His Leu Ala Leu Leu Ala Gly Leu Leu Val Gly
 1 5 10 15
 Val Ala Ser Lys Ser Met Glu Asn Thr Ala Gln Leu Pro Glu Cys
 20 25 30
 Cys Val Asp Val Val Gly Val Asn Ala Ser Cys Pro Gly Ala Ser
 35 40 45
 Leu Cys Gly Pro Gly Cys Tyr Arg Arg Trp Asn Ala Asp Gly Ser
 50 55 60
 Ala Ser Cys Val Arg Cys Gly Asn Gly Thr Leu Pro Ala Tyr Asn
 65 70 75
 Gly Ser Glu Cys Arg Ser Phe Ala Gly Pro Gly Ala Pro Phe Pro
 80 85 90
 Met Asn Arg Ser Ser Gly Thr Pro Gly Arg Pro His Pro Gly Ala
 95 100 105
 Pro Arg Val Ala Ala Ser Leu Phe Leu Gly Thr Phe Phe Ile Ser
 110 115 120
 Ser Gly Leu Ile Leu Ser Val Ala Gly Phe Phe Tyr Leu Lys Arg
 125 130 135
 Ser Ser Lys Leu Pro Arg Ala Cys Tyr Arg Arg Asn Lys Ala Pro
 140 145 150
 Ala Leu Gln Pro Gly Glu Ala Ala Ala Met Ile Pro Pro Pro Gln
 155 160 165
 Ser Ser Val Arg Lys Pro Arg Tyr Val Arg Arg Glu Arg Pro Leu
 170 175 180
 Asp Arg Ala Thr Asp Pro Ala Ala Phe Pro Gly Glu Ala Arg Ile
 185 190 195
 Ser Asn Val

 <210> 35
 <211> 73
 <212> PRT
 <213> Homo sapiens

 <220>
 <221> misc_feature
 <223> Incyte ID No: 2461076CD1

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

 <210> 36

WO 01/98353

PCT/US01/19862

<211> 376

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 195/517CD1

<400> 36

```

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

```

<220> 37

WO 01/58353

PCT/US01/19862

<211> 216
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 866038CD1

<400> 37
 Met Met Tyr Trp Ile Val Phe Ala Phe Phe Thr Thr Ala Glu Thr
 1 5 10 15
 Leu Thr Asp Ile Val Leu Ser Trp Phe Pro Phe Tyr Phe Glu Leu
 20 25 30
 Lys Ile Ala Phe Val Ile Trp Leu Leu Ser Pro Tyr Thr Lys Gly
 35 40 45
 Ser Ser Val Leu Tyr Arg Lys Phe Val His Pro Thr Leu Ser Asn
 50 55 60
 Lys Glu Lys Glu Ile Asp Glu Tyr Ile Thr Glu Ala Arg Asp Lys
 65 70 75
 Ser Tyr Glu Thr Met Met Arg Val Gly Lys Arg Gly Leu Asn Leu
 80 85 90
 Ala Ala Asn Ala Ala Val Thr Ala Ala Ala Lys Gly Gln Gly Val
 95 100 105
 Leu Ser Glu Lys Leu Arg Ser Phe Ser Met Gln Asp Leu Thr Leu
 110 115 120
 Ile Arg Asp Glu Asp Ala Leu Pro Leu Gln Arg Pro Asp Gly Arg
 125 130 135
 Leu Arg Pro Ser Pro Gly Ser Leu Leu Asp Thr Ile Glu Asp Leu
 140 145 150
 Gly Asp Asp Pro Ala Leu Ser Leu Arg Ser Ser Thr Asn Pro Ala
 155 160 165
 Asp Ser Arg Thr Glu Ala Ser Glu Asp Asp Met Gly Asp Lys Ala
 170 175 180
 Pro Lys Arg Ala Lys Pro Ile Lys Lys Ala Pro Lys Ala Glu Pro
 185 190 195
 Leu Ala Ser Lys Thr Leu Lys Thr Arg Pro Lys Lys Lys Thr Ser
 200 205 210
 Gly Gly Gly Asp Ser Ala
 215

<210> 38
 <211> 233
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3669704CD1

<400> 38
 Met Ala Trp Thr Pro Leu Leu Leu Pro Leu Leu Thr Phe Cys Thr
 1 5 10 15
 Val Ser Glu Ala Ser Tyr Glu Leu Thr Gln Pro Pro Ser Val Ser
 20 25 30
 Val Ser Pro Gly Gln Thr Ala Arg Ile Thr Cys Ser Gly Asp Ala
 35 40 45
 Leu Pro Lys Lys Tyr Ala Tyr Trp Tyr Gln Gln Lys Ser Gly Gln
 50 55 60
 Ala Pro Val Leu Val Ile Tyr Glu Asp Asn Lys Arg Pro Ser Gly
 65 70 75
 Ile Pro Glu Arg Phe Phe Gly Ser Ser Gly Thr Met Ala Thr
 80 85 90
 Leu Thr Ile Ser Gly Ala Gln Val Glu Asp Glu Ala Asp Tyr Tyr

WO 01/58353

PCT/US01/19862

```

          95          100          105
Cys Tyr Ser Thr Asp Ser Ser Gly Asn Asp Arg Val Phe Gly Gly
          110          115          120
Gly Thr Lys Leu Thr Val Leu Gly Gln Pro Lys Ala Ala Pro Ser
          125          130          135
Val Thr Leu Phe Pro Pro Ser Ser Glu Glu Leu Gln Ala Asn Lys
          140          145          150
Ala Thr Leu Val Cys Leu Ile Ser Asp Phe Tyr Pro Gly Ala Val
          155          160          165
Thr Val Ala Trp Lys Ala Asp Ser Ser Pro Val Lys Ala Gly Val
          170          175          180
Glu Thr Thr Thr Pro Ser Lys Gln Ser Asn Asn Lys Tyr Ala Ala
          185          190          195
Ser Ser Tyr Leu Ser Leu Thr Pro Glu Gln Trp Lys Ser His Lys
          200          205          210
Ser Tyr Ser Cys Gln Val Thr His Glu Gly Ser Thr Val Glu Lys
          215          220          225
Thr Val Ala Pro Thr Glu Cys Ser
          230

```

<210> 39

<211> 163

<212> FRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1415179CD1

<400> 39

```

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

```

<210> 40

<211> 235

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1664792CD1

<400> 40

WO 01/98353

PCT/US01/19862

```

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

```

<210> 41

<211> 94

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2079396CD1

<400> 41

```

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

```

<210> 42

<211> 85

<212> PRT

<213> Homo sapiens

WO 01/98353

PCT/US01/19862

```

<220>
<221> misc_feature
<223> Incyte ID No: 5396119CD1

<400> 42
Met Ala Ser Asp Leu Asp Phe Ser Pro Pro Glu Val Pro Glu Pro
1      5      10      15
Thr Phe Leu Glu Asn Leu Leu Arg Tyr Gly Leu Phe Leu Gly Ala
20     25     30
Ile Phe Gln Leu Ile Cys Val Leu Ala Ile Ile Val Pro Ile Pro
35     40     45
Lys Ser His Glu Ala Glu Ala Glu Pro Ser Glu Pro Arg Ser Ala
50     55     60
Glu Val Thr Arg Lys Pro Lys Ala Ala Val Pro Ser Val Asn Lys
65     70     75
Arg Pro Lys Lys Glu Thr Lys Lys Lys Arg
80     85

<210> 43
<211> 901
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1403326CD1

<400> 43
Met Glu Ala Asn Gln Cys Pro Leu Val Val Glu Pro Ser Tyr Pro
1      5      10      15
Asp Leu Val Ile Asn Val Gly Glu Val Thr Leu Gly Glu Glu Asn
20     25     30
Arg Lys Lys Leu Gln Lys Ile Gln Arg Asp Gln Glu Lys Glu Arg
35     40     45
Val Met Arg Ala Ala Cys Ala Leu Leu Asn Ser Gly Gly Gly Val
50     55     60
Ile Arg Met Ala Lys Lys Val Glu His Pro Val Glu Met Gly Leu
65     70     75
Asp Leu Glu Gln Ser Leu Arg Glu Leu Ile Gln Ser Ser Asp Leu
80     85     90
Gln Ala Phe Phe Glu Thr Lys Gln Gln Gly Arg Cys Phe Tyr Ile
95     100    105
Phe Val Lys Ser Trp Ser Ser Gly Pro Phe Pro Glu Asp Arg Ser
110    115    120
Phe Lys Pro Arg Leu Cys Ser Leu Ser Ser Ser Leu Tyr Arg Arg
125    130    135
Ser Glu Thr Ser Val Arg Ser Met Asp Ser Arg Glu Ala Phe Cys
140    145    150
Phe Leu Lys Thr Lys Arg Lys Pro Lys Ile Leu Glu Glu Gly Pro
155    160    165
Phe His Lys Ile His Lys Gly Val Tyr Gln Glu Leu Pro Asn Ser
170    175    180
Asp Pro Ala Asp Pro Asn Ser Asp Pro Ala Asp Leu Ile Thr Gln
185    190    195
Lys Asp Tyr Leu Glu Tyr Gly Glu Ile Leu Pro Phe Pro Glu Ser
200    205    210
Glu Leu Val Glu Phe Lys Gln Phe Ser Thr Lys His Phe Gln Glu
215    220    225
Tyr Val Lys Arg Thr Ile Pro Glu Tyr Val Pro Ala Phe Ala Asn
230    235    240
Thr Gly Gly Gly Tyr Leu Phe Ile Gly Val Asp Asp Lys Ser Arg
245    250    255
Glu Val Leu Gly Cys Ala Lys Glu Asn Val Asp Pro Asp Ser Leu

```


WO 01/98353

PCT/US01/19862

260	265	270
Arg Arg Lys Ile Glu Gln Ala Ile Tyr	Lys Leu Pro Cys Val His	
275	280	285
Phe Cys Gln Pro Gln Arg Pro Ile Thr	Phe Thr Leu Lys Ile Val	
290	295	300
Asp Val Leu Lys Arg Gly Glu Leu Tyr	Gly Tyr Ala Cys Met Ile	
305	310	315
Arg Val Asn Pro Phe Cys Cys Ala Val	Phe Ser Glu Ala Pro Asn	
320	325	330
Ser Trp Ile Val Glu Asp Lys Tyr Val	Cys Ser Leu Thr Thr Glu	
335	340	345
Lys Trp Val Gly Met Met Thr Asp Thr	Asp Pro Asp Leu Leu Gln	
350	355	360
Leu Ser Glu Asp Phe Glu Cys Gln Leu	Ser Leu Ser Ser Gly Pro	
365	370	375
Pro Leu Ser Arg Pro Val Tyr Ser Lys	Lys Gly Leu Glu His Lys	
380	385	390
Ala Asp Leu Gln Gln His Leu Phe Pro	Val Pro Pro Gly His Leu	
395	400	405
Glu Cys Thr Pro Glu Ser Leu Trp Lys	Gln Leu Ser Leu Gln His	
410	415	420
Glu Gly Leu Lys Glu Leu Ile His Lys	Gln Met Arg Pro Phe Ser	
425	430	435
Gln Gly Ile Val Ile Leu Ser Arg Ser	Trp Ala Val Asp Leu Asn	
440	445	450
Leu Gln Glu Lys Pro Gly Val Ile Cys	Asp Ala Leu Leu Ile Ala	
455	460	465
Gln Asn Ser Thr Pro Ile Leu Tyr Thr	Ile Leu Arg Glu Gln Asp	
470	475	480
Ala Glu Gly Gln Asp Tyr Cys Thr Arg	Thr Ala Phe Thr Leu Lys	
485	490	495
Gln Lys Leu Val Asn Met Gly Gly Tyr	Thr Gly Lys Val Cys Val	
500	505	510
Arg Ala Lys Val Leu Cys Leu Ser Pro	Gln Ser Ser Ala Glu Ala	
515	520	525
Leu Glu Ala Ala Val Ser Pro Met Asp	Tyr Pro Ala Ser Tyr Ser	
530	535	540
Leu Ala Gly Thr Gln His Met Glu Ala	Leu Leu Gln Ser Leu Val	
545	550	555
Ile Val Leu Leu Gly Phe Arg Ser Leu	Leu Ser Asp Gln Leu Gly	
560	565	570
Cys Glu Val Leu Asn Leu Leu Thr Ala	Gln Gln Tyr Glu Ile Phe	
575	580	585
Ser Arg Ser Leu Arg Lys Asn Arg Glu	Leu Phe Val His Gly Leu	
590	595	600
Pro Gly Ser Gly Lys Thr Ile Met Ala	Met Lys Ile Met Glu Lys	
605	610	615
Ile Arg Asn Val Phe His Cys Glu Ala	His Arg Ile Leu Tyr Val	
620	625	630
Cys Glu Asn Gln Pro Leu Arg Asn Phe	Ile Ser Asp Arg Asn	
635	640	645
Cys Arg Ala Glu Thr Arg Lys Thr Phe	Leu Arg Glu Asn Phe Glu	
650	655	660
His Ile Gln His Ile Val Ile Asp Glu	Ala Gln Asn Phe Arg Thr	
665	670	675
Glu Asp Gly Asp Trp Tyr Gly Lys Ala	Lys Ser Ile Thr Arg Arg	
680	685	690
Ala Lys Gly Gly Pro Gly Ile Leu Trp	Ile Phe Leu Asp Tyr Phe	
695	700	705
Gln Thr Ser His Leu Asp Cys Ser Gly	Leu Pro Pro Leu Ser Asp	
710	715	720
Gln Tyr Pro Arg Glu Glu Leu Thr Arg	Ile Val Arg Asn Ala Asp	
725	730	735

WO 01/98353

PCT/US01/19862

```

Pro Ile Ala Lys Tyr Leu Gln Lys Glu Met Gln Val Ile Arg Ser
740 745 750
Asn Pro Ser Phe Asn Ile Pro Thr Gly Cys Leu Glu Val Phe Pro
755 760 765
Glu Ala Glu Trp Ser Gln Gly Val Gln Gly Thr Leu Arg Ile Lys
770 775 780
Lys Tyr Leu Thr Val Glu Gln Ile Met Thr Cys Val Ala Asp Thr
785 790 795
Cys Arg Arg Phe Phe Asp Arg Gly Tyr Ser Pro Lys Asp Val Ala
800 805 810
Val Leu Val Ser Thr Ala Lys Glu Val Glu His Tyr Lys Tyr Glu
815 820 825
Leu Leu Lys Ala Met Arg Lys Lys Arg Val Val Gln Leu Ser Asp
830 835 840
Ala Cys Asp Met Leu Gly Asp His Ile Val Leu Asp Ser Val Arg
845 850 855
Arg Phe Ser Gly Leu Glu Arg Ser Ile Val Phe Gly Ile His Pro
860 865 870
Arg Thr Ala Asp Pro Ala Ile Leu Pro Asn Val Leu Ile Cys Leu
875 880 885
Ala Ser Arg Ala Lys Gln His Leu Tyr Ile Phe Pro Trp Gly Gly
890 895 900
His

```

```

<210> 44
<211> 1040
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7690129CD1

```

```

<400> 44
Met Ala Ser Thr Gly Gly Thr Lys Val Val Ala Met Gly Val Ala
1 5 10 15
Pro Trp Gly Val Val Arg Asn Arg Asp Thr Leu Ile Asn Pro Lys
20 25 30
Gly Ser Phe Pro Ala Arg Tyr Arg Trp Arg Gly Asp Pro Glu Asp
35 40 45
Gly Val Gln Phe Pro Leu Asp Tyr Asn Tyr Ser Ala Phe Phe Leu
50 55 60
Val Asp Asp Gly Thr His Gly Cys Leu Gly Gly Glu Asn Arg Phe
65 70 75
Arg Leu Arg Leu Glu Ser Tyr Ile Ser Gln Gln Lys Thr Gly Val
80 85 90
Gly Gly Thr Gly Ile Asp Ile Pro Val Leu Leu Leu Ile Asp
95 100 105
Gly Asp Glu Lys Met Leu Thr Arg Ile Glu Asn Ala Thr Gln Ala
110 115 120
Gln Leu Pro Cys Leu Leu Val Ala Gly Ser Gly Gly Ala Ala Asp
125 130 135
Cys Leu Ala Glu Thr Leu Glu Asp Thr Leu Ala Pro Gly Ser Gly
140 145 150
Gly Ala Arg Gln Gly Glu Ala Arg Asp Arg Ile Arg Arg Phe Phe
155 160 165
Pro Lys Gly Asp Leu Glu Val Leu Gln Ala Gln Val Glu Arg Ile
170 175 180
Met Thr Arg Lys Glu Leu Leu Thr Val Tyr Ser Ser Glu Asp Gly
185 190 195
Ser Glu Glu Phe Glu Thr Ile Val Leu Lys Ala Leu Val Lys Ala
200 205 210

```

WO 01/98353

PCT/US01/19862

Cys Gly Ser Ser Glu Ala Ser Ala Tyr Leu Asp Glu Leu Arg Leu
 215 220 225
 Ala Val Ala Trp Asn Arg Val Asp Ile Ala Gln Ser Glu Leu Phe
 230 235 240
 Arg Gly Asp Ile Gln Trp Arg Ser Phe His Leu Glu Ala Ser Leu
 245 250 255
 Met Asp Ala Leu Leu Asn Asp Arg Pro Glu Phe Val Arg Leu Leu
 260 265 270
 Ile Ser His Gly Leu Ser Leu Gly His Phe Leu Thr Pro Met Arg
 275 280 285
 Leu Ala Gln Leu Tyr Ser Ala Ala Pro Ser Asn Ser Leu Ile Arg
 290 295 300
 Asn Leu Leu Asp Gln Ala Ser His Ser Ala Gly Thr Lys Ala Pro
 305 310 315
 Ala Leu Lys Gly Gly Ala Ala Glu Leu Arg Pro Pro Asp Val Gly
 320 325 330
 His Val Leu Arg Met Leu Leu Gly Lys Met Cys Ala Pro Arg Tyr
 335 340 345
 Pro Ser Gly Gly Ala Trp Asp Pro His Pro Gly Gln Gly Phe Gly
 350 355 360
 Glu Ser Met Tyr Leu Leu Ser Asp Lys Ala Thr Ser Pro Leu Ser
 365 370 375
 Leu Asp Ala Gly Leu Gly Gln Ala Pro Trp Ser Asp Leu Leu Leu
 380 385 390
 Trp Ala Leu Leu Leu Asn Arg Ala Gln Met Ala Met Tyr Phe Trp
 395 400 405
 Glu Met Gly Ser Asn Ala Val Ser Ser Ala Leu Gly Ala Cys Leu
 410 415 420
 Leu Leu Arg Val Met Ala Arg Leu Glu Pro Asp Ala Glu Glu Ala
 425 430 435
 Ala Arg Arg Lys Asp Leu Ala Phe Lys Phe Glu Gly Met Gly Val
 440 445 450
 Asp Leu Phe Gly Glu Cys Tyr Arg Ser Ser Glu Val Arg Ala Ala
 455 460 465
 Arg Leu Leu Leu Arg Arg Cys Pro Leu Trp Gly Asp Ala Thr Cys
 470 475 480
 Leu Gln Leu Ala Met Gln Ala Asp Ala Arg Ala Phe Phe Ala Gln
 485 490 495
 Asp Gly Val Gln Ser Leu Leu Thr Gln Lys Trp Trp Gly Asp Met
 500 505 510
 Ala Ser Thr Thr Pro Ile Trp Ala Leu Val Leu Ala Phe Phe Cys
 515 520 525
 Pro Pro Leu Ile Tyr Thr Arg Leu Ile Thr Phe Arg Lys Ser Glu
 530 535 540
 Glu Glu Pro Thr Arg Glu Glu Leu Glu Phe Asp Met Asp Ser Val
 545 550 555
 Ile Asn Gly Glu Gly Pro Val Gly Thr Ala Asp Pro Ala Glu Lys
 560 565 570
 Thr Pro Leu Gly Val Pro Arg Gln Ser Gly Arg Pro Gly Cys
 575 580 585
 Gly Gly Arg Cys Gly Gly Arg Arg Cys Leu Arg Arg Trp Phe His
 590 595 600
 Phe Crp Gly Ala Pro Val Thr Ile Phe Met Gly Asn Val Val Ser
 605 610 615
 Tyr Leu Leu Phe Leu Leu Leu Phe Ser Arg Val Leu Leu Val Asp
 620 625 630
 Phe Gln Pro Ala Pro Pro Gly Ser Leu Glu Leu Leu Leu Tyr Phe
 635 640 645
 Trp Ala Phe Thr Leu Leu Cys Glu Glu Leu Arg Gln Gly Leu Ser
 650 655 660
 Gly Gly Gly Gly Ser Leu Ala Ser Gly Gly Pro Gly Pro Gly His
 665 670 675
 Ala Ser Leu Ser Gln Arg Leu Arg Leu Tyr Leu Ala Asp Ser Trp

WO 01/98353

PCT/US01/19862

```

680      685      690
Asn Gln Cys Asp Leu Val Ala Leu Thr Cys Phe Leu Leu Gly Val
695      700      705
Gly Cys Arg Leu Thr Pro Gly Leu Tyr His Leu Gly Arg Thr Val
710      715      720
Leu Cys Ile Asp Phe Met Val Pro Thr Val Arg Leu Leu His Ile
725      730      735
Phe Thr Val Asn Lys Gln Leu Gly Pro Lys Ile Val Ile Val Ser
740      745      750
Lys Met Met Lys Asp Val Phe Phe Phe Leu Phe Phe Leu Gly Val
755      760      765
Trp Leu Val Ala Tyr Gly Val Ala Thr Glu Gly Leu Leu Arg Pro
770      775      780
Arg Asp Ser Asp Phe Pro Ser Ile Leu Arg Arg Val Phe Tyr Arg
785      790      795
Pro Tyr Leu Gln Ile Phe Gly Gln Ile Pro Gln Glu Asp Met Asp
800      805      810
Val Ala Leu Met Glu His Ser Asn Cys Ser Ser Glu Pro Gly Phe
815      820      825
Trp Ala His Pro Pro Gly Ala Gln Ala Gly Thr Cys Val Ser Gln
830      835      840
Tyr Ala Asn Trp Leu Val Val Leu Leu Leu Val Ile Phe Leu Leu
845      850      855
Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile Ala Met Phe Ser
860      865      870
Tyr Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu Tyr Trp Lys
875      880      885
Ala Gln Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg Pro Ala
890      895      900
Leu Ala Pro Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu Leu
905      910      915
Arg Gln Leu Cys Arg Arg Pro Arg Ser Pro Gln Pro Ser Ser Pro
920      925      930
Ala Leu Glu His Phe Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg
935      940      945
Lys Leu Leu Thr Trp Glu Ser Val His Lys Glu Asn Phe Leu Leu
950      955      960
Ala Arg Ala Arg Asp Lys Arg Glu Ser Asp Ser Glu Arg Leu Lys
965      970      975
Arg Thr Ser Gln Lys Val Asp Leu Ala Leu Lys Gln Leu Gly His
980      985      990
Ile Arg Glu Tyr Glu Gln Arg Leu Lys Val Leu Glu Arg Glu Val
995      1000      1005
Gln Gln Cys Ser Arg Val Leu Gly Trp Val Ala Glu Ala Leu Ser
1010      1015      1020
Arg Ser Ala Leu Leu Pro Pro Gly Gly Pro Pro Pro Pro Asp Leu
1025      1030      1035
Pro Gly Ser Lys Asp
1040

```

<210> 45

<211> 2508

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2101688CB1

<400> 45

```

cggtctctctc cgggggtgct tggcacagcc tgggattcct cctctctgct gctcagatca 60
gtttccctat cggcagcagc gggcagggcg gcagcggcgg cggcagcagc cagcagtcgc 120
gcgtggggaa catctcggca gccacagcgc ttctcccgct ggcgcggggc tccagcttgg 180

```

WO 01/98353

PCT/US01/19862

```

ctgccccggg tcttccctg ccaagtttgg ggtgcgcctg cccccccac ccaggtctgc 240
ttctcttcga agcgggaagg gcgccttgcg ggatccctgc gccctccaa ccggaacctg 300
ggtctagagc tccccaaagc gaggcctcgc ccaggactcc tgcctcccca accctgacgc 360
ccgggggggg ccccggggac gtacgcgcgc ggaagggaag cggcaaaagg gcccatggcg 420
cgcttgactc gtccgcttgg tctgcnagtc ttgggggggc tctggatcac gggtcgctgc 480
ttctctcggg taacaaagag gaagtlggag gtgcgcagcg gaactgaugt gcaagacctc 540
aacctctcgg acgtcactcg ggcgcgcctg tgggancagt ttgatgagcg cgggtatctg 600
aatgcacaaa agtggcgcgt ttgtgacgac cctataagc tgrtatgcttt caacacagcg 660
gagagtgggc ggatctccag caatggggcc atccgcgaca ctgcgcctct gatgtcac 720
ctcttggtgt attgcacgga ccttccaccc actagctaca tcalcacctt ccacacagcg 780
gccctcccaa cgtcgtccag gacctccgc agtgtattaa accgcacccc taagcatctg 840
atccgggaaz tcatattagt ggatgacttc agaatgacc ctgatgactg taacagctc 900
atcaagtgc ccaagggtga atgcttgcgc aataatgac ggcagggtct ggtccggtcc 960
cggattccgg gcgctgacat cggccaggcg accactctga ctttctctga cagccactgt 1020
gaggtgaaca gggactggct ccagcctctg ttgcacaggg tcaaaaggga ctacacggcg 1080
gtggtgtgac ctgtgctga tatcattaac ctgcacaccl tcaactacat caggtctgct 1140
tcggagctca gagggggggt ttgactggag ctcacactcc agtgggagca gctctcccca 1200
ggcagaagg ctggggcctt ggcacccacg gggccactca ggaactctat cactgctgga 1260
gggtctcttg tgatgcacaa agcttgggtt gattacctgg ggaatataga tatggacatg 1320
gaacatcggg gtggggagaa ctttgaatcc tcttcccgag tctggatgtg cggggggagc 1380
ctagagatgc tccctcgag cgggtggggg ccgtctctcc ggaagagca cccctaggtt 1440
ttcctcgatg gaattgccaa ccgltatata aagaacacca agcggacagc tgaagtgtgg 1500
atggatgaat caaagcaata ctatbacct gcccgcccat tgcctctgga ggggcctctc 1560
gggaatgttg agnccagatt ggcactgagg aagaatctgc gctccagag cttcnagtgg 1620
taactcgaga atatctacc tgaactcagc atcccaagg agtctccat ccagaaaggc 1680
aatctcgac agagacagaa gtgcctgga tctcaaggc agaacaccca agaaacccca 1740
aacctaaagt tgagccctg tgcacaggtc aaggcgag atgcaaatc ccaggtatgg 1800
gccttccat acacccagaa gatcctccag gaggagctgt cctgtcagt catcaccttg 1860
ttcctcgag ccccgatggt tcttgccttt tgcagaatg gagatgacg acagcaatgg 1920
accaaactg gttcccaat ccagcacata gcatcccaac tctgctcga tacagatg 1980
ttcgtgatg gcaccgagaa cggcaaggaa atctgctga acccatgtga gctcctactc 2040
atgagccagc actgggacat ggtgagctcc tgaggacccc tgcacagag agcaaggccc 2100
atgggtgtgt gcttccctgg accagaacag actggaaact gggcagcaag cagcctgcaa 2160
ccactcaga calcclygac tgggaagtg aggcagagcc ccccgagaca gggcagctg 2220
tctcagggag gacagggaaa aacatccaaa gccaatgggg ctcaagaca aatcccaat 2280
gttctcagc cgttcaagtt ccagtcctgy ccagtcattc cctgallgt atcggagac 2340
agaaacctaa tgggaagtgt ttattgttcc ttttctaca aggaagcag tctctggag 2400
ccagaaagaa agcctcttct ttctactagg ccaggactac attgagagat gaagaaatga 2460
ggtgttttcc aagaagaaa aagagaaact tagaaaaaa aaaaaaaa 2508

```

<210> 46
 <211> 4034
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5453330CB1

```

<400> 46
aacacgacac atgagcccggt tagcctgtgt gctcctgccc tgcgtgctgg cgtctctggc 60
tcacagactc tctttagagg ccccaaccgt ggggaaggga caagcccccg gcatcgaggga 120
gacagatggc gagctgacag cagcccccac acctgagcag ccagaaacgag gcgtccactt 180
tgccacaaca gcccccacct tgaagctgct caaccaccac ccgctgcttg aggaatttct 240
acaaagagggg ctggaaaagg gaaatgaggga gctgaggcca gcactgccct tccagctcga 300
cccaactgca ccttccccc caagtccctt tcccgctg gccacacagg acagccgcc 360
tgtctttacc agcccccact cagccatggc tgcggtaccc actcagcccc agtccaggga 420
gggacccctg agtccggagt caggttcccc tatgcttcca atcacagctc cctacacctc 480
aggcccccag atggcagtg ccacccatag cccaggggag atagccagca ctacaccccc 540
cagcagagcc tggacaccaa cccaagagg tcttgagagc atgggaaggc cgtgggttgc 600
agaggttgtg tccaggggcg cagggatcgg gatccagggg accatcacct cctccacagc 660
ttcaggagat gatgaggaga ccaacactac caaccacac atcacacca ccatcacacc 720
agtccagaca ccaggccctt gtatgtggaa tttctcaggc ccagagggtc cctcggactc 780

```

WO 01/98353

PCT/US01/19862

```

cctacagac ctcagctccc cccatgatgt tggcctggac tgccttctct acatctctgt 640
ctacccctggc tatggcgtgg aatcaaggt ccagaatata agcttcgggg aaggggagac 900
agtgcctgtg gaaggcctgg gggggcctga cccactgccc ctggccaacc agtcttctct 960
gctgcggggc caagtcctcc gaagccccc ccaccaagcg gccctgaggt tccagagcct 1020
ccgcccacgg gctggccctg gaactctca tttccctac caagcctata tccagagcct 1080
ccactttccc cgtcctccag ctbatggaga tgtgactgtc accagctccc acccagggg 1140
tagtgccegg ttcacttggt ccaatggcta ccagctgaag ggcgcaggc atctccctg 1200
tcloaatgct accagccct tctgggattc aaaggcgccc gctgcctcg ctgctggcg 1260
cggagtgatc cgcactgcca ccacggcgcc catcgtctct ccaggcttcc cgggcaacta 1320
cagcaacaac ctacactgta actggctgct tgaggtctct gaggcccgag ggttaacct 1380
gcactttgag aagggttccc tggcagagga tgatgacagg ctcatcatic gcaatgggga 1440
caacgtggag gcccccacag tgtagcttcc ctatgaggtg gaatacctgc ccatggagg 1500
ctgctcagc tctggcaaac acttctttgt tgagctcagt actgacagca gggggcgag 1560
tgagggcatg gccctggcgt atgagggcct ccagcagggc cattgctatg agccctllgt 1620
caalacggt aacttcagca gcagcacacc cacttacct gtgggtacca ctgtggagtt 1680
cagctgcgac ccggctaca cctggagca ggcctccatc atcatcgagt gcttggacc 1740
ccacagcccc cagtggcatg agacagagcc agcctgccc ggcgtgtgca gggggagat 1800
ccagactctg gctggggtgg tctctctccc caactggcca gagccctacg gctgtgggca 1860
ggatgtatc tgggggtgtc atgtggaaga ggaacaagcc atcatgcttg acatccagct 1920
gctggcataa ggcctctgtg atgtgttacc ctctctagat ggggatgacc tgaagcctg 1980
ggtctctggc caatactcag ggcctcttag caacttcaat ctcttctact caatggctg 2040
tgtccactat cagttccagt cggccccggg gacgtcagt ctgggctacc agcagggctt 2100
cgtcatccac ttcttctgag tgcctccgca tgcacatgt cggagctgct ctgagatccc 2160
caatggctgg aagagcccat cgcagcctga gctantgac ggcacctgg tccacttaca 2220
gtctacccct ggttaccagg taagtgggac cagtgtctct atgtgctagt gggacctaac 2280
ttggagtgag gaactgccc ctatgcagag ggtgacttcc tgcacagat ctggagatgt 2340
ggagcaacag cgaactccca tatccagccc caagtttccc glgggggcca cgttgcaata 2400
tatctgtgac caggggtttt tgctgaiggg cagctccatc ctccctgccc atgatccca 2460
ggctggcagc cccaagtga gtgcccgggc cctcaaatgt ctctgggacc agctcaagcc 2520
atgcctatgt ctcaagtccc ctgaggaatg tgcctgaagt cctgagagag agctcaaccc 2580
agcagggccc accatccact tctgtgtgct cctggctat gtgctgaag gccagggccc 2640
catcaagtgt gtcctggggc acccttgcga ttggagtgac cccccacca tctgtaggcc 2700
tgcttctctg gatgggttct acaacagctg cagcctggat gtcgccaagg cactctgccc 2760
ctcagagacc ctgggtgtgt cccacatgg agctgtcacc tcttggcacc tgggtggcct 2820
gtgtgtgttg taggaggtg tatacttcta ctctccagg ctccagggaa aaggttccct 2880
gcagctgccc cgcctccgct cccgcctcta caacggact accatagagt caggtttga 2940
caatccact taagagactg gatcttctc ctlttgagga gacgagaga tatgaagtct 3000
ccatctaggt gggggcagtc tagggagtc aactcagact tgcaccacag tccagagcca 3060
aggctccttg ctctcgtgtg tccctccacc tccgttatat accactagg aggagatgcc 3120
accaagccct caagaaggtg tgccttccc cgcctggat gccaccag gctatttttc 3180
ttgtgtgcat tgcacacttg gggccttcca ttgggcccac gtcagggggc atctacutgt 3240
gggaaagaca tagctggagc acaagcatca acagccagca tcttgagcct cctcatgccc 3300
tggaccagcc tggacacac tagcagagca ggaatccctb tctccacatg accacatccc 3360
cgccttgcca tggcaacctg cagcaggatt aacttgacca tgggtgggac tgcacccagg 3420
taactctcac agcgccatca ccaatggcca aaactctct ctacgggtgac ctctgggtag 3480
tcttggcag ccaacatcag cctcttggga ggtctctagt tctctaaagt tctggacagt 3540
tctgctctct gccctgtccc agtggaggca gtaattctag gagatctctaa ggggtttag 3600
gggaacctac gccacactca ggttgggtct ccttgggccc tcatgtctca caccaaagca 3660
ggaaacgcca ttttccactg acccactat accctgagga aagggagact tctctccat 3720
gtttatttag ctgttgcaaa catcttccc ctaatagtcc cctctccat tccagccact 3780
tgtcaggtct tctcttgac cactgtgta tgggataagg ggaaggggtg ggcatactct 3840
ggagaggagc agaggtccaa ggacccagga atttggcag gaacaggtg taggagagcc 3900
ccagggggcc gccccagagc tgggtgaaa ccacttttga catgtaagt attatatgg 3960
gtctgggtct cagccagaga caactcttt atttctgttg tttcttatt aaaatgggt 4020
ttttgggaaa aaaa 4034

```

```

<210> 47
<211> 645
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

```

WO 01/98353

PCT/US01/19862

<223> Incyte ID No: 4362432CB1

```

<400> 47
ggtagaagag ccagccatg ttacaggaca aaggccglog ctttgaasa gcttgaagt 60
cagtttgctg ctgagtagag aagacctttg caaacagaga ggggagattt tctctgtag 120
gttgaagaca agagcaggtc ctggaagata agattcccg catgtttatc ctccgtggtg 180
ttttgggac taattgacct ctttggaact tccggggct catccctctt cagtcattca 240
atgagagctc acataaatcc aggcctgtac cccggttctt atggggacat cagtcacatg 300
aagacctctg gggccacttg tgatgcaaac agtgtgatga actccgggat ccgtggtttt 360
gaattgttgc ctgagatgga tttaggggac ataaaacctt accagactct gatcaagaa 420
gtcgggcaga gacattgggt ggacccctgt gtcctcgag ccactatctc cagggaagac 480
catggcggat ctgtctgca agacggctgg gaccacagg gacttaattt tggcttgatg 540
cagcttgata aacaaacgta ccaacctgtc ggtgcttgg atagcaaga gacactttca 600
caggclactg ggattcctaac agagagaatt aaggcaatcc agaaaaaatt ccccaagtgg 660
aggttgctgc agcacctaa agtggtgttc ttagttttta agtcagguat tgaagcatt 720
gccaccccat cggacataga caatgacttc gtcaatgata bcattgctcg agctagctc 780
tataaagac aaagcttcta ggcaaaagctc tgtgggtggg ccagggttggc agagtgctca 840
gagctg

```

```

<210> 48
<211> 2300
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5308104CB1

```

```

<400> 48
tttagggcgc agtgtgctgg acaaagcgggt caggccgggc ggcatgtgtg gggggggccc 60
cggtaggggc ccggccctgg cggcccccgc catgaatggg ctgtgctga gtagactctg 120
ctgctctctc tggtagccgc cctgcccggg ccgcatcgct gccagagctg ccllcttccc 180
ggcggaggcc acctactccc tggtagctga gccgagccg gggcctgtgt gggccggggc 240
cggcccttgg gggacccctga gagctcttc gggccacccc gggcgctgga agctgcaact 300
gagggagcgt gccgacttcc agtacaacca gccgagctg gccacactcg aggtcttccc 360
cacaaagagc gcccggggca accgcgtctc ctgcatgtat gttcgtgtcg tgcctgtgtc 420
caggtacacg gtctctttct cgcacggcaa tgcctgtgac ctgggcccga tggagcagctt 480
ctacattggc ctgggctccc gcttccactg caactcttcc tcttaccact actccggcta 540
cgtgtccagc tcgggcaggg cttccgagag gaactctat gccgacatcg acgcccctg 600
gccggccctg ccgaccaggt accgcatcag cccggacaga atcactctgt accggcagag 660
catcgccacg gtgcccacgg tggacccggc ctgagctac gagtgtgccc cggtaattct 720
ccattccctc ctgaggtctg gtttgcgtgt ggcttttccc gataccagga aaactactgt 780
ctttgatgct ttcccagcca ttgacaagat attcaaatgc accctctctg tgttggctat 840
tcattgtaca gaggatnagg tcatcatttt ctcccattgg ctgagcagtg accgagcctg 900
tcccagacc gtggggcccc ttgggttga aggggctggg cataatgaca tagagcttta 960
tgcaaatcac ctgaaagac taanaacagt catatctcac gaacttctca attcctgaag 1020
acaaacactt gatellacat catttacgtt gaacagaaga gtctctgttt tggcaactgc 1080
tttaacttgg tagctgtaaa ggttgaata ccatgaagaa gtgcccaccc tttaggggtg 1140
tcaatcaaa gagctgtag aaatccagtc tctgtatctt agaggtagtt ctgctaattc 1200
aocaaacag tttaactgaa cagtcgtgat tccagcttcc attacttgc aggaatggga 1260
atgagagctg aatgtaggga caattttcta gtgctgtata aagtagcttc gcatctgttt 1320
ctcaactctt tccatcattt ctgacattca tgcaggactt gccctgttgc caaccaatgt 1380
ctoggtattt cactgcaagc tctcttctgt cacttgata catgggttca atccatttgt 1440
gaagctgtga tagtgtactt ggaagctag tgtgttgaaa attcctttat tatttttgt 1500
taacatgctg attcttccc gacaaatgaa ctgaaggata atttactgga actctctgt 1560
acagctctat caactgtaac catataaata taactgggat attcttaaac aaaaagaaac 1620
tegggttttt tttaagtga aatttactac tagcccaag agttttacta ttttgattgt 1680
ctgggttggt taacaaagag cctagctgac ttctctctg taaagtcttc ctgttaggt 1740
tttttaaat acgtanata ttgcataca cattgtgcat agattcttca tggtagatat 1800
gattttcttt tccaggctac aacaaagac tgcagattcc ttgtttgtaa tggtaattgt 1860
tgaatcattt ttgttaatat gtttttattc ctatgttttg ctattaaaaa ttttaataca 1920
tttccagac aaaaattcca agtttatgct ttgagaat tttgttaata aatttctact 1980
aaactaatct ttttagtcta ggaattattt ggggtttgac actggaagtt ggcacaaata 2040

```

WO 01/98353

PCT/US01/19862

```

agratcagaa ataggagatg cttacacattg ctatactact tgtgttggtt aggggtttgg 2100
atttgggggt tctttgggtt taattttttt ttccacattt aaaagcctta aatgtactgt 2160
aagcctcaga tcttgggtaca actggactgc ggttgaltgc cagtttgggt actgtttgctt 2220
gggtggcgga cagtgggttg taatgggaata aaggatgcac ggtacagaaa aaaaaaaa 2280
aaaaaaa 2300

```

```

<210> 49
<211> 1587
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3092736C81

```

```

<400> 49
ggaggacata attacacctgt cctagctctct tctcatctta catttccctg tagccactgg 60
gacatctgtg gtttcccttc ctagctctctg tctctctctc atgctcttgc tgggtatggg 120
cactgttggg ggaaggtctat tctgttcaga ggggcactga ctcttctaatg gtttaccaca 180
agtggtatgt tggagacaca gtctgcatgc tgcctcagtc cgggcagacc ctactatccc 240
aggagctgc tgcgtctggc agtctctctc tgcctcttat ttccacagtg gtaagagacc 300
acttgaccaa gcttaccgcc atggccacag gccgagtgcc tcaactctat ggttggaagg 360
gtcggagcaa gccgagtgc tcaactctgt ccttggaata agccttttcc tctatllcag 420
acctcagcga gggcgaacaa gaggctcgtt ttgcagcagg agtggctgag cagtttgcca 480
tcgcggagac caagctccga gcatggctct cgttgcatgg cgggactccc actgatgact 540
cctatgatga ggaactttgt gggggaatgg acacagacat gactgggcag ctgcccctgg 600
ggcgcacact ccaggacctg ttcacccgcc accggttctc cgggcctgtg cgcacagggt 660
cgttgagacc tgagagcgac tgcctacaga ccgtgtcccc agacacctg tgcctcagtc 720
tgtgcagcct ggaagatggg ttgttgggct ccccgccccc gctggcctcc cagctcctgg 780
gcgtatgact gcttctctcc aaactgcccc ccagccggga aagtgccttc cgcagcctgg 840
gccacatgga ggcacagacc tcaactctaca actgcacctc cagagagttc tgccttctcc 900
cccgaggaga ggaacagacc ccttgaagg actgcagcc actctgacca ccaactaacg 960
gcagctggga accgagcggg caagcctctg acctggcctc tctctgggtg gtgtccttag 1020
atgaggatga ggcagagcca gaggacactg gactcacatc atgcttggca gtggcatgca 1080
tccccgggt actgcagggg gcagagctcc tctgcacag tggggctca aggtctccag 1140
cagagctaca cagctagagg gactctgggg agcctcctct tctcgttgt gtgttttgca 1200
tgaaactgtt tggagaggag gccagggcct ggttgggggc gcatgtctg cccctactcc 1260
cgggctcttg cgggggttg cgggggcttc tggggcatgg ctacagctgt ggcagacagt 1320
gatgttcatg ttcttaaaat gccacacaca ctttctctcc tcggataatg tgaacacata 1380
agtggtttgt gactgggctg tgtgatggtg ggttgaggag gggccagaca acccccacac 1440
ctccccatgc ctctctcttc tctgcttttc ttctctcttc cagatcctg tgcagtgctt 1500
gatagatcca accccactg gaggggctgg ctctctgccc cccggagcct atgggttgag 1560
cgtctctcca agggccctgc cagctgg 1587

```

```

<210> 50
<211> 669
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3580257C81

```

```

<400> 50
cttgccatgaa caacaaacca actcaccact cctgacacca tgaatcacta cggcagctac 60
taccgagggc tgggtacacg ctgtggaggg ttggttgccc tgggtatgg ctatggctgt 120
ggaatgtgga gcttctgacg accgggttct ggtgttggtc atggaggcta cggatattgc 180
tctggctttg gaagctacgg atatggctct ggttttgagg gctacggata tggctctggc 240
tttgagggtc atggatatgg ctgctgccc ccatcttaca atygaggata cggattctct 300
ggcttttatt aatgaattg ctgaatttgg aagcagagga gaaacctcca aatgtgtttg 360
gtcctgtccc gtgcttctat tcaaaaatc cactctatgt ccttcagcat caatggagag 420
atatttagct atgttaaatc tttaaaatg atttaagctg cttctgtgaa tattgttgtt 480
cttttctctc tgggaatccg tcaactgaaa tcatallcat gtataactac cctaggctgt 540

```


WO 01/98353

PCT/US01/19862

```

tottttaaat ttgttttatt ttctttacca ccttaacttt gatgttatct atcaagatcc 600
tegcwaaac ttgtatttgy cttttatota ataaatgaag taaagtataa atatttttaa 660
aaaaaanaa

```

```

<210> 51
<211> 1463
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3634758CB1

```

```

<400> 51
atgcacatct agctggttgg ctttggttag agctcccgct agactttcgt tgggcccctag 60
gatttggtag ccccgaggtg tgggctctct ccagttaccag actcatttca glaccagcct 120
ttgggaagtc ggtggaatag ctgggtctctc taggtaccag gatagaattga cggccctagcg 180
ggcgccgggc gccggcgagc tgcgtgagcg gcgagctgga acccccagac cagctcaaac 240
ggagaccaaa actcgaagct tggagaaatt agcagysaat ggcggaatgag gcgttggttt 300
tgcctcccca taacggagtg gtctctggag tgtacaagtc cgcggagcag gggaggttgg 360
aaanaggagc atgtattact aaactggaaa acatgggggt tccggtggga caaggattga 420
tagaaaaggtt tacaaaagat actgcaaggt tcaaggatga gttagatct atgagttca 480
tttgtaaaga tttttggact acggtattca agaaaacaa cgcacactca aggaacaaac 540
atcagggcac ctatgtactt caggacaaca aatttcgcct gcttactcag atgltcgcag 600
gaaaacagta tttagaacat gcactcaagt atttagcatt tacgtgtggc ttaattcagag 660
gtggcttato aaacttggg ataaaaagra ttgtaacagc tgaagtgct tcaatgcctg 720
ctgcgaatt tcaggtgat atacagaagc tgtagaacat actgaaatgc aaggcctcaa 780
cagtgtaag agataaatta ttcatgtaaa agtatttcaa gtatgtatga tttatttaca 840
ttgttcgat ttgtacagg agtaagcat tatttttacc aatttaaac agalcaaaag 900
agatgaagg acattctgac atgacataca cttaacnaaa actattcaaa atgaaaaccg 960
gatttcaaat aaccagacan caagatgag ggccttctt ttaaaccttt tbatttggt 1020
agaglgalat gtaattagcc atagatggag aaacaaagct caggttttgt tgaattagca 1080
tgagagaaaa ttatgtacca acaagattat ttgtggaag atgagacaaa ttttgatana 1140
gtatgaattt cttttattt aaaaagcaaa catactaaat ttttttatt ttaattctta 1200
taatttatta agaatgttta caactgtata aggtttcat atatacattg tctgtctgta 1260
tatatttaata catatattgac tgcctaattt gtttatanaa ttaattttt ttaattaggt 1320
ttcattctct cagagctcca ttaattgaat caaaatgaaa tccagattag tttaaatgt 1380
aatctagtg ctttagggcc aaagaatatt aggtatgttt gaaaagaatt tttgtattta 1440
ttcctgttac agttttgact ttc
1463

```

```

<210> 52
<211> 1686
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4027923CB1

```

```

<400> 52
gtaaccacac gccacttggg atcaggttag taggtgcacg gccaggtgtg gatgcagaca 60
tccccgccta tctgttcccg ggtctgtgtg cgcacaggtc caggcctctt ctgctctga 120
ggagcgcag acgagtcaca ggcgcgcgcg cgcgcctcgc gcgctcggc gcttgcgcag 180
tgaagacccg gccgagcccg gcgcctggtt gcagcccgag agccccagc gctggtgggg 240
cagaaccccg gccccctgng tcttctggtt ccagagctgg tctctacagt ttacgtctcg 300
gcgaaagtga gtaggggggt ccttcgacac gtgctgagct ggttggaacg cgaagccgcg 360
gccctcgcgc acagccactt ccgagggcct ggggtcgagt tccccggcgt cggccaggct 420
ccgggcggcg tagccttctt ctccggagcg ggcgccttct cctacgcaga cttttgtcgg 480
ggctctctgc tgcaccaact gccctgcgtg ttttcagagc ccttcacgca ggtctggggc 540
agccggcgcg cctgggtgac gccgcggggg aggcgcgact tcyaccacct gctacggacc 600
tacggagacg tggttgtaac agttgcacac tctgggtccc aggaatacaa ctgaaacccc 660
aaagagcaca tgactctcag agactacac accctactga agagatcat acagggcgcg 720
tactctcttc ccaggggctg tcttactctc aaagactggc acttctgcag ggaatttcg 780

```

WO 01/58353

PCT/US01/19862

```

gtgaggagcgt ttttccacct gctgtgtac ttctgttcg actggtgaa ttagttcttg 840
gatgcactgg atgtggatga ctaccgcttt gtctacggg ggctgcggg cagctggtcc 900
cctgtccatg ctgacatctt cgcctccttc agctggctcg tcaatgtctg tgggaagag 960
aagtggctcc tcttccctcc agggcaggaa gaguucctgc gggaccgcca cggcaacctg 1020
cctagagagc tgaactccgc agcaactctg gacacacac tgcacccacg gaacccagctt 1080
gtctgcccac ccttggagat cagcagagaa gcgggcagaa tgggttttgt gcccagttgc 1140
tggcccccac aggtgcacaa cctggatgac acaatctcca tcaacccaaa ctgggtcaat 1200
ggcttcaacc tggccaacat gtggcgcttc ttgcagcagg agctatgcgc cgtgagagag 1260
gaggtccagc agtggagyya ctccatgccg gactggcacc accactgcca ggtatcatg 1320
aggtcctgct cgggcaccaa ctttgaagag ttttccact tcttcaaggt catcgctag 1380
aagggtctcc tggctctgag ggagggagcc gctgagagc gtgctgggtl ggggttcgaa 1440
caggcagcct ttgatgttg gcgcacaca gagggtgtgc cctccttggt tggcaccccc 1500
gacttccaga gagtggacac cagcggcttc tcaaccagc ccaagagct gctgcagcag 1560
ctgagagagg ctgttgatgc tgcgtggccc ccatagcacc tctgtgagg atagagggac 1620
gggtggacga gaggcagcct cctgtctcgg ggcccttcca gaataaaga cggccctccc 1680
tgtgaa 1686

```

<210> 53
 <211> 2497
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc feature
 <223> Ineye ID No: 4348533CB1

```

<400> 53
gggaggcggg agctgctttg cctccaccga tctccctgtg cggccctcat gtgctgtgct 60
cgggtcaccc ggaagtccgc ggccttccgc acacggtagg gtctgcagac ccgctgcccc 120
tggcgtctga agtngtctgt cggccccgcy gggcccgccc atggagaagg ccaggagagg 180
cggggatagc gtccccgggg ggcctcctact gacatctgtg gttgttggtat ttcacacaa 240
gaggggttgc caggttgaat tctcttaccg gccctgattt cggggagatg gacatgcagc 300
ccacacttta cctgagaagt ggaagtattt gccctctctt gctttaccag atgggcgaca 360
caaccacccg gaagatactg tgrtttttca cttgcacccc agaaatgaaa atggagccac 420
agtattttgt atctcttctt atcgacaaat tgaagccaaq gactggaag taaggcaagc 480
agatatccad agagagactg ttccagaaaq tgcctgtgtt ctaagcaagc tgcctctgta 540
tgggtttact caagcaaac ttcaactcct tacaactgca tattttgaag agaaggtatt 600
ttcccaaatl tctattctaa aggagcttta tgaacatag tatagttcct lgggaggtgc 660
ttcattagaa ggaaccacag tatatcttgg tctgtcactl cagatctttg tcttctactt 720
tgcacacaaq gtcttaatcc tatttaagct aattctctct gaaaaaagg ttttttttta 780
tattttccca gtgaataaat tgggtgggtg actgatgact gtggttatcc tttttccagg 840
catgattgan catggtctca gtgactgttc tcatgataga cccggaaaa glatgtctga 900
agatggtggy ctccaggaaa gtaaccatg tgcagatgat tttgtttctg catccactgc 960
lgatgttcca cacaccaact tgggaactat cagggaagtc atggcagaaa accatggaga 1020
agatgttgc atgaagactg aggaagcttt gtcccaagtg gaagacagca gcaaaaggca 1080
ggaaaccnat gataccaatc aatatttgaa acctccact cgcctcatct cagatctctc 1140
agaagtgcac tgggaactt tggatcttag tgtcttcagc gaccccaact tgaagaag 1200
ggaaagctgt ggaatcacac agacaaattt gtttccaaag gactctgtcc cctcagagag 1260
tcttccaatt actgtacac ctaagataa tggggacag gtatctctga taccagggtc 1320
catttcgggt ttggaagagg atcagtatgg calgcocccg gccctcttca caaagggtta 1380
tctgtgtttg ccttadctgg caltgcagca gcatcatctt cctctcagat tcaocgttgc 1440
gggtttgttt gctggagcta ctaacatctt ttttcganaa cagaacaacc tcaagtatgc 1500
cattgtggaa gtgaagaag ctcgatctca gatccatgat ccagaactca ggaagctgct 1560
taaccacacc actgcaganc taaggttgc agactacct gtgaggcac tgaactgaaa 1620
tcgggttgac gtcttcttag atggcagcgg ctgggagggg ggtgagaaat ggaatccggc 1680
ccagtttgcg gtctacattc atgcctgctt ggcctgcacg ctgcaattag acaatgaaaa 1740
gatattatgc gactatggga caacttttgc tacaagatgg aagaatactc acaactacag 1800
ggtgtggaac agcaacaagc atccagcact tgcagaaata aatccaaacc atccatttca 1860
agggcaatgc tcaatctcag acatgaagtc aaggtctctca catctgttcc agaatgtga 1920
agtggtgcaa aaatttgaa acgtatgggt cacaactgag cggaaatgttg tacaacacgg 1980
aaagctgtt ggcagtcag ttgagagagc tttttccagt gaaagagcag ctatgtcttc 2040
atagcttccc actttccaca cttccacctc ccaagttctc actgagccac cagatgagaa 2100
gacttgagca agggctcaga ggtgtctatt gctttctgag gtttaagtgt cccctgtctg 2160

```

WO 01/98353

PCT/US01/19862

```

lc:gtgtctc ctaggtctgtt actagccaca gatccacagc aggggaccat atgtcgaaat 2220
gtttacatgg atgttgcctc aagtgaatgt ttcgggatgc cgaatgatg aaacccacagc 2280
catagcaggg atggctttcc aggttggggg ttcaatctgac taacttttatt taagtctcag 2340
cctgatlaaa acatacagtg aaccttctaa tgaattttga gttttggcat ttgaagtctc 2400
gtacatcggt ttaactttag agagttttta ctatgtaaa ttttggcttg ttttggcttc 2460
tgaattctta gtaggtcact gtaatttata ctgtcta 2497

```

```

<210> 54
<211> 1783
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Tncyte ID No: 452185/CB1

```

```

<400> 54
ccgtctgctc ctccggaaag ccccgagaaa ctaccagaac ggggtccctg gggccagtgt 60
tgggagcgta cgttgaggtg gcagggctgc tgcgcccgtg cggaaatccg ctgtggagga 120
ttagccctct tggtagctcy tgggtttctg ccgttttggg accctctact gtgtagttctg 180
aacttccagc agtggagga gggcagagag agagctggaga tctgtggctc atgcagtgcc 240
gagagccagt acgpcaggaa gccatgatgg gactttcagt tcttttcttg gtggagcgca 300
gcacagggga gagggggtgc ccacagatgc tagaagctta aaggtcgtg ggggttagcg 360
gcctcaggac ataaggtttc cagacaggac ggggcaggtg gtgatgcagc ttaggtctgt 420
gagttctgac gggaaacagc ccccgccgac tccgtgtcgc tctgaacgga aaagggctct 480
ccctacacag gaacgggggt ctctggggga tccagtgcag tggccaaaca cagagcccg 540
accctctctc ctgtagagag tctgtgtctc cctctgggat gtaccctcag gtggagaccc 600
gcacagctc ccgcctccat ctgaaagagg cctcaggcat ccgtctctgg tccctgtctg 660
ttggactctt gtogattggc ctggctgctg cctactacag cggagatagc ctgggtctga 720
agctctctta cgtcacaggc tgcctgtctg tggctgtgca gaacttggag gactgggagg 780
aagcctctct ccacaagagc acagggaaag ttgttttgaa gacgttcagc ctctacaaga 840
agctgtgac tcttttcaga gctggccacg accaggtggt ggtctctctc catgtctcc 900
ccgacacagc gtctctctcc tggtagaat cccagcggtl caggtgtctc cccaggggca 960
gtgggggggt aacatccagg gctactctg ctgtgcacgc tgcagctaca ctgtggagag 1020
tgcctctccc cagggagccg cctcctctgc ttagtcagag atctcggggc atagacacat 1080
ctgcccagc gttcttcaca gaaggtctc tntgtctctc acctctagct ttagcacaag 1140
ctgtgagccc agaaaaaagg tccgtcttc tagtatctg cagtgtctaa aaacggggag 1200
ctccagccgg ggcaggtgtt tcatgccagt aatccagcca ctctggagg ccgaggtggg 1260
aggttgtctt gaggcaggga gttaagacc agcctgggca acacagcaag atcctgtctt 1320
tgtaaaaaaa ctacccaaac agaaaaaact gggagattt ctgcagaaat ttagttccag 1380
cctctctcga acctgggaag acctggcagg agggggcttg gctctcggct acngactctt 1440
cccaccccg taggagctga argccaaaca tcttgacctg ccagtgtctt tgtctctctg 1500
agtatccca ggtctctcca ggtgcggtgt gcaccgagcg cgcctggcct gatgctctgg 1560
cctgtgagct ggggactcct gggccctgtg agcccttagg cggcaggccc aggaatggct 1620
gggttaggaa gggaaacact ttgcccacg tctgtctgtg acctgggtga aagccgacag 1680
gagggagatg ggaacctcct cttcagtagt ggtgcagat cctctgtctg aggcaggggt 1740
caccataacc ataatatcc cttcacgtgt caaaaaaaa aaa 1783

```

```

<210> 55
<211> 1461
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Tncyte ID No: 4722253/CB1

```

```

<400> 55
ccagactcgg tcccgccggc tttaaggggc ccgggcgggc aaggcgacg gagccaagt 60
ccgggcagca gccclgggal gccctgagcc gggccgggag gaggctctcg ttgcttagca 120
acggggcccg gggcctgttg gcggagcctg caccgtggct gcgacacggg ggggggctc 180
agcgggagcc ggcggaggag cgggcacccg ccattggcac aggcacgggc cattggcgca 240
tgcgtagggc gcggccttgg ggcggggccc acccagtggc ggaactctct gaggggcggy 300

```

WO 01/98353

PCT/US01/19862

```

tggggggggg gggcgtgggg cgggagcggg gccgtcagta gtcggctgag gacggggggc 360
gggcgggggc tgggatgggg cggggcctgc gctgggtggg cggggggggg agggagacag 420
ggcaagcgcc cgggtggggg ccgttgggtg gtggcgggct gaagggtgtg ggcggagcag 480
cgtcgttggg tggcgggggg cgggacggga cgggcatggc cctgctgtgt tggcctgggt 540
gcttcacggc gggcgtgggc caggcctgtc tgcaetgcca cagcaacttc tcccaagagt 600
tctcttcta cggcaacat gtaacttca agtccctgtt aatgggggac atccccgtgt 660
caggggggct gctacacgac tgggggggac aacggtgaa gggagtgcac ctggccatcc 720
cggccaagat ccccggggag aagctggacc aagtggtgac atcagtgtao cagatgatgg 780
atccgtgtga cagggggag atgtacttcc ccgggtatlt ccccaacgag ctgggaacaa 840
ctctcgggga cagggtgac ctcatccaga ccgcatcat cgaacggcgc atcgactgtc 900
agacccgtg tggatcttcc cagtacgaga ccatctontg caacaactgc acagactgcg 960
acgtgcctgt ctttggctat aactgcagat cctcgggcca gtggagtao gctgtccagg 1020
gcctcttga ctaataaat aactggcaca aacaggacac gacatgaga ccaagctctc 1080
ctgccttctc ctggcctggg acacacagag ccccccggc cttctgtgta tggccagcct 1140
taaggtgtct gtagcccca cacttggcca acctgacctt ggaagatgtc gctgagtgta 1200
tcagcagcga ctgacagcag ctgggcttgc cccagggcaa cgtgggggca gacatcaga 1260
tggcagcgcc ctgctgttca ctctgagctt gggctgtctg tggctcagga cccctctccc 1320
gaccccgga agagctgagc tggcgaaggc caggagggcg gaggggagg atgggggtg 1380
ggctgtcgc agcaccagcg cctgggcagg tccgcagagc tggggatgt gattaaagt 1440
cctgtgttt ctaaaaaaa a
cctgtgttt ctaaaaaaa a 1461

```

<210> 56
 <211> 2116
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> misc_feature
 <223> Jaccy ID No: 4878134CB1

 <400> 56
 cggccgggga gtcgctgggg agggggaaga ctgggtactc ggatccgag cctgagtcgc 60
 cggccgggga ggggggctgg cgggggagct ctccgacacg ggcaccttct 120
 ggcctgacgg gatcgtgtc ctgaaggccc tagccttctg gtacttctgt gcatlctctg 180
 tggcttctga tgggacaaag cagcctatcg gtgacaggag gctgcttccc tgcagatgtt 240
 tctgagagaa cttctgagag taacttcagg acaggagag ctgggaagtc ttcagctaca 300
 tgcacacat ctcttggctg atggactgtt cagacatgaa ctcaacctg gacttgcctg 360
 ctcttctctg actgggacat tctgttcttg taactgttcc gggctgctcc aacatgcttc 420
 tcatgtctgc cctgtggggc ctctacatgt cctgtgttaa tggggccat gtctgtactt 480
 ctltggagtg gtagtccagg cttctggaga cgggttctct ggggatcttc ctgtgctctc 540
 ttttgaagct gtaaggctg cccacagata cccacacat cgggatgttc ctgtggggct 600
 tccgtgtgct gatcttcagg atcatgttg gaggaggcct gatcaagatc cgggggggac 660
 ggtgtggcg agacttcacc tgcattggact tccactatga gacccagag atggccatc 720
 ctgtggcgta ctacccgac cactcaacct ggtgtgttca tggcttggag acgtccagca 780
 acaacttcat cyagctctg gtcgcttct tcttcttct cgtcttctc agcctgagca 840
 tccgggggt gctgagatc ctgttccagg ccgtctcat cgtcagcggg aacctgact 900
 tctgagctg ctgactatg gtgcccagcc tggcctgctt tctgagagc aacctgggt 960
 tctgttctcc ctctgggcca ggcagctga aggacagat tctgagatg cagagggaca 1020
 cctgaggggg cggcccgag cccagattcg gctcgtgtt gggcggtgca gccacgtct 1080
 tggcggggt cctgtcggcc tggctcagcg tggcgtgtt cctcaactg ctgagcttca 1140
 ggcagatcat gaacacccac ttcaactct ttcacatct caacacttac ggggctctg 1200
 cagagatcac caaggagcgg gggaggtga tctgcaggg caccagcag tccacagcca 1260
 ggcgcccgga tgcctgttg gaggactacg agtccagtg caagccaggt gacccagca 1320
 gacggccctg cctcatctcc ccgtaccat accgcttga ctggtgatg tgggttcgg 1380
 cctccagac ctacagacac aacgactgga tcatccact ggtggcaga cctctggcca 1440
 cggacggcga ggccttgtcc ctgctggcac caaacctct cgggggggag ccccgagca 1500
 ggtgggtcgg aggagagcac tacagtaaca agtccagcg tctggagct ccccgagca 1560
 cggagggcaa gtgtgtgttg cggagagga tctgagctc ccccgagcgc cccagcctg 1620
 agggagctga gccctacttc agggagcgt agggagcgt cagcaagctt tgcctctccc 1680
 gacccagaaa taaggcgga gacccagccc ctcggcggt cagcaagctt tgcctctccc 1740
 tggcggcgg ccaagctggg catcgccag agggagctg agggagagc ggtgggagc 1800
 agtcccgag cgggggatcc aggttgggt ctgtgttcc atactgtgg ggtccagag 1860
 ccttgcacac ctggggcccc acatccagcc caaccttgt cccagggcg cgggggctct 1920

WO 01/58353

PCT/US01/19862

```

gatctccac ccatccacc ctctccac aggcaccg ccggcgctg cggccacc 1980
gagttgagac aalgccacc ctgaccact cctccactac agccctgacc atagaccag 2040
ccaggtagct cttggggtct ctggggtct aggggtctg tctgtttccc tcttcaatgg 2100
tgtgttcca gccagg

```

```

<210> 57
<211> 702
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Ineye ID No: 5050133CB1

```

```

<400> 57
ccgagggccg ggcccaacca cgggtcccca ggcagcctcc gccagccgga ccccgctccc 60
ctcttgatgc tgcctgctga cgcctgacag cgggagccca tgcctgagcg ggcgcscgag 120
ctggcgctct tctgacccc cgaagctggg gccagggcga agggagtgga ggggacatgc 160
gaagggatgc tctccagctt ggaaggtctt tgcagctggt ctgacctgat caggagtgat 240
acttcacaga tctggggagg aacatcccca gtcttaagg ccnaactgac aaaaatgcgt 300
ggcactatgc ccaagtgga cggctagag gccttcgtca agatgggttg acacacagtc 360
gcctcctctg aagcagagct gcttcaggct gagggggacc atggggctct ccttcaggcc 420
ctggggaggt ggtggggtgc cgcagggtct cctccttcca ggaacaagtc aactgcacgc 480
gtgcccgtga cgtacagact gccacactg tataggacgg aggaactatli tctgtgggac 540
ggcggggagc caccagacca ccccgccauu tgcctctggc ctttgtgagc tttgtggtct 600
tcccatcagg aacgtgggaa agtgacattg tgtacagct gcagctlggg ggtttcttct 660
tgttatctct gtttatttta tattttaaaa atatttaaaa aa 702

```

```

<210> 58
<211> 2613
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Ineye ID No: 563024CB1

```

```

<400> 58
tctcggaaac cattttactg ggctcgagtt tagcgcgcc cggaggggtg gggaccttga 60
gtcatgctcc tacaaccacc gctagagttc ctctctcttg agtgcagagg tttagactgt 120
gtcttttgtt gccgaaagtc ctgcagttct cacagcgacc tggcagaaaa agtctgtccc 180
aaatgtttgt aaatcctccg ttgggcaacc cgccttcagc ttctgctgtg atcttgtcga 240
gpcactaagc gtgcagtatt agcagagagc ggggtggcag agtgcctggc ctgaaggtca 300
tgttgcaggg gtaactatgc tgttttaggg gccgggcaga gggagggagc actgaccacc 360
ccagagggcg cccctttagc tgcctgtctt tggcgggcgt cgtctccacg gccgcccagg 420
taccocagcc gctgtctgac tgtcttcttc ctttaacgag caggcttggt caccgcattt 480
aggcatccgc gctctccgag gaactgctct gcctctcggc cgttgaaagt gtagaggggt 540
ccatagtttg gtccactttg actcctctct cctgcocgga tcttlaagg cctctctgic 600
ctcccggtct ccggtcgtgc ccgggtctgt gcgcgggtcc ggcctccgccc tgcctctgac 660
atggggcgtt ccagctcttc cgcgttgagg cgcctcggcc tcccagcccg gccctggccc 720
agtggtctcg ggtcgccgca gctagagctg gccgcgttgg cctgggggac tctcgccttg 780
cgcggcgcat ggcccaggcg gcccgggcgg ctgcagcagg tgggcaccgt ggcgaagctc 840
tgagctctac cgttgaatcc ctgcaaaagg gtgcgggtga gccaggctga gtcgacggcc 900
atgggggtgc gcagcggcaa cctgcgggac aggtttttgg tgggtattaa ggaagatgga 960
ccatgtgtca ctgcgcgaca ggagcctcgc ctgcgtctca tctcatcat ttatgagat 1020
aactgcctga tcttcagggc tccagacatg gaccagctgg ttttgcctag caagcagcct 1080
tctcaaaaca aactccacaa ctgcaggala tttggccttg ccattaaagg cagagactgt 1140
ggcattggcg cagctaatga gttcaccac tcttgcacaa ctgaagccta tagattggtt 1200
caatttgaga caaacatgaa gggagagaca tcaagaaac tcttccccc tcttgatcag 1260
aatltccagg tggcctaccc agactactgc cgcctcctga tcatgacaga tgcctctctg 1320
gtgagattga atccaggat ggagagaaaa ataaaatgg agactttcag gccaaatatt 1380
gtggtgaccc gctgtgatgc ttttgaggag gatacctggc atgacctcct aattggtagt 1440
gtagaagtga aaaaagtant ggcattgcgc aggtgtattt tgacaacggt ggcctcagac 1500

```

WO 01/98353

PCT/US01/19862

```

actggagtcac tagacagggaa acagccactg garacccctga agagctaccu cctgtgtgat 1560
cctttctgaga ggggaattgta caagtgtgtc ccactttttg ggaatctatta ttcaagtggaa 1620
aaaaatggaa cccctgagagt tggtagccct ggtatccgga tgggtgtatg atgagtgatg 1680
gatccactag ggtgatatgg taagaagcctt cagcaaccag gaggatctga ctgagatctt 1740
aaccaacgca gaaacgatac atcagcaaat ccttatctac cagccctcaa ctatctttac 1800
ctgggaacac aatctcgact ttgtactttt caaagtgtgt tatgctccag gttaaatgcaa 1860
ggaaagtatt agagggggaa atatgaaagt atatatataa abttttaggt ctgaaggctt 1920
taaaaaaat taagctcttc aaaaatgcta ttttgaatgt latcatggct attacacttt 1980
tacttctctga ctltcaatatt gatgaataaa gcaagcttaa tgaatcaact aaaaagctgc 2040
aaaaatgttt ttaaaatgtg tgccttttat taccatcag tatatgtttt gggagaatg 2100
ggaaagcaaca gatcactgtg tccctgtgtg caggagcgtat gttaccacac tcacaaatgc 2160
ctaatatctg tctttatgtg gcccttgagt cctgttgact ttccactcat gtgcltttta 2220
ctctagcatt atggaatctg ggcctgactt gagtatggaa attctcttat agacttagtt 2280
ttaagtctct atcacacctt tactaagcua catcaagata atctgtttgt gtgtaactgc 2340
cagatatacc acctggaatt ccaagtaaga taaggaaag gatyacatlt aaaaagaaat 2400
ggaaattttg gagtgggaat gaaaggaaga cagcatgaac atattttttt cagtgcacat 2460
aatctttctg taacaagaaa acgaaccaat ttggtatgat ctttaagcaaa aactactcat 2520
gaatagtat gttgatgat tccctactt acatctttat ggtttctttg taaataatua 2580
atgtgaatct caatccctgaa aaaaaaaaaa aaa 2613

```

<210> 59

<211> 1778

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5677286CB1

<400> 59

```

gtctgccctt tgaatgggtg gatctgggct caccgcaacc tcggcctcca gggltcaagc 60
gatttctctg cctcagcttc cagtagctg gattacagg catgoyccac cagcccgga 120
teavlllgtt ttttttagtg agacaggtt tctccatglt gtlcagctg gtctcaaat 180
gacctcaggl gatccgcca cctcggccc caaagatgc ggaatcag gctagagcca 240
ctggcctgg cgggacaaa atgttttggt ctgtgtgta gtttggttt ttgtcgga 300
actttatgca ctcaagaglt ggcctgtgt ctggcctgtc cccgagtgct catggcagc 360
caaggcctgg ggaatgctca ggcagggtg atgtatttgc caggccagct cccgtccctg 420
ggactcagv gagttagctt atccctgaag gctagltgc ttctgctca tgtgggtggc 480
ccctctcag ggttctcagg ctkccaggga agtgggtgcc acagcgttgc ctccagctc 540
actggggctg gctggacttt gctctgacaa caccagctg gctggggccc tgcaccttc 600
cagggtctag ttccctgltt ggcgcacta tggcataggc cctgtgttg aacccgatga 660
gctgacccca caaatgccac ccgcctgct cccaggtt cagtggctta acagctgttg 720
tgcagggaac cagcttaag aatctttctt gctttctcaa actctccag aatctctgt 780
gactggaagg gggagtgact caggccctca acctatgca ggttaagggt ctactctctg 840
ggtgggcatg ccaglttct cagtggccc agcggggccc agcctccag gaagctctga 900
gaagctggtc tggctccag gtatttttag ctacagatcc agccccctc catcaccca 960
acagtcagat ccaaatcaaa cgcctccag gcagtgagg ctgggttct gccagctcc 1020
tggccagcag ggtgggtgg gcaagctgg gcaagctgg ctgggttct gccagctcc 1080
gcccggctgg aaagcagac ttgtgttat ttccgctc caattctgt caggttggt 1140
ctgtcataag caggatcac agcccatct ctggatgga gaaagtgar actgaggtc 1200
gagpctttt aggaagctca ggaacccctc tatgggtctc agtgatgac tccacagct 1260
ctgttctctc tcttccact cttagagtg cttggctccc tctgtctc ctggggaacc 1320
tcggcccccag cctgtctctc cagcagctc acagctctc cctgttacc ctgagggagc 1380
tcaggccccc gctgttagct ggtgtgtct gctctgtct ccgactctg tggagctgg 1440
caggcaactc catgatctga ccccggttac ctgacagcc ctgctgtgct tccctctca 1500
tggccagccc acccnagaac ctgaagagt tttctagct ccatgcatct gccagctgg 1560
gttaccacac ctactttccc tgcctgctt cagtgctgc caggctagt gtgcccaga 1620
gogctcagcc ttcagtaag ggttccctg ctcccaact ccatgcaat gcttcccta 1680
agactgtgac ctccctgag gctgggacac aactgctct ccatgcaat gttgcacaa 1740
aggaaaaaaa aaaaaaaa aaaaaaaa aaaaaaaa 1778

```

<210> 60

<211> 1234

WO 01/98353

PCT/US01/19862

<212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 6436791C31

<400> 60
 ccgagcagag accgctgct gggcggcgc cgttatccct taggtctgga gaggggacat 60
 ccgagcagag gccacttgcc gccagggccc agctcgltca gctccggctg accacagagt 120
 gccgcgggcy gccagagggg ccggaacccc aggcgcgttc gtcctctgtt cggcagagcc 180
 cgcgctgctc cgggagccgc tctggcagcg tatgctgcca cggggactga agatggcgcc 240
 gcgaggtaaa cggctgtcct ccaccccgct ggaatcctg ttctttctga accgggtgga 300
 taatgtaccc tatctctctg tggactttt catattctct tataaggtg tccgtctacc 360
 atatccaaac gctaacctag tactggatgt ggtgatgctc ctcccttacc ttgggaattga 420
 agtaattcgc cgtttttttg gtacaaaggg aaacctctgc cagcgaaaga tgcacclcag 480
 tattagcgtg gccttgacct tcccatctgc catgatggcc tctattacc tgcgtctgca 540
 gccctacgta ctccgctcgg aagccatcat gaaaggcacc ctgctctctc tctgtggctc 600
 agagcttlla cttgaggtgc tcacctggc tgccttctcc agtatggaca cgaattgaag 660
 taacgaattt cagcagcag ccactcagg tgaacccaca cgtattgctt ctggctctct 720
 agccacacca gtcgagattg gtggggcagg ttgtctctag aagcgtctgt tggctttct 780
 tcagcacaga catttgggca agcaactcag cataaggcca gtgggtacca tctctaac 840
 caggaccalc agcccaagg actctctcac actccagtat agggaggggc aaggttatcc 900
 ccactcctgc cctctcaga nccagtcctc tgcctgacct aagttctctc ccttgatcac 960
 cggcgccaga gcactctctg tggacacatc aggtcctctg ggcttcaagc aggacctgag 1020
 cccactgctc cctgtacagc ctgtctata cctgtccacc atgagcagc agagctctat 1080
 gtgtgtgggt ttccagagtc atgtgaaagc ctctccccc aatcctcaga gactgagctc 1140
 cccactcttt ttatgagctc atagtgttat tttctctacc tctctatgaa actnacttta 1200
 tttctataata aatataatct tctgtctaaa aaaa 1234

<210> 61
 <211> 863
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1820972C31

<400> 61
 gtaggtgctt cagccatggc atggatccct ctctctctg cgcctcttgc ttactgcaca 60
 ggtatccatgg actcctttga atggactcag gccacgtcaa cgtccgtgtc ccaggacag 120
 acagccacca tctctctctc tggcgagaag gtgggaagta aattcttttc gtggtatcaa 180
 cagaagggaag gccagtcctc tctcttaato atctatcaga atgggaagcg gccctcagag 240
 attgtctgac gattctcttg ctccaagtct ggggacagcg ccaacctgac catcagcaga 300
 gcccaagccc gggatgaggg tgaacttttc tctcaggtgt gggacagcag cactcgggtg 360
 ttgggctggag gaaccaagct gaccgtctca ggtcagccca aggtctcccc ctgggtcaat 420
 ctgttccccc cctcctctga gtagcttcaa gccacaaagg ccacactggg gtgtctcata 480
 agtgaacttt acccgggagc cgtgacagtg gctgggaagg cagatagcag ccccgcaag 540
 cggggagtggt agccaccaa accctccaaa caatgcacaa caagtaagc ggccagcagc 600
 taactgagcc tgacgcctga cagtggaag tccacagaa gctacagctg ccaggtcacg 660
 catgaaggga gcccgctgga gaagcagtg gccctacag aatgttcata ggttctaac 720
 cctcaccccc cccagggagc actagagctg caggatcccc aggggagggg tctctctccc 780
 ccccccaagg catcagcccc ttctccctg cactcaatna accctcaata aatattctca 840
 ttgtctactc aaaaaaaaaa aaa 863

<210> 62
 <211> 2521
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature

WO 01/98353

PCT/US01/19862

<223> Incyte ID No: 3286805CB1

```

<400> 62
gcttgcctggc cgtgtgttag ggcctggtgag tggctggggc tgtctgagcc atgaaacaact 60
tcaggggccac cactcctcttc tgggcagcgg cagctatggc taaatcaggg aagccttcgg 120
gaagaaatgga ggaattggga gttcaaaat gcaagatgc cttgaaacta cctgtctcgg 180
agttcctaac tggagggggc tgggaacatc tgggaatgt gscatcgga cgaattatgg 240
aattgactta ctcaactgc aggaacaacg aggatggaca gtaatacacc cctgatgaaa 300
tcttcacacat tcccagaaaa cagagcaacc tggagatgaa ctacgnaate ctggaatcct 360
gggcaaatat ccagagtagc acctcctact ccalcaacac agaactctct ctlttltcca 420
aagtaaatgg caagtcttcc actgagtctc agaggatgaa gacctccaaa gtgaaggacc 480
aagctataac taccaggtt caggttaaga acctcgtcta cactgcaaaa atcaacccaa 540
cttttagact aagctcaggt tttaggaagg aactccttga catctctgac cgtctagaga 600
acaaccagac gaggatgggc acctacctgg cagaactcct gggtctcaac tatggcaacc 660
acgtcaccaac cagtgtcgac gctggggctg ctcttattca ggaggaccac ctacgggcct 720
ctctctccca agacagccag agcagtcgta gtcccgtagc cgcctctgct ggaactgcct 780
ttcaaaacac cgtgaacttc aaatttgagg aaaaatatac ctgcagaaat gtcctcacca 840
agagtaacct ctcaaaccca accaactcca ggggtcagag cattggaggg gttcctcttt 900
accaggcct caccctccag gctcggcagc agggatcac caaccactg gttggcctcg 960
accgctctgg cclgcggctg catlctctca tcaaccacca cctgctactt gacttgccag 1020
gccccttggt ggaagaaatg tcaagacag tggaaactgc tgtgaaggcc tattatccat 1080
tcaaacacaa ccttggtgc acagatctca actctcccaa gttcaaktll caggccaaca 1140
cggatgatgg ctctgcgag ggaatatga ccaactctct tttcgttggg gcttatcagg 1200
aagctaccca gctctcaggg aatagggtat tctcctctct ccaaaagtgt gagcagaagn 1260
atccactcac tggtagtctc tcttgccctc ctggtcactc cccgtgcac ctgttatccc 1320
agatccacga ggaagggttc aaccacctgg agtctcatg aagltgcact ctctcgtct 1380
ctcgcagacc cgtgtgtgga gatgtgttcc aggtggcaaa agctgaattt agggctcttt 1440
gggtgtgtgg cagcagccaa gtacctgaaa actcaggaat gctttttggg ggccctcttc 1500
gcagcagagc cataaacccc atgacaaatg cacagtcag cccagccggc tactttccac 1560
tgaactctt tgaaaacccc aaggtatgtg tttctcagga ctatgagttg ggaagcaggt 1620
ttcgggtccc ctttgggggg ttctttagct gcaagcttgg gaacccctg gtgatctctg 1680
ctatatccag agatctaggg gaacgctctc tgaaaaagtg ccccgggggc ttcagccagg 1740
accagcctc cclcagagat ggaagccaa gtctctatgg agtcaaatcc gggctcttca 1800
caggaggggt cctgccccct gccaggtcc caactctcac ccggcaatcc ctactgagtc 1860
aggtggccac caatactgac atagtggaca attctggaaa tggggalcc tggatfaag 1920
actccagac caccagtggt aggtggggag aaccgataga gctgggggg gccatgaatg 1980
tcacccatgg ggaatgtgtt ggtctgtcag gggggcctgc agcgaagtc acagtggggg 2040
tcacccatct tctgggtgtt gttatcacct tggccatcta cgcacccgg aagttaanga 2100
agaagacata tcaggcaatt gaggaaggc agagtctggt tccaggcact gacgcaactg 2160
gagacacac ttaccaagag caggggcaga gtccagctta aactctccc cgaaatggt 2220
ttctctctac tccagtgttg tcaattgctg ccactctgtt ttcttaagca tlgaaatggc 2280
aagtgcacac aaaaactctg gctcggagag aggcctctca ccaagaggt ttgaagtacc 2340
acagcgtgtg ggaaggtggg aaccaggata ccaattcatt tccacagag acacagagaa 2400
gtgagtcaca gaatttgagc ccgtctctct gactgcccag ccagagacac tgattttctg 2460
aacctcttca cttgtctctg cctcttaagc attaaacat tctctaaaa aaaaaaaaaa 2520
a 2521

```

```

<210> 63
<211> 1765
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3506590CB1

```

```

<400> 63
gcactggaag tggccggtgt ttccattcgg tgatcatcac tgaacacaga ggaactaccca 60
tggaggttgg gctgagctgg gtttctctcg ttgtctcttt aagaggtgtc cagtgtcagg 120
tgacagctgtt ggagctctgg ggaaggctgg tccagccttg gaggtccctg agactctcct 180
gtgcagctct tggatccacc ttcagtagct atgctatgca ctgggtcccg caggctccag 240
gcagggggct ggagtggttg gnaatttat catatgatgg aagcaataaa tactangcag 300
actccttgaa gggccgattc acctcttcca gagcaattc caagaacacg ctgtatctgc 360

```


WO 01/98353

PCT/US01/19862

```

aaatgaacag cctgagagct gaggacacgg ctgtgtatta ctgtgcaga acgggagaa 420
ggagcccccga taatttgggtt gcttttgaata tctggggcca agggacaabg gtcacgctct 480
cltcagcttc caaccaaggcc ccatcggtct tcccccggcc gccctgctcc agggacacct 540
ctggggggcac acgggcccctg ggtcgccctgg tcaaggacta cttcccga aa ccggtgacgg 600
tgtcgtggaa ctacggggcc ctgaccagcg ggttgcaac ctbcgggt gtcctacgt 660
ccacaggact ctactccctc acagcgtgtg tggcctgctc ctccagcagc ttggggaccc 720
agacctacac ctgcaacgtg aatcaacagc ccagcaacac caaggtggac aagagagt 780
agctcaaac cccctctgtt gacacacac acacatgccc acgggtgcca gagcccaaat 840
nttgtgacac acctcccccg tgcaccaggt gccagagccc caaatcttgt gacacaccc 900
ccccagccc aggggtgccc gagcccaaat ctgttgacac acctcccccg tgcaccaagg 960
gccacgaccc tgaactcctg ggaggacgt cagtctctct ctcccccga aaacccaagg 1020
atacccttat gatctccgg acccctgagg tccagtgctt ggtgggtggc gtgagcagc 1080
aagaccccga ggtccagttc aagtggtag tggagggctt ggaggtgcat aatgccaa 1140
caagctggc ggaggagcgc tacaacagca cgttcctgtt ggtcagctc ctacccgcc 1200
tgacccagga ctggctgaac ggcaaggagt caaagtgc aa ggtctccac aaagccctcc 1260
cagcccccac cagaaaaacc atcccaaa ccaaggaca gccccgagaa ccaaggtgt 1320
acacccctgc cccatccccc gagggatga ccaagacaa ggtcagctg acctgctgt 1380
tcaaggctt ctaccccag gatctgccc tggagtgagg gacaaatgg cagccggaga 1440
acnaactaaa caccacgctt cccatgctgg actccgagc ctccctcttc ctctcagca 1500
agctcaccgt ggcacaagc aggtggcag aggggaacat ctctctatgc ctctgtatgc 1560
atgaggctct gcaacaacgc tacaacgaga agagctctc ctctctctcc ggttaaatga 1620
tgccatggcc ggcaagcccc cgtccccc gctctcgggg tccggcagg atgcttgga 1680
cgtacccctt gtacatactt cccagccacc cagcatggaa ataaagcacc cagcctgac 1740
ctggccctct gcaaaaaaaa aaaaa 1765

```

<210> 64

<211> 1264

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Inyte ID No: 003600CB

<400> 64

```

ccctgctcca gtcacacccc gaagctgact ggtccacgca cagctgaagc atngggaaac 60
lcatcgccgg actaatcttc cttaaaalll agacttgccc aglaaggact tcaactgacc 120
ttcctcagac tgaactctgt ttccagatata tacctcaagt cactgagatc tccagacccc 180
tgcgggtggc actactgaga gacagagtgcc cagggtggtt cctgaagtg cctgagcccc 240
aacttatcag caaggagctc atcatgctga cagaagtcat ggaaggtctg catggcttag 300
tgatcgccgt ggtgctccct tctctgagg cctgctctct caccgccatc aactaactgc 360
tcaggcagga catggccccc aagaytgaac agatactgaa accggccagt clccaggttc 420
ccaggccccc cctggccccc catcatccac ctgctgtcaa agagatgaag gagaactcga 480
cagagagaga catcccaatg tctgattccc tttaacggca tgcacagcac acaacctcag 540
alagcttgga tagctcctgc agtctgctc ctgcttgcca ggcacagag gatgtggatt 600
acacacaaat cgtctttctt gacccctggg aactaaaaaa tgaactcccc ctggactatg 660
agaaactaaa ggaactcaca gattatgca atgtcaatcc agaaagacac aagccagttt 720
ctgggtatct tgcacaacct gctctgtctg agccagcggg atatgatcaa gggccatgt 780
gaattccaaa taatttttaat ggggtccagt tctctatgga ttcttacct taattttag 840
ggaatgcnat tttttccccc ttaaaccaag catggggctc acaactctat ggagacaggg 900
caaaagaat gtggagaaga aactgataa ataacacagc gtctcaaga cccatggact 960
cctggtctgt acccaaaaaa gctgtctgtt ctcaaaaaa aaaaacaagg ctlygctggg 1020
aaaacagccc aatgccccgg caagaaaggt tgaatcaga tgttaggaag aactttcagg 1080
taagtatga gaactatgga gtccatcagc agagatagta gtgaagctc tcccagggg 1140
aaattttaaa aaggttgaat cagctgttgt agagtctat ttggcaatc catggtttaa 1200
tgacttccct ttgagctctt taattattgg caataaaca ottctttaaa agtttttaaa 1264
aaaa 1264

```

<210> 65

<211> 3416

<212> DNA

<213> Homo sapiens

WO 01/58353

PCT/US01/19862

```

<220>
<221> misc_feature
<223> Inzyle ID No: 1251534C31

<400> 65
gootcctcna cctcctcgtc tttctccggg gaacettgac gaagccttcc gcttgccctt 60
gootcctcgc gaatcccgcc cgcgcgcggc ccttgaggag caggagaaga cgcagccggg 120
cgcgcgcctt tagagggggt cccggcgcgc ccttgcccgc tcggccgcna cgcctccggg 180
ggcagccctt cctcctcggc cccgccttcc tcttgcccgc agcgcgcgcct cctgcgcgcg 240
atggggcctc taggactcga ggcgggttag gctcctaacg tagtgccac ggcactcaac 300
gaacggtag agttctacgc cgggacgtgg tcatcgagc ataagcgtgt ggaataatgg 360
cctctgatgc agtctccttg gctatcacta agtataagca cctcttatct cctgtttgtg 420
tggtcgggtt caaatgggt gaaggaccgc gacccttctc agatgcgtct agtgcctatt 480
atctataatt ttgggatggt ttgtctaac cctcttatct ctagagagtt attcactggg 540
tcataatag cgggatatag ctatatttgc cagagtggtg attallctaa taatgttctt 600
gaagtcaggc tagctgctgc tctgtgtgtg tactttgtat ctcaaggagt tgaatatttg 660
gaacagctgt ttttctctc gagaaggaag acaacccaag ttctttctct tctgtgttat 720
cctcactgta ccatgtttac cctgtgtgtg attggaalla agtgggttgc agggggaaga 780
gootccttgy ggcgcgcgtt gaattccttt atcagtgga ttatgtactc atactatggg 840
ttaactgcat ttggcccttg gattcagana tatcttgggt ggaacgcata cctgactatg 900
ttgcaactga ttcaattcca tgtgaccttt ggcgcacagg cactgtctct ttcaactgac 960
tgcccttctc ccaaatggat gcactgggct ctaattgctt atgcaatcag ctctcatatt 1020
ctctttctta actttctacat tcggacatac aaagagccta agaaaccaan agotgggaaa 1080
acagactgga atggtatttc agcaaatggt gtgagcaaat cagaataaca actcatgata 1140
gaataggga aaagcagaaa aatgggaaaa gcataaggag attaaattga actgggcttt 1200
aactgttgtt gacgtgagg acaactcctc atactatata aaatttcagg gaataacgaa 1260
gcaagcagga gcttggggtt ggggagaaaa gacaatgctg cctctatgctc tagtaactct 1320
tagactgagt aaagtgttaa taccataccc agatgtttta tctatgaagt ttttatttta 1380
aacatttttt ttaaaaaatta gcttgatat tctcagacc aaagcaatca ttaagtgaat 1440
ttggggattt tcccctgttt caactccagt tgcataaagg atgagatttt tctgttatct 1500
tctagtactt cctctctcgg tctgaatttt agacgatcac agaaacgctc tttatgaatt 1560
atttggataa attactaat atcttatcta cgcactgaaa tcaactgtgt tcaactttct 1620
tgtccaaagg tgaaaatggg tatcacctta aacttgcaaa ttggaattca ttgtcgacc 1680
ggaatgtgta aactctctca cctctagtaa gagtataatt ttggtgttaa ttaaatcttt 1740
aaaaatcgtt gatctctgta gaattctaga gcttgatga tgatggtgtt ggtgaataa 1800
agaaagcaatt gcaatgaagt cttgtctggt gacccagaga tcaactgac litagggaca 1860
aatcactgtg gggaaacaat ttttctggtt gaaagggcag catltgaata ctctgttag 1920
tagcagaata atattatgaa aattagactt atgtctgtat tgaaacatga aacacactat 1980
glcttatta gtaacatcat aagtagtata ctttatgtgt ctgttagaat atgttgattt 2040
ttatcaggct tctctgtttt tgattatggt ctgttctgta ttttttatat gtgggaattt 2100
acctactctt tctgttgaaa agaacattta aaattaaata aattttaakt aaaaatacaa 2160
ggagcttctt aatgcaaat tlaalgtaa ctttcaaat cactagtatt ttttttgctt 2220
ttatgacaaa tagcatcac caaacctttt tgtgaacta tcttctctt tcaatgtgtt 2280
taattttgga gtaagctttt ccttgtagct aagtgtcag atcttatcta tcaactaggt 2340
atgaagtata aaacattttt ggtgcaatat ccttgactcc ttgggtgtaa agattgttaa 2400
attcaatgct tgaatgtaaa aggtgtgttt aaacacaaa atgaataaaa gtgagagtag 2460
tcagaactat aacattcaat ttgttatcta caatgaagt atttcaalga atataagtga 2520
acaaactgga ataaagtagg aaagaatttg tatcatgttt tactacatag gttaatlttt 2580
taagggatgt tgcnaaggga ttactagaga aagacaaaat gtgacaaaa aaagacatga 2640
atatltctta agtatctcaa caacatgca agctgctgt tgaaggatgt atgctgtttg 2700
tacaactatb ttcagaatat ttgttagctt ataactattt battgtgat taaatlaaa 2760
taotttttcc ccaaggcat gcagtcata gaattacaga aaatttgcaa catataaagt 2820
agtttgatct aagaggatc aacacacttg tttgttgtct cagtgtgtaa tgaactgagt 2880
ttglaaactt ttgtgaacat tclgtactgt ttcccaagag ctatctaat cctgtacact 2940
gatttcagca caataaata acttctgctg tgggaaaaal atgattlttt ttgtttaatc 3000
agtttattat aatgtaaat gctgtctaga tgtttacttg atttagtraa tctcaaatgt 3060
ttctctgctg ctaatttctt atatagaana aatcaacta gaantgggt atctacactt 3120
ctcggtaaaa agcttatltt ctatagtagc atatttttao tttaaaatt agctgtcttt 3180
ggaaagactt ttacttcaga caaaaaat accattttct tctaaatta tcttaacata 3240
aatggggagc tctgttat ctcttaata aactattcag tcttaactat aagtctcttt 3300
atggagatg attalaagat gtcataaata cagccttttc cagtgccttt cataatttct 3360
ctattatcac agtatgttg catgcttat atctagaaac tgaatgaata agatg 3415

```

WO 01/98353

PCT/US01/19862

<210> 66
 <211> 2289
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1402211C81

<400> 66
 ggacatgccc ggagttgcag taccctcagg aaggtagcgt cttgactcgc gtggcgtggt 60
 tctgtgcctt gggagagcat gaatgggaag cggccagcgg agcccgccc agcccggtg 120
 ggaagaaagg gaaagagga ggtgatggcg gagttttcgg agcclgttac ggaagaaacc 180
 ttgaaaagc aggtggctga ggccggagc cgcaggagcg cgttcagta cgaagtcatt 240
 gtcatggaca tggacccctt tcttcactgt gtgatccca acttcactca aagccaagac 300
 ttcttggag ggcttcagaa ggaactgat aacttggact tcatagaga gtataagat 360
 ttatataagt tccagcagtc tgatgatttg aagaagagaa gagagcctca catctccact 420
 tttaggaaaa tctgttttga aaatttccgg tcttggcttt ctgatalttc taaaatttac 480
 ctggatcaca ccatggact gtctgtgtgt aaatattgaa tcatgtatgc ctgtgtgtgc 540
 catgatgat agctggaggg gcgcgggatt gcttctatcc tgtacatggt tcttccctgg 600
 gacagggagc tgggtgttac cctggacctg tccagcattg atgaacactt tccggcgaag 660
 cagattgtca agtctcttat ccttctgtgg aacaaactgg ttttctttga agtatctcat 720
 gtgtactttc accaggtgtc tgaagtgtct tctgaagaaa agtcacgttt gtctataagt 780
 ggctgtgtttc atgttccatc attgaotcgg cctcccaact actttgaacc ccccatactc 840
 cggagccctc acatcccaca agatcatgag attttgtatg attggatcaa cctactctat 900
 ctggacatgg attaccaagt tcaattcaa gaagagtttg aagaagttc tgaattctc 960
 ctgaaggagt ttcttaagcc tgaataatc scgaaggtct gtggggcctt ggagcatgga 1020
 catgtggagt gggagcagcg aggtccctct aacaaagggt tttatgagaa agctgaggag 1080
 agttagcttc ctgagatatt gaaggtgtgc atgaagtatt tccgtctga ggcactattc 1140
 ttgtgtctct ccaacttcac aggtctgaag ctctatttct tggcccttcc ggaagagat 1200
 ggatgaatg ataaaaaga ggcagaaacc actgalaaca ctgagaagg gactagccat 1260
 agtctctctg agccagagaa taatccagtg gctatcagca acacagcca acagagcaat 1320
 ggccagagay acccagagcc agaggaatb gaacaaaga aagaaacaaq tgttcccatg 1380
 tgcctagggy aactgagga ttggagacc ggtcactaaa ctttaattca tgaccctagc 1440
 aaggctgaat ttgcctlaga cttaattctg tactgtggct gtgaagcttg ggagccagaa 1500
 tatggcggtt ttacttttta cattgccaac gglaaagatg aagagctgci aacagtgat 1560
 cccgaagaca attctttggc attggtctac agagacagag agactctgaa atttgtcaag 1620
 catatttaac accgaagcct ggaacaaaag aaacacttcc caaacagaa acgtttcttg 1680
 gaotttcaat tcatctatta tgaatgacag cactgggcaa agctgacaa aatgtgtgac 1740
 ctctcaatt actgggaagt ctgaagagc taagcatgga gtcaaggaga actacatggt 1800
 agtctgtctg acagtgttct taaaactggt tgtcttttac taggactcat aatgatgtc 1860
 ctcaacagag adcttgagct tgcagctaaq taattattct ttgattaaaa aaaaaagct 1920
 ggtttttttt tttttaacat ttatctcttt ttcatattg gcttcttccg tgaattttta 1980
 agttcaattt gtttttattg aggtaaaaa ttataacat aaactgacc agcttaacca 2040
 tttttaata tgcattccag tggattaaat acattctcat tgttgtccag ccatcccat 2100
 catcatcttc cagaagtttt caatcttccc aatttctgt ccatctgaac aataactccc 2160
 catcttctct tccctagaca acagccatcc ctlttgtctc tatcatcaac tctctactc 2220
 atattctcca tgaagtgga acatcacagt atttgtctt ttgtgactgg ttctacttag 2280
 cataaagtc 2289

<210> 67
 <211> 4480
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 162347C81

<400> 67
 tccagggagc tccgggggccc gcgcgctgct tctctgagcc aggcaggcac tgcggggagg 60
 cggcggtgac aaccagggcg gtgggtgacgg gcaccgggct cgtgggtgag acacagtaac 120
 ctgggtgaac tctgcatctg gaaagctgaa gactgaagaa agataagaga catctgactag 180

WO 01/98353

PCT/US01/19862

tctggaaaca gggacatctt tggaaactcg ttttcaatca cagtaaacct ttgaagtgtc 240
 atcaattgga attgatttct tcatcttatt ctgcctattg ggaagaacat gctctcaagg 300
 atttttaagt ttcccttttag tttaacatgaa ctlttgtagga aacagagccc ttaaaagggc 360
 tgggaatcac aagaagagat tgaagacaga gaagcttgcc ctgttttctt tgcctcttca 420
 aagaaagaga tttaacagctc aaactttaga cagctgttgt ccagcttttag ccctcaagag 480
 agaatcaaat taacacacca ttgcacagct acaagacctg gtgaagctag gctgtcggag 540
 tggtagcatt gagaagacta cctaaagag acaagagactg cagtaaacaa agctcctctt 600
 taagttttag aggggcctca ggttccctct tgaattgaaa tagaataaga aacacagggc 660
 acacctcttt taggtgcagc tcaactttta tgaactgtia gtcgtgagg lactatagta 720
 tctcagaaga atttttcttt gcccaaaagt tttttgcoat accctgatat tctctctctc 780
 tcttgagagc ctgcctccat caacagagctg tactcttgatc tctctgactg tccatgtttt 840
 ccaoctgcaa ccatttgcat gtgtacagcc tactgtttgt ctccagtttt taaactgtac 900
 aagltgtgtt tctlaactct cctctctgct ttgttctggg gaggtggtta ttcactatct 960
 ggaatcacct tcccccctcc catgtgtctt cctcactttg agatcttttg acccttggtc 1020
 ttatttgaga gggggaaggg tgcataaagt ttctgtttcc ctggttttct ttgttaetcc 1080
 tctctgttgc ttccctctcc cacttttctt gtctgttctg ccgctgtgtg ggcctgggct 1140
 atgcctgagg cagattttcc catcagagct ccaacatgcc ccagagagct ggaagagat 1200
 tcaaacccag caagatctgc ccggtctctg cagccgcoat ctctctagtg ggaagctaga 1260
 cactctcttt tgccttttag tctccaggac taagctgtta tgtgtcactt gcatgtccca 1320
 tctcaacagc aattatgttt ctcttttgtt tggcccaact cagcatggcc accctcaagg 1380
 acccaggagt ttctctctga gcttggaggg atgagagcaa gtagagatgt ttccgggtct 1440
 cctcttcaaa aacagtgagc ataaaggga tccaggtgag catgaatgg ttgtccacct 1500
 ggcgcttcta cgttccctct cgtatgtccc actgcagtgt ctgtgacaa tgtgtggagg 1560
 aatttgatca tccctgcccc tgggtgaata actgtattgg tgcctgggac tccgctttat 1620
 ttttctcttt cctcctctcc ctgacagccc acattatggu ctgttttggc ttttgctccc 1680
 ttatgtctct ctaccacata gaggaaactct caggggtccg ccagctgtgc acaattggag 1740
 taatgtgtgt ggtctggtta ttctctctcc ctgtagctgt cctcagctga tttcaagtgg 1800
 ttctgtgtgg caggggagcc acaacaaag aacaggttac ggttaacttc cagggaggtg 1860
 tgaacccctt caccatggc tgcgttaaca atgtcagccg tttctctg agttctccag 1920
 cccccaggtt ttgggggaga ccaaaagaa agaaagacaa ttttaactga cctcctctcc 1980
 ttgcacaga agttccagat gggcagataa ctgttaagat catggaatg ggcctcagc 2040
 gagagctgag gagaacaaag tctaaaggaa gcttgagat aacagagagc cagctctgag 2100
 atgtgaaccc tccacctctc cctaaagcag acctgagccc ttacacaggg ttgggaacac 2160
 accctggctt ggtactaat gaggatagta gcttatggc caaggacagc cccccgacac 2220
 ctaccatgta caagatccg ccgggttaca gttagagcag taagctagct gcatgtccgc 2280
 attctctcag cgcgaagtgc agtctgtggg acagcttgaa ggaagcaacc tcaattgcag 2340
 agagcagcgc tcccccagc taccgtccag agcccagctt ggaacagagc agctctcgtt 2400
 ctctctactt tggcaaaagt tttaacttgc alccactatc cagtgtgtca cgtctctcca 2460
 gctcagctc accccagggc acaggctttg agctgggcca gtlgcactcc attcgttcag 2520
 agggacccac ctccacttcc tataagagcc tggcccaaca gaacagcaat ggaagcctat 2580
 ctatgacag ctgtctcaca ccttcagacn gccctgattt tgaagcagtg cagggcaggg 2640
 ctgagccaga cccacttcta ggtatacct ctcccttctc gtacagcagg ctggccacgc 2700
 aacgggaagc tgaagggcac ccagtttgg tgcctactgt cccaacacac cagagccctt 2760
 caccagtcgc ttacgacat ctgtcggccc acattgtggc ctctctctcag gaacagagag 2820
 agttgtctgc ccagtccccc ccactccggg gctgtgagga agaaccaggg ttggggagct 2880
 caggcatcca gtcaacacca ggcctgggccc atgcccctgc tactagtctc tctcagatg 2940
 attcaagagc atcacttttg ggaagactc cactgggagc cccagctgtc cccgcttttg 3000
 gcaagccaga tgggctaaag ggcggggggg tagggctccc tgaacagagc ccaacagccc 3060
 catctctgg ccgctcagag tcttaagca gccaaaagc caacctgtgt gctctcagga 3120
 cagaagaagt ggccttgagc caatctatga caccacaaga ctgaagatag ctgaagacca 3180
 cclacaccaa atccaaaggg cagcccaaga gcttaggtct agctctccctt ggcacagccc 3240
 agccactctc cagttagccc acgaggggag gactcaagaa ggtgtcaggg gttggtgata 3300
 ccacatagta gatctgggtg tgaagcttgc gcaactcccc tcccacatgc ctctggcctc 3360
 accccaagg gccccaggtg gccactctcc ttccctcaag gggctccctc ccgctgcatg 3420
 gacatttttt aaacacacga ttccaaagag atgagagagt ttttctaaa tgcagtaggc 3480
 ttggggagtc ggaaggttgg ggccttgaga ctgggttagc aacccccctt ttatctttt 3540
 aagaccttcc ctctcttgat ccttggaaca gactcagtg acatttgtgc aattgtctgc 3600
 cctggagggga accagatcat tttaaacna gaataaatt tttttattat tgttaaggat 3660
 tctatttttt toctctctgc cgtlaacagg tgtgtgtgta cataaalat atatatatat 3720
 atattataaa latcaaaagaa attatatact tatctggga tgggaaatag agggagggac 3780
 acatatacgg aggggagatc tactcttccc attcctcaga ccagcaggaa aagggggag 3840
 acgtcagctc ttctctgtg gttcctctcc atttgtccca gttactaat acggaalatg 3900
 catctctctc tgggtctaa ggtgattagg aagaagcctg gggagagctg agtctggaat 3960

WO 01/88353

PCT/US01/19862

```

tttgttcaca agaggggaagg acttgagag gagattagt tttctaggct cattggcatt 4020
tggtttccct aggaagggg tcaaaacttc aagacactgg tgggtggggg agatcaggaa 4080
aataacttgg cctagctcaa acaatatttg ataactccct ccttggggga gaggaattag 4140
agtgtgctcc tactggcccc ttggagcttc ccttagctta cccagttaac ttgattttan 4200
aatccaaaggc caggagagaa gaattcaaaa agcaattatt tcatccat gccaaaacg 4260
gggcataggg aaaaagatg gaggcctag gcccctcaga ttgtcccttg ggggttaacc 4320
ctcagccac ctaactatgg tgcctggtag agggatcac tgggttttaa cttctaaaa 4380
ggggagatcc cagcctccac aaagaggccc ttttattttt tattctgatt agccatttta 4440
aaccaacgag gaataaaaag aaatcctgat ctaaaaaaa 4500

```

<210> 68

<211> 1568

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1706443CB1

<400> 68

```

ctgtctatgt gccctctccc actgaacagg cagccagggg cctgtgcgag ctctgtctcg 60
tcacacagct catggggcgt gtgagggcca tctaccctca gctgctcttg gccclgctca 120
ttcaggccca ttaccacatc ggctccaacc tgcctggctg cgtgctctct cccaaggaca 180
ccacagaggg ttgacacgccc tctcccttcg tactgtgctg ctgggtgggt aaagtggtag 240
aaacccctgt actgagagtg ggtgtctctt atgagaccac gtttctggag gaccagggtg 300
gctggggagt catggagcag gtggagagcc accaccgggg agtggccttg ctggcaaggg 360
ccatgggtga gtactctctg caggagctgt gcccatctct ctacutctct atcccgctcc 420
tggagcagag cgacagagaag cacaggatca cggccaccgc cttctctctg gagctcctcc 480
agatggagca ggtgcgcagg atcccccagg aatctctctt ggggcggatg gcagaaggcc 540
tgagccacca cgaccccatc atgaaggtgc tgtccatttg aggcctggtc atctggccc 600
gcaggtctga gaagaccgcc aaggtgaagg cctctctgct cctcatggtg aagggtccta 660
agaacatgga tgggatgctg gtggtggaag cgtccacaa cctcaaggct ctcttcagg 720
ggcgggacca gaagctgatg gacagtgcgg tctatgtgga gatgctgag atcctgtctg 780
cgctcttcag cgacgcagaa gaggcgtgct cctctctctg catcaacctg tatgggaagg 840
tggctcagaa gcttcgggca ccacgcactc aggcattgga ggaagagctg gtccagacct 900
tgggtccctt actgtgacc atgcaggagg gcaactcaca ggtkagccag aagtgtgtga 960
agaacctgtt cgtgtgtctt tactctatgg ctggggagtt cccaagaaga gcttatagcc 1020
ggagcgcctg ggacacacaa cagcagagag tggccaaaat ttgcaagtgc cttgtgaaca 1080
cccaccgaga cagcgccctt atattctca gccagagcct gtagtatgct aagaactcac 1140
gggcttccct ccggaagtcg tcagtcagt tcatagggtc cctgttcccc tgcattggag 1200
gntaatgac agnagatcgt ctgaatgaag tgaagctgc totgtataac ttagagacatg 1260
accagaagc atcagtytgc atctacgag cccaggtcca gacacacatc ctggccagct 1320
gctggcagaa cctctggctg ccgcacggga actcatgggt gtgttactca gccaccaccc 1380
accgctggag ccccgctgtg gagaacctgc ccacttccca ccagcgccgc tcttggatca 1440
tgaggcactt gggctctctg aagatgtctt tgaagaagt cgtccctga gccccaacc 1500
ctcctcaggg tgggttgagtt ccagccatgc tccctatana tgtcatgttg cttaaaaaaa 1560
aaaaaaa 1568

```

<210> 69

<211> 1887

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1748627CB1

<400> 69

```

acgaagtccc acgcccacc cggttctttg tttttttcaa aataaacata ggttatattt 60
tttcaaacat gagcatctat atattaaaag gcttttcaal ggcgacacat ggtgttccc 120
acctgcaatc ccagcacttt cggaggctga gccaggtgga tcaactgagg tcaggagttc 180
gagacacgcc tggccagcat ggtgaaaccc tgtctchact aaaaatacaa aaattaghta 240
ggcatgtgtg tggcagcctg tagtccacc tacttgggag gctgtggcat gagaatttgt 300

```

WO 01/98353

PCT/US01/19862

```

tgaatccagg aagggagggt tgaatgggc ccagatgaca cccctgcact ctactgtggg 360
cgacacagca agactctgtc gaggggggga gtagcttcog gaaagggtac tgcatttccc 420
gtttctacct ccactgcacc cgtttattgc gtcttgctcc tgggtccacag agcctaaacc 480
gacacaccca acacgcccgc cggagttaaa gctaaaggaa ggcaggggga agtaattgaa 540
tgcagggggc ggaagccaga ggcacactgg gggagcagga aacggggggg ctctcgtctg 600
gggcatggag gctgtaccc taaagtaact tccgttaaca aacggaggtg cggagggagg 660
gttcccggga tctgaccggg gctctgcttg tggctgggc ggtggctctt gaggtgtgog 720
gtcttttggg ggttcgggaa gggggcccca ataccctctt tcttcaggtc ttaagaagct 780
ggcgtgtgtg caataaggaa cttaaaacaa tggaaaggcg gaaagtggag aggaggagtc 840
ctaagtcttt taagtccacc tgaactcagg ttgtccatgc aaaaaaaat gcaattccag 900
tgagttaaaag cacagggttt tcaattcttg catcacagtc aacttcacag cgaaccaagt 960
taaaaagagt gatgaagaa aagcccaaac ctcagggttg agaggggcaa ggcgtctcgt 1020
caactccgat ccagcactcc ttctcactg atgtctcaga tgttcaggag atggagagag 1080
ggctgctcag tctttgaal gatttccact ctggaaaact tcaagcattt ggaatgaat 1140
gttccattga acagatggaa catgttcggg gaatgcagga gaatttagct cgtctgaatt 1200
tggagctcta tggggagtta gaggaaact ctggggataa gaaaaaaca gccagtgaat 1260
ccaatctgga tgggtcttg tccagcttag aagaattgaa ttcttccata caaaaactcc 1320
atttggcaga tgcacaagt gtcccaaat ctctgttag ctaaaaaia atgtagtttg 1380
ctttctgtg attgaagag aagcagcag ctcttcttcc ccagccaat ccagtagcag 1440
aatcatcttc ccaacacagg atctcaaga attaaagtgc aagttcaatg atcgtttcca 1500
ctactgctt ttagttaatt tgaactcgtg taattcgcca taactggagg cctaggcctt 1560
gtcgaaagct tcatcaggaa atgtgacgca gagattgtgc tgcgtgctt catattgttg 1620
ccttatggga ttatacttga aatgcattgt gctgctgctc ttgacagggc ggaagggatt 1680
cttttctctg agctcaacaa acttatttcc gtgttgatcg caagctgttg atgcacaggc 1740
gtcttggtgc aagcccgagt tcaatttatg tttagacatc agtacacatg ctgggtgtgt 1800
ttccaaagcg cagtgtatag aatlaaata agtatttgtg agatttgcta ttttttaata 1860
aagtattgt ttaaatata aaaaaaa 1887

```

<210> 70

<211> 569

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Inocyte ID No: 1818332CB1

<400> 70

```

cggctcagg ggtgggggt gggggaggcc ctggagcgcc ggggtgatgg cggggccggt 60
gaaggaccgc gaggccttcc agaggctcaa ctctcgtac caggccgcc attgtgtcct 120
tgcacaggac ccggaagacc aggggtggc gagggtttac tgcatactg agaggccat 180
tgcgaagcgg ctgctcttgc ggcagatccc ctgggtgag aggaactctt gtcgaggctg 240
ctcttccctc ctgctccggg gctcaactg ccccnagcc caggagcgtt gcaggggaca 300
ggcgtggacc gtacagacct gctcaactg ccagcgccgc caacgcttc tcaatgatcc 360
cgggcattta ctctggggag acagggctga ggcctcgtc gggagccaa cagaltccaa 420
acactacaaa ccttgcctaa acacagccca ctccatttca gacggccttc ctgaggagaa 480
aagtcagact cagggttcca gtanccagtg atggattcac cccatccccc aaataaagtt 540
taactgtttt acattccaaa aaaaaaaa 569

```

<210> 71

<211> 2338

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Inocyte ID No: 1822832CB1

<400> 71

```

aacgtaggct gtaaggtct cccctgtgca aatggggcat ccacttgaga gtcaagggtc 60
tgtttgggtg gcaaggatag ccacttctga aggtagaag aaaaataagc caccaaattg 120
gtatctttct gtgaattgga catcgtgctt agagtcacca ttttccccc aacctggagg 180
aataagattt gtaactgca ttttatagct gaggaatctg actcaacaaa attaaattac 240

```

WO 01/88353

PCT/US01/19862

```

tccactaagc aattagcaat taaccaagtc tttctgattc agaaacccag ctgttccctg 300
ttcatatcca gccctcttat tgggabcagc atctgcccctc ttctcagtcg agcaagatcc 360
aggcagatca cactgggcttc ccagcactga atctggctca aggggacatg aaatttgatc 420
gggtcaltgg gctcaggagc atccactcca aaatlagcag tacaggaaag ggcgatggcc 480
ctaaacagca ttgacaggca gatcccatg taatctaaag ggcagagact ctctcactgt 540
ctgctctctc ctctgctctc cctctcagggc tgaccccaaa ttggacaccc ctgcatccat 600
gtccatcaaa caacacagct gccgtttctt ctgctgctt atgggaagct cccctctctc 660
ctctcgtttt ctctctcttc tgcctatcca caccgtgcac ttctcccttt ccttaagaaa 720
ccaccatcaa ctttagaggc agggaaaggc gtggtctctg caggaaaagc cagaatcccc 780
tctacccagc agagagagag aggaatggct gaatgttttc tccccagtc caaggcactg 840
ggtcttggtc gggttgaagg ttccaaagct ctctctctgt gtgtcggtga gttctggcca 900
acctgaaccc tctctgatag gccattggag ttcatcgagt cticagggac tccccatggc 960
ctggagatct ttgctctgct tacacgggag agggagaatg atttatagag aacatcatct 1020
aaatccactc tgaccattgt gtggccacac ttgctagatt gctatagctt aaatctagca 1080
ttgtgaagag acggggggagc ttggagctgc acaaacccgc gtctggactt ggctctttac 1140
cttggaaagt gaatgatctc gccaggaact ttagccctcc tgggtctcaa ttcttttctc 1200
tgtttctctg gaatgaggat ctctgctggg tggttgggtg atgtgggggc tgtgtgaaaa 1260
cagctctgca atacaagcca aaatagaatc atttctccac agagtatgaa ggtcaaatga 1320
pagaatccat ttcaattaaa tggaaaatta aaatggtaaa aaatgcaag ctgtattgaa 1380
agttctcagc ttctctaaa gggagctttt gactatgtaa gaaatctgca ctggtcccc 1440
ctaaatatac aaaaaagttg aaggaggcag aaggaggagc gatgcacgat gggcaaggac 1500
ttcactctgc gttgctggct ttgaggatgg aggaaggagg ccacaacact agaagctcga 1560
gccctcagaa gctagaanaa gccaggaccc gattctttcc ttgagccctc agnagggaac 1620
cagcccccgc agcaccttga ctttagcctg gtgagatctc cttaggactt ttggcaatca 1680
gaactataag acgaanaatg aagccactga gtcctgagct gtttcttgca ccagcaatag 1740
aaaactaatg cagggcccaa gaaatcactg gtgatgagat cgggaaagtg ggtcaggag 1800
gtctggatct gtcgatgagc ggggaaagtg ggtcaggag gtctggatct gtcgatgagc 1860
gggggaagtg ggtcaggag gtctggatct gtcgatgagc ggggaaagtg ggtcaggag 1920
gtctggatct gtcgatgagc ggggaaagtg ggtcaggag gtctggatct gtcgatgagc 1980
gggggaagtg ggtcaggag gtctggatct gtcgatgagc gggggaagtg ggtcaggag 2040
gtctggatct gtcgatgagc gggggaagtg ggtcaggag gtctggatct gtcgatgagc 2100
gggggaagtg ggtcaggag gctctgactc gtcgatgagc gggggaagtg ggtcaggag 2160
gtctggatct ggggtggggc tctggagctg aaggggaaat callygttca ttgtctatcc 2220
ttttgalttg attgaattt ttatataat atgtgaatt ttacataaa atttttttcc 2280
aaaaataaat aaaaaaagg ggtcttttgc aacctctc ctatctaaaa aaaaaaaa 2338

```

<210> 73

<211> 481

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1832219Cai

<400> 72

```

ctcggggcgc gttggtccgg tgcgtccctgc tctacagcta tggccggggc agctgcagct 60
ttccggcgcct tgggggcctt gtcggagctc ggggacttgg gcttcgcttc ctacggggcg 120
caggcgccc aactccaga tgcctacggg agggagctgt ttgaanaggc caacaaacac 180
cacttcttac acagcctggc cctgllagg gtgcccactt gcagaagcc acctgggct 240
gggttatctc tagcttcagg aacgacctta ttctgcacca gcttttacta ccaggctctg 300
agtgagagcc ccagcatcca gaatttgccc cctgggggag ggaacctgct actcttgggc 360
tggcttgctc tggctctttg agtcccttt ttcttaatta ctgggttttc tgggcagttt 420
tttttttaaa ggtttggagt aagaagagga ttaaaaagga aaggcaata aaaaaaaa 480
a 481

```

<210> 73

<211> 1255

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

WO 01/98353

PCT/US01/19862

<223> Incyte ID No: 1899010CB1

```

<400> 73
cggacgggtg cggacgggtg gtgctgctg gtctagctg gcttacagg cggcggcgg 60
gtgtgtgttc tctgttaaga gtgtactc cccgggttg atctggaat ggcactcctg 120
ggtcagacgg tgaagtcggc gtctctgacg acgctggcgt ggagtgaag ggcagccagg 180
gacagccccc ggcgcccttc cgaggttaga cggcccccag ttcggggga tcatggcatt 240
ggtggtggac cgaatggggg gccctggcg aatcccccgc gggctctgt tcaactgct 300
gggtgccatc tgcatttgtt tctcaacaa atggatttat gtgtaccac gcttcccaa 360
catgagcctg accctgggtc acttctgtgt cactggctg ggttgcata tctgcagaa 420
gctggacatc ttgtgcccga aaagtctggt gccctccagg cctctctcc tggccctcag 480
ttctctggc ttgtgtgtct tcaactaact ttctctgag acaacacaa taggcaacta 540
tcagctggcc aaggccatga ccaagccgyl galcatagcc atccagacct tctgtacca 600
gaacaccttc tcaacccaga tcaagctcac gctgattcct ataactttag ggttaactct 660
aaattcttat tcaatgtga agtttaattt ccttggatg gtgtttgtgt ccttgggtgt 720
tttagttaca tccctttatc aagtgtgggt aggaagcaaa cagcatgaat tacaagtga 780
ctcaatggag ctgctgtact accagctcc galgtcatct gccatgttc tggctgtgt 840
gccctctctt gggccatgt ttggagagg agaatattt ggtctctgt cagttctgc 900
ttgtctatgt gtgtgctat ctggagact agcttctctt gtagaactat caatttatg 960
ggtcaltggg aacacttccc ctgtcacta taactgttc ggaacttca agttctgct 1020
tactttatct gtaggtatgt ttatttttaa ggtatcact tcaattaat aggccttgg 1080
cattttatgt acattattt gccattctgc ctatccac ttttaactca gtagaactga 1140
aggaagttag agtaaacctg cacaacgtcc tttaattgggt ttttggtag aagaaatgt 1200
tgtcccaaga agataaaaaa tattgttaag tgtgcaaggt attaaaaaa naaaa 1255

```

<210> 74

<211> 875

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2008768CB1

```

<400> 74
tagctgagca cgcctctga gccgtcgtt gacacccagg cactctagta ggcctggcct 60
accacgaacc agcaggagag agaaagaaca ggcagctgt gagaagccaa ggcacccgag 120
tcagtcatgt cactaaggc ggcacagggc gccagccag agccagcacc agctccacct 180
ccacccgggg ccaaacccga ggaagacaa aggaaggtta aggaagcacc ggcacacact 240
cgaagggcgg tcagagacca taaggagcca tcggacaaac ctcaaaaggc ggtgcagccc 300
aagcacgaag tgggcacgag gagggggtgt cgcgcctacc ggtgggaatt aaagacagc 360
aataaagagt tctggctctt ggggcacgct gagatcaaga ttcggagttt ggaactcttc 420
aacgacctga ttgcttctgc gttcttctgt ggaagcclgg tcttctgtgt gagaagtcgg 480
cgatccatga atctccacta cttacttctt gtgactctta ttggtggagg tggagttttt 540
gcttttatcg atgtgtgtct tcaagaagac cacttcagag gcaagagagg caaaagcat 600
atgtgtgttc ctctccagg aaaggaaaaa ggaacccagg agggcaagg accagaccc 660
ggaagccac cagaacctgg caagccacca gggccagcaa ggggaagaa atgacttga 720
ggaggtctct ggtgtctgaa acggcagltt attttacagc uatagtgttc cactctcttc 780
cttctctctc ttctggaatg gttttctttt cctttttcat taccaccttt gcttggaaa 840
gaatggatta atggttcta aaagccnaa aaaa 875

```

<210> 75

<211> 2188

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2070984CB1

```

<400> 75
cggcagcggc gacggagagg gggcagcag cagtacggg agcagcagc tggacgggc 60
tggcgtctgg gccatgaacc cgtgttaagg cgcaggctgt gcagcaggg gtgcggggga 120

```


WO 01/98353

PCT/US01/19862

```

ggaggaggag gacgcgcggg tgaagtcttc ccccatgaac ctgaggggcc tcttccagg 180
cttcaacccg agtaaatcc toatctatgc ctgtctgctg ctgtctctctg tgcctctggc 240
ccttcgtttg gatggcatna taccgtggag ttactgggct gtctctctctc caatctggct 300
gtgggaagtta atggtcattg ttggagccct agttggaact ggaatctggg caggaalcc 360
lcaatctga ccagagggag aaacgttclg ggacltlaaa gccatcttga ttgcaaggg 420
catctacttg ctcllgttga tgtttgaagt tclgtctctg gacagcaatcg agagaggaag 480
ccattctctg cctctggctc tcatgcgcct gtctctctgt tcccggctgt ctgttgcagc 540
ttgggttttg ggtctctgac atgacaggtc actagagtta gaaatccctg gttctgttaa 600
cattctccag tttatcttca ttgctctcag actggacaa atcatccact ggccctggct 660
tcttctgtgt gtccctgtgt ggacltcaat gtcctctctg tgcctgttgg tctctacta 720
catttctgtg ccgclcttgt tcttgcctc talgpalgtg attggcagag agcgcaggac 780
aacactaacc atggccctga cctggatgac catctctgt cccctcttta cttttagat 840
lctgtctggt cacaacctgg atggccacaa cgcctctctc tgcclccga cttttgtccc 900
cctttgcttc togttgatga cgtctatggc aaccacatt ggacagaagg gaggaaacca 960
ctggcggctt ggtatccgca aagatttctg tcagtttctg ctlgaaatct tcccltctt 1020
acgagatatt ggaacattt cctatgatct ccatcagaa gataatgaag aaacccaaga 1080
gacccagtt cggagccccc ctaaaatgc accatgttlt cgaagaagag ccaggggtgt 1140
cattaccagg agcctctgga agtaltgtct cccctctccc aaatlaata tcgaaatgcc 1200
agattagatg ccaactcccg ggacagagct taagtggact gggacgcact ctctccgct 1260
tctctccccc cctctctccc ccgcagagac agaacagta ctggagctgg gtctccaggt 1320
acgtccatct catgctctgt ttgcctcag cgcctctcag caatcaca cgaaggagag 1380
cggaactggc aggttaccgg ggtgtgtgtgc agcagatcgc galcccagga agagtgtgag 1440
aaacaggtgt ggaatacgt ctgtctggga ggggtccag gtaacccct tcccgctag 1500
cccacactgg agatgtgtgt ttgcagagcc ccaggaagga acatctgtct atcaggtgct 1560
gaactcccaa tcaaaagtat tgtttgaaaa tgaattcttc caccggctgt acctccgca 1620
gtctgtgtgt caclccagg tctccagcac tccaggtctg tggctgggc agtccagtag 1680
aactgttaact atgtctctga tgcaccagct gtttagacac agcagagctc ttttctctgt 1740
tcttactgtg gaagtattt cctcttgggc atgtgacag cagttttcta tagctccag 1800
gctagagcct tcttaccgga gtgacttcat gcttctatag agagtattta tacaatgtt 1860
ttagctactt catatgcggt gtttaacct agttctgtac agcatattct gtlccagtat 1920
tttttttcaa gcttgtgtgt taggcacatg cctctctgtg cagagtgga cgcctgtggt 1980
aacctccccc ccccccccg tgggtgtcca cgcctctatg ggcacattgc acttctgccc 2040
tggaaclctg gcaggtacgt agtagctgtc actgccacaa cggcaacacc aagcagaga 2100
tggcccatgc tttctgacg cttccagaat agtggctagc ttcaaacctg acaagcgtg 2160
cttgaagcgg gaacactaga gaagtgtg 2188

```

<210> 76

<211> 1361

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> Incyte ID No: 2193240C31

<400> 76

```

tctggcctag ctctcaggga aagcttggtgc catcacctc cgtctgtta cgttggcct 60
ggcwaagcac tggacagcag cgaatgtgct atgaaggaa atcttcaaa gatattctga 120
aaagaglagc tgcacatgca ttgatctcac tgc-ggctgt cagtagaaga gacacgaaac 180
atgctttgaa agccaatctt atagacantt gcatggagca gatgaacac ataaatgcac 240
aacggaacct agattctctg agcctctggga agcagacatt gaaaaaaag gagggtggtg 300
ttattaaaga gtttaagcatt gccatgcagc tcttaagaaa ctgtctttat caaaatgagg 360
aatgttaaga agcagctctt gaagctcacc ttgtccctgt cttgcactct ctctggcctt 420
ggattttgat ggatgattca ttgatgcaaa tttctctgca gctcctttgt gctctatactg 480
caaatttccc aantggttgc agttctcttt gttagtcaa ttgtggacaa caccctgttc 540
aagctacaca tagaggagcc gtgagcaact ctctgaltgt tttgtactta aagltggctt 600
cccagatgco actggagaac acccgttct agcagaltgt ttttatgctt ctllcaaac 660
tggccttgtc gcatgactgt aaaggagtaa ttacgaagag taactcttta cgaactctcc 720
tctctctagc attgcacaaa ggaggaaata aacatctaa taactgaact attctttggt 780
tgaagttaact cctgaatata toatctggag aagctgggca acaaatgatt ctgaggcttg 840
atggtctgct agacttacta acagagatga gcaaatcaca gcaacagagc agccctttat 900
tgcctctctt tacltttcat aatgtttgct tccgcctcgc aataaacc 960
ctaagaaaa agtcaattact gtccttgctg cctgtctgga aagtgaagaat caaatgtctc 1020

```

WO 01/98353

PCT/US01/19862

```

agaggtattgg agcagctgcc ctttgggctc tgatttaca ttatcagaag gcaaaaacag 1080
ctttgaaaag cccatcagta aaaagaagag tggatgaagc atactcctta gcaaaagaaa 1140
ctttcccaaa ctcagaagca aacctctaaa atgcctalta tttgaaatgt cttgaaaacc 1200
tcgtgcagct ccttaattct tcttgagtcg catgggagtc tacacettga agctgacagt 1260
catcaaccag gtagctaaag tgaagccag ctgtgtgtag cagctgttac ctgaaagcyl 1320
gtacactctc tcaaaagtat tgatccctt ctctccatg agaaaagaaa ctgtgtgtac 1380
tccaaacacg tccagtggtg gcagctctcc agagtaata gcagctgaca actttctgtg 1440
ccttttcttt tctgttgaaa aggaatagaa agttctaggg acataaacat ttttcccttt 1500
ttctatgcca ttatttttgt aaaaatccta tttaacagtt atttaataa acagatttll 1560
g -561

```

<210> 77

<211> 1777

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 223517/CB1

<400> 77

```

ccaaatcag cgaggagggc agcttcgggg agctggggac catgggtcc aggtcaagc 60
agaatccaga gaacacattt gaagtatat ttgaagtggc ctatcccagc aagggtggca 120
ctctttccaga tcttgagggtg cagaggcaat tcccggagga ctacagtac caggaggttc 180
tacagacttt gaacagtttt tgtttccctt tctatgtgga cagcctcaca gttagccag 240
ttggccagaa cttcacgllc gtagtcaact acattgacag caaacagaga ttgggttct 300
gcagcttate ttccaggagcg aagagctgct tctgtatctt aagctatctc cctgggttcg 360
aggtatttta taagctgctt aacatcctgg cagattacac gacaaaaga caggaaaatc 420
agtggatga gcttcttgaa actctgcaca aactcccat cctgaccca gtagtgtctg 480
tccctctcag cctgcattct tattttactg tgcctgatac cagagaactt cccagctac 540
ctgaagatag aatctgaca gaattatttg tggctgtgga tgttaacaa atgttgcatc 600
tgtacgcaag tatgctgac gaacgcggga tactaatcat ttgcagcaaa ctcagcactc 660
tgcctgctg calccacggg tctggggcga tgcctaaccc ctgtactgg cagcaggtgt 720
acatcccctg gtagcggcgy calctgtcgg actactgctg tgcctccatg cctacactca 780
taggaatcca tttaagctc atggagaag tcagaacat ggcctggat gatgtgtga 840
tcttgatgt ggacacacac aacctggaaa ccccttcaga tgacttcag agcctccaa 900
acgaactgat ctcttccctg aagaacaggg tgaaaaaggt ctcacacac actgggagtg 960
gttggtccag agcgttctct aagccccagc ctgttctctt cgttagctac cgaacgctc 1020
tgaacactga gccggaggag ccgaacactt tctgtgagga agccttcgtg tccactacc 1080
gtctccggagc cagcaggcag ttcctgcaga acgcaacaca gctgcagctc ttcaagcag 1140
ttattgatgg tggattagat cttctcaatt ccggcgaagg ttccagtgt gttttttgag 1200
aggaaatcaa catgggctgag tacgttgcca gtgacaaact gtaccatcag tggctctcca 1260
ctgtccggaa aggaagtggg gcaattctga atactgtac gaccaaagca aatccggcca 1320
tcagactgtc ctacagttc gcaaaagatc atgcuaaat ggaataaaaa aggtgaaaa 1380
accgtttgna gcaaaaggac attgcgcaga atggctgcgc cccacccca gaagagcagc 1440
tgcaaaagac tgcaacgtcc ccactggtgg aggcacagga ccccaagctc cagagaagac 1500
ggcgcccat caccgtccac ttggacagg tgcgccacc tgcctccat gtttttaaga 1560
gaccaaagag caacatcgca gtggaaaggc ggaagagctc tgtccagagc cctgagcaaa 1620
aaccaatgca aacacagctt aactccaa tctacagaa aagcattacc catcttggcg 1680
ccaggttccc gacgagaggc tggacctctt cctcacatg acttaagcag ttgctttccc 1740
agactgggca gaggggtga cttcgcagtg tgtgcca -1777

```

<210> 78

<211> 1841

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2416227CB1

<400> 78

```

cgaaagagaa acccgagggg cgccggggac tggccggggc tctgcagggc tcaagtggac 60

```

WO 01/58353

PCT/US01/19862

```

ccatgagctc ccagagctaa cccctgaaca cccaggcggg caaaggggctg atgtcggtag 120
tcccacatcct ggaggggcag gctctggcca tctgtccttg ccatggcgct cgggcacctc 180
gcccctccttg ctggccllct cgtgggagtc gccagcaagt ccatggagaa caggggccag 240
ctgcccgacl gctgtatgaa tctggtgggc gtcaacgcca gctggccagg cgaagtcttg 300
tctgggtccag cctgttccag ggcctgggac ggggagggga ggcgagctg cgtccgctgt 360
gggaacggaa cctccagag ctacnaaggg tccaggtgta gaagcttgg tggccgggt 420
gcgcatttc ccatgaacag aagctcagg acccccgggc ggcacatcc tggggctccg 480
ggcgtggcgc cctccctctt cctgggcacg ttctlcailla gctccggcct catcctctcc 540
glagclgggt tcttctacct caagcgctcc agtaaacctc ccaggggcctg ctacagagaa 600
aaacaaggct cggccctgca gcttggtgaa gctgtgdaa tgatccccc gccacaglev 660
tcagtacgga agcccggtta cgtcaggcgg gaggcgccc tggcagggc caggatctcc 720
gctgctctcc cggggcgagg ccgtatcagg aatgtctgac ctggaggcgg agacacagcc 780
aucgacttgg cggcagggac cgggagccg accccttggc ggggaaccag acaagtggt 840
ggcatcgccc ggcgcgggg acagtcttg gacagcctc ggccttgagt cctcccgct 900
cccagcgagc gaagccaaag ggtcccggg cgcctgaggc tcttcccac catagccalc 960
tcgtttatcg gaacaggagc aggcctccat gagaactcag agcttcagat cggggccttg 1020
gggggtctcg gcccctccg gaacacggg gaggccccc cgcctgcagc caaagctggc 1080
acgactcttg gggcaggttc cgtctgacct agaaaagcca ggggctctag tggcgtgccc 1140
tcacagagcc acagcgggca ggcctctccc agcaacacca caaccagtgg ccggagaccc 1200
ctctgaggaac agtgaggctg gctcctcgtg cgttccagcc ggtgcgggag cagtgaggtg 1260
gaacagagcc aggaacagag tgcctgagac caaggttgct ctgggctgtc ctcccgtgtg 1320
ggcgagagcc caagcagcca gcccccatt tccgaagcty caggatagag cttctcttg 1380
atctctgttl ltaagcagaa attcatttgt ccgaanaagtc ctccagagct cttgggccc 1440
gctcggatcc cctggagccc catgctctgc tgatccccgc ccacgtgggg naggccacca 1500
tctaaccccc acaagtnact gctcaactgc acctgccag gctgcccctg cgttgagttc 1560
tggggtctct cccggagttc ctgggagaaa ggcgcgtctg tggccgctc ccgcacgcca 1620
ggcccggttc ccccgctggg tctcagagc cctgcggcac cggcacgctc tgccttagca 1680
tgggaccccc attcgagggg tggcctggcc ttccgggttc caacgctct tggcgaagtc 1740
cctgttggtt gccatcatg tctccgggac ctgggcccag gggagcgtgg gggcactggg 1800
tgtctgata taagctggc attactcaaa aaaaaaaa a 1841

```

<210> 79

<211> 1616

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Ineyte ID No: 24610/6CB1

<400> 79

```

agclggagac agagaggctg gccattgcgg gcaaaagcag aggcctagat caacagtgc 60
aaggactggg atgagggtag tgggtgtggg agtgaggaac catgatgag taagagacag 120
ttagtgctct ctgattccagg gaaggtggca gtytagtgac agaaacctga aggggagttg 180
atgggagaga tggcccaact gcttgggggg agaguccagg ctggagggcg gacctgggg 240
aalgagggcc agtgcatacc tcaactccc agctcctgt cctgggggaa ttctgggccc 300
cagatcccga tggtagcata gtcaagcttg gctgttcca cctcttctc aaaggaaggt 360
tctgggctcc cctacccnht gtaggggtgg caacgatccc tcttctctc aagcattlat 420
gaacacacct cctaaagtgt agctcatgca gacccaaatt tattacttta gctaatgatt 480
gaactactgc tctgttttag gacccaggga taccacatuu aggcgtgcgc cagagccctg 540
tcccctggag ccttgggttt caaataactgg ttctacctgg gagtggggag tagttggcct 600
ccagccttg catcctcagc cattgaatta gtttaaatgt gaagactctt agaaacaggt 660
aagcgtgcca gaagaatttg ctgctgctgt tattattatt ccagacatta ttataacag 720
aaagccgttc tgagctcaga agcgaactct cagcgggtcc ttctctaccc agaagaggtg 780
ttcataggaa cctgttagto agagtgggaa gggcccaagg gtcactaggt cactgttttt 840
ttgttggttt ttttaagttt tataattaat ataaaaatgg ggtclgcga tgtttggcat 900
gttgttcttg aacttttggc ctcaagcagt cctcccgcct cagcctccca aagtgtctag 960
attacgggca cagagcccca catccagcca agctgggtca ctttggaact aggcaggggt 1020
ggallgcaga cgtggggcac atccagccc ctgatttcag ggtccacat cacacctgc 1080
cctgctcaga atctggggg ggtgtccag aaaaaggtt ttacseacca gctgtgata 1140
agtcagaaat tttagaagcc actgtctca ttataaggc ataaggtca aggtatgaa 1200
tcccctgtc ttactcttt cttagaactc tgggttttta catccagtg aaagggtac 1260
tgtctcttgg ttgtgggac aaggcatgtc ttatgtgtt aaaaagata accactgaca 1320

```

WO 01/98353

PCT/US01/19862

```

aggttttttt tgaacctttt aaaattttact ttccagctgt aatccaggga ctttgggggg 1380
ccgaggccagg tggatcactt ggggttagggg tttagagcca gctgggtcaa catggtyaaa 1440
ccccgtctct actaaaaaat acaaaaattta gctgggtgtg gtgggtgtgca cctggantct 1500
cagctatttcg ggaggtcaag acaggggaat cgtttgaacc caggaggcag aggttgcagt 1560
ggcgcagat. cagccactg cactttagcc tgggtgacag agcggagctc catgtc 1616

```

<210> 60

<211> 1434

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1957517CB1

<400> 60

```

ccgccccgct cgttggaggc ccgggcgcga tggaggccgg ctgaggagcg ccgtgcctc 60
gctctggctc cccacaggtg caggaaagcg ccgcccagcc atggacgggg aggagcagca 120
gccaccgcac gaggccacag tggacactgt tctgcccgtc gaggcttcag agccggctcc 180
cagggtgctt tctggagacc cccagaacct gtccgacgtg gacgccttca acctgtctct 240
ggagatgaag ctgaagcggc ggctcagcg gcccaacctg ccgcyactg tgaacaaatt 300
gttggctggg gacggggaga ggtgtacgt ggtgggaca gccacttca ggcagcagag 360
caagaggagc gttgtgaaga ccatccggga ggtgcagct gacgtggtgg tcttggagcl 420
ctgccaalat cgtgtgtcca tgcgaagat ggaagagagc acgtgtctgc ggaagccca 480
ggagctcagc ctggagagac tgaagcagcg cgtgaggcag aacgggtcga tcttggggct 540
gatcagagtg ctgtctctga aggtgtctgc acacatcaac gacagctgg gcatggccac 600
aggtggcgag ttccaggagg ccttcaagga ggcagcagag gtgcctttct gcaagtcca 660
cctgggtgac cgaacctcc ccgtcacct caagaggggc atcgacgcgc tctcttctc 720
gcagagagtc aggtcggctt ggggcctggt cttcctgtca gaccccatca gcaaggatga 780
cgtggaacgc tgcgaagcga aggaacctat ggaagagatg atggccgaga tgaattggca 840
gttccdagac ctgaacgca ccatcgtctt ggaagcgagc gtctacataa cctacatgt 900
ggccaggccc ggcggcgccc tgaagctgac tgggctctat gacgcagagc ccaggagtg 960
cgtccccccc gtggtcgtgg gcttctgtgg catgggccac gtgcctggca tgaagagaa 1020
ctggagcacc gacctcaaca tccaggagat catgaccgtg cccccggcgt ccgtctccgg 1080
caagatgctt cgttgggccc tgaaggccc ctctctcagg ctgctgggct acagctgata 1140
ctgagatggc cgcgcgcgcg cgaagcaggt cctgtcgtg ccggcgcggc agtaactgct 1200
gcagaggtgt accgagggcc ggcacagta ggaagclgct cccgcgcgc tgggcccct 1260
paggagccag tgcctccgcg gcacttctgg gtgccaaggt cactctagcc cgcgcagggc 1320
ccttccacc cccatgggg gcttgggccc ggctcggctt gctctctgg gccagtcacc 1380
cctccccag ccaacccaaa taagaggatta ttaactgtc tgaanaaaaa aaaa 1434

```

<210> 81

<211> 2085

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 866038CB1

<400> 81

```

ccgcgcgcgg ccccgccatg gtgtcctgga tcatctctcg cctggtagtg ggtgatgg 60
gagctggggg ggaggaycca ctgcccctgg ctgaggagcc cagtaacct gctgcctcc 120
ctaggctcct ctttggcacc ctgtacccag cctattcttc ctacaaggcc ctgaagacaa 180
aaaaactgaa ggaatatgtg aataggatga tgaactggat cgtctttgce ttcttccca 240
cggccggagc gctcagggat atagtgtct cttggttccc cttctacttt gaactgaaga 300
tcgcctctgt gatctggctg ctgtccctct acacnaaggy ctccagcgtg cctacagca 360
agtctctgca cccaaagctg tccaaacaag agaaggagat cgaagagtae atcacgcag 420
cccgagacaa gagctalgag accalgalga ggttgggcaa gaggggccty aaccttgcg 480
ccaatgtctg agtcacagct gccgccaag gccagggggt gctgtcagag aagctccgca 540
gcttcagcat gcaggacctg accctgaccc gggagaggga cgcactgccc ctgcagaggg 600
ctgagggccc cctccgacc agccctggca gcttctlgc caccatcgag gacttaggag 660
atgcccctgc cctgagtcta aggtccagca caaacccggc agattcccg acaggagctt 720

```

WO 01/98353

PCT/US01/19862

```

ctgaggatga catgggagac aaagctccaa agaggggcaa acccatcaaa aaagcgccca 780
aagctgagcc actggcttcc aagacacuga agacccggcc caagaagaag accctggcg 840
ggggcgactc agcttgagcc cctccacccc cgcaggctgc agagcagaga tgaagcctca 900
gggggggctc aagaccagac cctgctccca cctgtatgca gtagcctagg tgtctcaggc 960
ccctaggccc ccagatggc catttcoggt gctggccag tggccactct tctggaaagg 1020
gcttggaaaa gaggaaggag gccagctgl gggggtttag ggtagagggt ggaccagagg 1080
ctgaggacgt agccaccaca ggaggtgggg actggtcggc ctccaccgct gtcccgctgc 1140
agcccccgcc tgcaccacc acccagtgc ttgctgaagc ccatagcaat ccgcttctca 1200
gaggtcctat cgtgtttcca ctgctogct tggtttggga gcaggggagg ggaagtctca 1260
gccagatgg accaaggagc ggcttgagg cactggggg aaggggagca caggggagg 1320
acgttgggga cctgggtgg ggcttcagg tgcagctgt gatggaagac agggattggc 1380
ctgtcttcca gcaccagga tggccaggcc agagcgcag ctgggggctc ttttcttgt 1440
cattgggtgg ggctgagtc cccatgttcc caccataaaa aggggggttc agggcttgt 1500
gagtggtgtc tctgggtct agggctcggg gtatgttggg tcaaggactg tccctcagc 1560
agtgcctcc tccaccctg agcccccag tcatctggcc ctttccctgc tcaaccctcc 1620
atctatggtc ctgagctcca gaggaccag ccatgagag aacgggggtc tggggggcct 1680
ctacacggtc ccatgacct tgcctccctt tgggtcaccc cattgccacc gtgcccctgg 1740
gttgggctcc cgtgtctctc agggccacc cctgtctct ctgggtacag cccctgtcga 1800
gttgggctct cctctcggc ctgggggtc atccctctt gctcagggtc cccatctgt 1860
accaagaga ctgggcctca agaccctgg cagaaccag caccggccc 1920
agcttttgc accgcttccg tcaaggccag gaggaaacag tgaaggagaa agagaaatgc 1980
aggagccgg gggtcccg ttccttggga agaggttccc cattggacct ttggccctgg 2040
atgagccaat aaacaaact ctggcaccct aaaaaaaaa aaaaa 2085

```

<210> 82

<211> 904

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3869704CB1

<400> 82

```

ggggagccca gctgtgctgt gggntcagga gccagagctc tgggaatctc accatggcct 60
ggacccctct cctgtctccc ctctctcctt tctgtacagt ctctgaggtc tcttatgagc 120
tgacacagcc accctcggtg tcatgttccc caggacaaac ggcaccggct acctgctcgt 180
gagatgcatt gccaaaaaaa tatgcttatt ggtccagcca gaagtccggc caggccctgt 240
tctgtgctat ctatgaggac aacaaacgac cttccgggat ccttgagaga tcttttggct 300
ccagctcagg gacatggcc accctgacla tcatgtgggc ccaggtgag gatgaagctg 360
actactactg ttactcaaca gacagcagtg gtaatgatag ggtgttcggc gtagggagca 420
agctgacgtt cctaggtcag cccaaggctg cccctcggtt caotctgttc ccaacctct 480
ctgaggagct tcaagccac aaggccacac tgggtgtgtc calaagtgac tcttaccggg 540
gagccgtgac agtggcctgg aaggcagata gcagccctgt caaggcggga gctggagcca 600
ccacacccct caaacaaagc aacacaaagt acggggccag cagctacctg agcctgagc 660
ctgagcagtg gaagtcacc aaaaagtaca gctgccaggt cagcactgaa gggagccag 720
tggagagac agtggccct acagatgtt cataggttct cctctctcag ccccccac 780
gggagactag agctgagga tccagggga ggggtctct cctccccc caggctcaca 840
gccccttctc ctgactcaa taaccctca ataatatctc tcatgttcaa tcaaaaaa 900
aaaa 904

```

<210> 83

<211> 1496

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1415179CB1

<400> 83

```

gcgacagaa gttttagag ttccaggagc tggctgtgca ggcctctcgg gccttctgt 60
acatgctgc tggggagcgc cagctactcc tggcccccag ctctcaggga cagacgtgtg 120

```

WO 01/98353

PCT/US01/19862

```

tgtatagaca catttggtgt gntctcagtg cctagggggt gccaggccct ctggtctct 180
cagggtgggc actggaggat ggggtggggc ctctgtgagc acccgtatg cctgtgtgaa 240
gacaaaggac caacctgcac tcatgccctt gtatccgcac accgccagtg tctacaggcg 300
caccocgttg actggcccag gaggtcgtgc atcatcactg caatcagcat gaacacagg 360
caccocgttg acacctaacg cccagttccc gtgaccccaa ctgcaggccc cccctactct 420
gcagcctttc ccatgctctg cccctgtctc catgacagag tggcgggggg tctgtggagc 480
gagggggggc gcctctctct gtctttgtgc ttctctctct gccctctggg tctttctgt 540
tgggtctcca ggcattgttt acctgctctt gctctccgtc gtagcccaaa gccatctct 600
tcatgcttcc cggggggcac tcaacctgc cccaccccg ctgcactgbc acaggagag 660
ggacggacac actcatcaca cactcaacag ccccgactgc acccggtgtc catctataa 720
calgtcaggg cactgtgaca taacagcaga ttacgacgtg cctaccaggc cctctgtta 780
ccttgcctct ctgcttggag ggggcccctt ctgctcacc cccagctgac ccccctgga 840
ctgttaagga tgaatagggt ggtgcagct tctgagggcc tcttggct cctgggggca 900
aagccctctt gcccgaggca atcaaacag cccagagagc aatccgggg agctgccagg 960
gggtgtgtgt tctgtctcc ttggcgtttc ctctatggcc gtgaatccca gctcaccat 1020
ggacagagga cctccctcca tctgaagacc ttctgtctct cctctctgta gctgtctct 1080
tctcaggtct ggaactgtcc ctggttccac ctgtctttgg gatgtggcc ctgagctggg 1140
vgggtctctt ggtgtgatta gagagctggc cagcttgcaa gcacctgag taactggag 1200
ggtgcccaga gtggagcttg gctgagggc gtgttgcag cctggccca ccaatccaa 1260
gagggccccc ggagcactgc ctttcagtug accgggtggg gactgtggg agggccccc 1320
ctagctgagc ttctgtctgc ctctgtctgc tctgagagtg cagggtatg ggaagggga 1380
caccatgtgc caaatgtgt gtagctctct ttagcagga taacagggt ctggggggca 1440
atgctatgat taactgtct caataaaaa gttcccgccc caaaaaaaa aaaaaa 1496

```

<210> 84

<211> 2837

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1664792CE1

<400> 84

```

gtgtcagctg cgcgcgcgc ctggcgagag ccacgcgcga ccacttccac atgcagtg 60
agttctctgt cgggcccctg atgcagacca tctccacaca gcactggag gtccgtgtgg 120
ccgcacttga agccacaggg gcactgtacc attttggcaa cgggaaglcc ggggagcag 180
tgctttccca tttgtctcag cgaactgttg atgacgcccc gcaaggtccg cgggagcttg 240
cctcctgtgt gggcggtgtg ctgctgtgtc tgcgtgaccc ttactctctc ttccacaagc 300
tcctcctctt gctgctcagt agctcaacg accaggtgac tgggttcagg cagctggctg 360
ccagcctctg ggagagagtt ggcctgcagt ggcagagga gaatgagg gagctgaagg 420
acaagctgga ctttgccctt ccaaccccac ccaatccac tccacatgag cgcgcctctg 480
tgctgggctg ccgggagctc gtcttcagga accctcccaa gatctctct gcccctgtgc 540
acgcatcac cgaactgggt gtggggaccc gactgaagtc ggcacagctg ctccagctgc 600
tgcgtctgca tggcgaggac cagccacgc agcacctgga ggtctctctc cggacctgt 660
tccagcctg caccgacag gaggcagccg ttgtccaaag ttgtaccaga tctgagagc 720
tctcgaggac gttgtcagc cctgaggtgt ttctgagct gatctctat agctgaaga 780
agacgctctc tgcctcggc ctctctgtgc tggcctccgc catgagggtg tgcctccagc 840
aagccctcca gccgcactg ccagcactg ccacagagct ggcacagccc cacatctgc 900
agccactga aaacgaacct taactggagc gctgtgtgt gtgtgtgac gctctggtgt 960
ctgtctgca tgaagactgt ggcgtggcca gctgtcagct ctgtgagctg ctgtgtgcaa 1020
tagtggccct cgcaggtgct accggcctga gggacaagcc acaggagagc atgagctcac 1080
tggccatggt ggggggtgtc agcagctgac aggaactcta ccgcaagcac attggtccc 1140
tcctggagcg ggtgacccgc tggcacttg actgagccgc acacgcgcg gagctctg 1200
agttcagctg catctgcga cagtcaggcc ctgcccctgg agagccctg ccaacagctg 1260
tgcacagctt gaggcctgt ctgcagcctt ccaagagccc gcagatgct ctgaagctgt 1320
tctccactct gtccacgctg ctgctcagag ccacggacac catcaatcc oaggggcagt 1380
ttccacgcta cctcgagacg gtgacaaagg acatctgtgc cccaatctg cagtggcatg 1440
cggggagagc agcgcgggac atccgcatgg ctgcccgtgt ctgctctctg ggcctccaca 1500
gcagcgaggt cctgtcgaca gacagatac gggacgtgca ggaacactg atgcccagg 1560
tctgagccac cctggagagc gattcgagga tgaagcagc gactctatg cgtattatca 1620
acagcttctt aaaaacctg ggcggcagtg cggatccaga gaactctac agatttatc 1680
ctgactctt aaaaaccta gatgacgtgt ccaacagctg gaggatgca gccgcctcca 1740

```

WO 01/98353

PCT/US01/19862

```

ccttggtcac cttggtgcag tgtgtcaagg gtgccaacgc aaaaactctac taccagagca 1800
gtgtccagta cctgtaccga gagttgctgg ttcaccttga cgtccacagag agggccatcc 1860
aggatgcgaat tttagaggte ctcgaagagg gcagcggggt gtccacagat ctctggttga 1920
gggagacgga ggtccatcac cacaagcacc gctcggccac ctactcgagag cagctcctgc 1980
agcatgtgca ggcctgtcca gccacacagt gaccacgctg gtttcagcca cggcacaacc 2040
ttgtccccc cttgaccaga gtttgtggcc tttaaatctc ataaacaagg cactctctgt 2100
ccagcagtga gantgtgana goaagaatgt actcctcagg acaccltgccc actcttcccc 2160
tggatataca gctctgagt ggaattgca ttttatgtga ctgtgtctgt tcatcgagag 2220
ggctcccaa catctcgagc tgatttgaa ttaaagtaa gtccagagcg ctctccccc 2280
agccactcca gcagcatctt agattttaag cctcagctgc gcagctggtt catgaactat 2340
tggctgcatc ctgcttgggt gccaccacc aaggltttta cctactaac aaaaagaaa 2400
goagccaaag cgtatagaaa gaactgaat ctcttttttg gttctgtcta ctgaattta 2460
atatctcagt gaacagacta aaaggaattt agaatcttaa caacttacca gattttctct 2520
gttttaata tactgggact ttaaggtht tatgtccggt caccgtactt ttaagctcga 2580
tgttaatgct aacagtgttg aaaaacaat ttcattgagt ctaattgttg ttgcccctat 2640
aggtagcagg aagctaaagt tgcatttccc tctcgacat tctacacca agtgcctaaa 2700
agatctcatt ttaagtgggt agtcttacc gaagccattg tgttcacag ggggaatgc 2760
cgtatatatt tttcaacaaa tattaacgtt tatcttttca tgtttgaaa ttttaattag 2820
aatatttgtt ttaaaaa 2837

```

<210> 85

<211> 1123

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2079396CB1

<400> 85

```

taagcttbcg accgccattt tttttttttt tttttttttt tttttactcg tgttgcatt 60
ttactgttaag ttgaaactg tcaatacaaa aagtlacaaa gaattgttta taaagaagaa 120
gaagaaaaaa gcaaacccaa ccaagctcaa gcttagatcc gtatttggat tcaagctcgtc 180
tcggggcgga gccaggctc accgctccg gattaaaga tcccccgag gagtgctgc 240
tgtgaactcg ggtggggagg ggcctcggtg ctgttccag caaccacaa cctctctct 300
agtgttgcga ggaatagcca ggaactcag ctgactgcat cagaacctga gaagctggag 360
gctgaagcca gacaccagcc tctcaggact tggggcaact atgagccac tcagctccac 420
aggctcgaat ctttgggag gggagggttc cagtctgcac tctgcccctg accatcaggg 480
caggggagat acactggcta tgggcatcat ctatagcagt ttctccagcc ccagccctag 540
aatcaggcca agctctcagc actgcgtggg ttgatcttg aggtattctt accacatcc 600
aggaactggg ggcctgcgct catggglttt gttgttaaga gacagggtct cgtctgttca 660
ccaggctcgg agtgcagtg catgatcac gctcactgta gcttgcact cccgggctca 720
aggaactcct cggcctcagc ctcccgagta gctcgacaa caggcctgtg ccaggactcc 780
tggcttgcct gtcctcatct gctatgact catgaatccc accagctacg acaagctcct 840
catcaagggt aggtcgggct cctgggtggt ggcacacat ctgggcccgg ccgacccct 900
gatcccgctt cctccctcca cagtggttc ctaggtcacc gcaattctgc cggggatcc 960
ccatggtgct catctgctgc aagcagacc tggatgttc cgcgaagttt cggggagatg 1020
tggagagngt cttccggag gccgccaaag tggclctcag cgtctgaag aaggcgcaac 1080
ggcagagaaa ggcaccgctc tgcctgtctc tctgaccag ggc 1123

```

<210> 86

<211> 1549

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5390115CB1

<400> 86

```

ttccggggga ggcgcgggga ggcgcgggga tgcacctga ttcaactgac ttggaaccag 60
aagatgttgg gctcttccaa gaagtcgaat ccttctcgg aggatgaaga tctgccagct 120
ctgaggaattc tggctgggaa gaaaagagc ctcaagcttt gaaggccatt ccatggttag 180

```

WO 01/98353

PCT/US01/19862

```

acagagctga tgtgaaaggt ggcacatctg ttgcccagg cacatgacct tctgaagtga 240
ccaggctgaa aggaagcnaa tactcgtgtg atctctcaac ccgtcactca gggtagcga 300
atcaagactc attggctcac tgcagcctag acctcccagg tggagcaatt cctctgcttc 360
agccttctga gttagctggga ctacagttgg ttctaaagag tggtagacta gaagagagct 420
caggcagcaa ggcacttggg caatggcttc tgaactagac ttctcaacct cggaggtgac 480
cgagcccaat ttctgggaga aactgctacy gtaaggactc ttctggggag ccalettcna 540
gctcactctg gtgctggcca tcatcgtacc catccccaag tccccagayy cggaggtgca 600
acgtctctag ccagaagtg ctgaagtgac gaggagagcc aaggtctgtg ttctctctgt 660
gacacagagg ccaagaaag agacaaaga gaagcgttag aagagagagc ctgagagagc 720
ggcggggcag ggaaggggtc ttggggacag cctctctagg aatctacatt gtgtccccc 780
gctctcagg ctccagggtc gaggaggtg tgaagcccta tgaagcaga gatctagaca 840
gtcglcaacg tcccaggtct ccagctgggc aatccacca ttctctctcc ttctgcttct 900
gtgacgggtt agagtcaggg gggctgaac acactgtgag catagactgt attaggtttg 960
ttcayaaagc ggttcagctc acagagtcac atttctctgc ttatgtatgt gtccctctct 1020
gagttgcccc ctctctgttg gtttaccta catttgggag tcatgtctca atgtctgaca 1080
gcacacccct tccatratc tgtgacacac agatctctct ctgactcagtc acccttcttg 1140
ctgctgtgta gacagagcca ggctcactc gtttgtttag gcaagatgc catggacatg 1200
cagcgttagt galcccaacta gctggagacg ccaggcccaag aaaaagcctg gctgagagac 1260
cagcagacag ccaggccggyy glaggcagtg cctgctctgt cccatcagg tgcaggggag 1320
ttggtcagag gctgcatat ttccctgggca caaactctct gaggccctga aatggtaggc 1380
tcytcacttt cagaacaccc tcttctcacc ccalcagag acgtctcagg tctgctcttt 1440
actctacact gtaactcccc catccctcca attcttctat tccctgacca gtgagaggt 1500
tcttggggga agtatggtga ataaactgac atgcatgctt caaaaaaa 1549

```

<210> 87

<211> 4820

<212> DNA

<213> Homo sapiens

<220>

<221> misc_features

<222> Incyte ID No: 1403326CB1

<400> 87

```

ctcacactct aagcgtgtga gaatagctgg gaccacagtt ttagctgtga gttcaactatg 60
ggggcaactc agtccccctt gtttgtgaa ccatcttacc cagaccctgt catcaatgta 120
ggaggaagtga ctcttgagga gaaaaacaga aaaaagctgc agaaaattca gagagaccaa 180
gagaagggga gagttatgcy ggcctgcaatg gctttattta actcaggagg aggagtgatt 240
ggaatggcca agaaggttga gcatccctgt gagatgggac tggatttaga acagtccttg 300
agagagctta ttaagtcttc agatctcgag gcttctcttg agacaaagca acaagggaagg 360
tgtttttaca ttcttgtaa atcttgagag agtggccttt tccctgaaga tctgctcttc 420
aagccccccc ttgagagctt cagttcttca ttatacccta gatctgagac ctctgtgctg 480
tcaatggact caagagagggc attctgttct ctgaagacca aaaggaagcc aaaaactctg 540
gagaagggac ctcttcacca aattcaccaag ggtgtatacc aagagctccc taactcggat 600
cctgctgacc caaactcggga tccgtctgac ctactttccc aaaaagacaa tcttgaaatc 660
ggtagaaacc tgccttttcc tgaagctcag tttagtaaat ttaaacagt ctctacaaaa 720
caacttccag aatatgtaaa aaggcaaat ccagaaatag tccctgcaat tgcnaaacct 780
ggagagagct atctttttat tggagtgagat gataagagta gggagctcct gggatgtcca 840
aaagcaaatg ttgacccctga ctcttttgaga aggaataatg aacacagcat atacaacata 900
ccttctgttc atttttgcca accccaacgc ccgaiaacct tcaactcaaa aallgtggat 960
gtgttaaaaa ggggagagct ctatggctat gcttgcatga tcaagataaa tccctctctg 1020
tgtagcagtg tctcagaagc tcccaactca tggatagttg aggaacaagta cgtctgcagc 1080
ctgacacccg agaactgggt aggcctgatg acagacacag atccagatct tctacagttg 1140
tattcaagatt ttgaatgcaa gctgagctta tctagtgagg ctccctctag cagacagatg 1200
tattcaagaa aaggtctgga acacaaagct gatctacac accatttatt tccagttcca 1260
ccagacattt tggaaatgtac tccagagttc ctctggagg agctgtcttt acagcatgaa 1320
ggactaagg agcttaataca caagcaaatg cgaactttct cccagggaat tctgactctc 1380
tctagaagct gggctgtgga cctgaaactg caggagagag caggagtcac ctgtgagctc 1440
ctgctgatac caagacagc caccctcatt ctctacacca ttctcagggg cagagatgca 1500
gagggacagg actactgcac tgcagacgc tttaacttga agcagaagct agtagaatg 1560
gggggtaca cgggaaggt gtgtgtcagg gccaaagctc tctgcttag tctgagagc 1620
agcgcagagg ccttggaggc tgcagtgctc ccgatggatt accctgcctc ctatagcctt 1680
gcaggcaacc agcacatgga agcctgtctg cagtcctctg tcttctctt actngcttc 1740

```


WO 01/58353

PCT/US01/19862

aggtctctct ttagtgacca gctcgctgt gaggttttaa atctgtctac agccacagcag 1800
 tatagagatat tctccgaaag cctccgaaag aacagagaa tcttcttcca cggcttaccct 1860
 ggctcagggga agccacacat ggccatggaag atcatgggga agatccaggaa tctgtttccac 1920
 tctggaggac acagaaatct ctacgtttct gaaacacagg cctcgaggaa ctttaacagt 1980
 gatagaaata tctgcggagc agagaccccg aaaaactttcc taagagaaaa ctttgaacac 2040
 attcaaacca tctgcaattga cgaagctcag aatttcctga ctgaagatgg ggaactggtat 2100
 gcgaaggcaa aaagcatcac tccgagagca aagggctggcc caggaaattct ctggatcttt 2160
 ctggattact ttcagaccag ccacttggat tgcagtgccc tccctctctt ctccagccaa 2220
 tatccaaagag aagagctcac cagaatagtt cgaatgcag atccaatagc caagtactta 2280
 caaaaagaaa tgcagtaaat tagaagtaat ccttcattta acatccccac tgggtgctct 2340
 gaggatcttc ctgaagccga atggtcccag ggtgttcagg gaaccttcg aattaagaaa 2400
 tacttgactg tggagcaaat aatgacctgt gtggcagaca cgtgcaggcg cltctttgat 2460
 aggggtctat cccaaagga tgttctgtgt cttgtcagca ccgcaaaaga agtggagcac 2520
 tataagtagt agctcttga agcaatgagg aagaaaaagg tggtagagct cagtcatgca 2580
 tgtgatctgt tgggtgatca ctttctgttg gacagtgttc ggcatctctc aggcctggaa 2640
 aggagcatag tcttgggat ccatccaaag acagctgaac cagctatctt acccaattgt 2700
 ctgactctgc tggctccag ggcacaaaca cactatctta tttctccata ggttggccat 2760
 taggaagac tccaaatcaa aatgctatgt aatgtctat ggttgaggt ctgctgttgg 2820
 tagaaactct tctttttagt tcaacaagta gagatttga cggagctgac acaaaaggtt 2880
 tggagctccc ccattttctg ctctcccttc aggggttctc tcccaaccc tttccagcag 2940
 cggctggctgc ccccatctct gacccctgac tctccagcc agaaagatgg tggtttctta 3000
 aaggaaactt agctgtctct cacaatgcc atctgtgtct tgcattttgg gtaaaagcca 3060
 taataaagaa aaactcagcc tgtggccttt cttctctcca agctggggt tctttttcta 3120
 aqtagcttca tgcagtttgt tgccttttaa aatttgctca gaactgttt ctgcgaagac 3180
 atggtctgtt agagctctac tggcctgagc agaaagcaat gttctctgaa tctctgacat 3240
 tctatcttgc tgtatctatt taggcttgg gagagtcgca agataatca gtcactgtca 3300
 gatataaat tttgtcagga caaagaatc cgttatgatt atttaacct ttaaaattgt 3360
 ggtctccaga gcttctctc agaatggccc agaccaagcc ttaatttga tagtgaat 3420
 taatggctac ttaagaaga aattatagac caagatgaaa tgaactaaa cctgtttgcc 3480
 ctgactctca tggagaagtg atatttagga ccaaatctgt gttctgaagg tgtgtatcag 3540
 cctaaaggtg atccagagtg ggaagagat gctctttaa afcaatgagt tcatgttcta 3600
 aactctcagc ttagtctatc gattgtctc ataacaccc tagttaagt tcatgttatt 3660
 ctgagactgt gattttttt ctatttactt tcatgtttg ctggccactt gttatggact 3720
 gaatttctgt gtcacacct tcaaccccaa atctccgtgt tgaagcccaa accctgcactg 3780
 tggagctggg gctgtaagg aagtaattaa ggttacatga agtcatggtg gggctctgat 3840
 ctgctaaagt tgggttctct ataggagag ccccaagaga gcttgttccc tccctcctgt 3900
 tgcactgaaa caagagggca tggagacaca cagagagatg gcagccacct acaagccag 3960
 aggagagacc tcaaatcaa acctcctgt ctggcagagc tcttgacac tgtcttggac 4020
 ttcagactcc cagactgga gaacacaaat tctgttctt cagcttctca gctctgttg 4080
 ttttcttatt gcagctgga aacacagctg tctgttatt tctcaaacag aaaaactctg 4140
 tacttaacaa tgcataatga attatttatt tgcctttcag tctctacaaa acgttctaaa 4200
 acactaatct aatatataac ngtaaatat ttcctaatct aatggaandc aagaaatcat 4260
 atgaccaata ttcacttat tggtaattct acttactga tttccncca gactgtgatt 4320
 ttgggaactt cctgtcctt ctccctgtt tctgttctta ttcactggaat tccagttatc 4380
 tggggtctga aatgcaggc tctcctaatc taagcaaat ctgacagatc agcaaaatga 4440
 gcaaatgtt tctttttct tctgactgc attaatcag atcaactca gcaataaaa 4500
 gctatctttg taagcttgt tactaataaa ttagctctat aagatcctg gactttggag 4560
 ttgttgcaat gcttttqaga gtaattcttt aaagctctaa tttcagctg tcttatctct 4620
 taatgtatta tgcacccact acaacatlt gaacaaatat aataatttaa aatgtataa 4680
 taattatgaa ttttgttta gaacaaagag gattactgat atttgttcc ctatgaatgg 4740
 caaaggttt agcttactac tgcatttctg ttttaataaa aaagtqaga gtttgtgtct 4800
 catataactg gaacaaaaa 4820

<210> 88
 <211> 3599
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7690129CB1

<400> 88

WO 01/98353

PCT/US01/19862

```

ggggcccccgt cctccagacc tggctgcagg acctgctgag tggctgggtg gtcggggctg 50
cccagagacc agggagccctgg abtgcacatg gggatctgca cccgggcatc ggcgggcatg 123
ttgggtgtgg cgtacagggg ccatcagatg gccagacatg ggggacccaa ggtgylggcc 180
atgggtgtgg cccctggggg tgggttccgg aalagagac cctcatca ccccaaggcc 240
tcgttccctg cgggttcccg gtgggcgggt gacggcgagg accgggtcaa gtttccctcg 300
ggatccacat actcgacatt ctctctgggt gacgacggca cacacggctg cctggggggc 360
gggacccgct bccgctggc cctggagtc tachtctac accagagagc gguccgggga 420
gggactggaa ttgacatccc tctctgctc ctctgattg atgggtatga gaagatgtg 480
accgaatag agagccgac cagagccag ctccatgct tctctgtgg tggctcagg 540
ggagctcggg actgclggc ggaagccctg gaagacatc tggcccccgg gaggctggga 600
gccagggcag ggaagccgg agatcgatc aggcgtttc tcccaagg ggaacttgag 660
tcttctcagg cccaggtgga gaggattatg acccggaagg agctcctgac agtclattct 720
tctgagatg ggtctgagga attcagacc atagttttga aggccttgt gaagccctgt 780
gggagctcgg aggcctcagg ctacctggat gagctcgtt tggctgtgg tgggacccg 840
glggaatctg cccagagtg aactcttcgg ggggcatcc aatggcggt cttccatctc 900
gaagcttccc tcatggagc cctgctgaat gacggccgt agtngtgag cttctcalt 960
tcccagggcc ttagcctggg cacttctctg acccgatgc gctggccca actctacagc 1020
ggggcccccct ccaactcgct catcggcaac cttttggacc aggcgtccca cagcgagggc 1080
accgaagccc cggccctaaa agggggagct gcgggagctcc ggcctccga cgtggggcat 1140
tgcttgaggg tggctgtggg gaagatgtgg ggcgcgaggt accctccgg ggggcttgg 1200
gaactcaac caggccagg ctctggggag agcatatct tgcctcgga caaggccacc 1260
tgccgctct cgtggatg tggcctggg caggccccc ggaagacat gcttctttgg 1320
gcctgttggc tgaacaggg accagtgccc atgtactctt gggagatggg ttcacatgca 1380
gtttctcagg ctcttggggc ctgtttgctg ctccgggtga tggcacgctt gggagctgac 1440
cgtgagggg cagcagggag gaagacccctg gcgllcaagt ttgaggggat gggcgttggc 1500
cttcttggag agtctatctg cagcagtgag gtagggggctg ccgcctccl cctcctgct 1560
tgcccgctc ggggggagtc caettgctc cagctggcca tgcagagctga ccgcctgct 1620
tcttctgccc aggaaggggt acagltctctg ctgacacaga agtgggtggg agatattggc 1680
agcctacac ccatctgggg cctgggtctc gcccttcttt gccctccat catctacac 1740
cgctcatca cctttaggaa atcagagagag gggccacac gggagggagt agagtttgac 1800
atggatagt tcatuagtg ggaagggcct gtcgggacgg cggaccacgc cagaagagag 1860
cgctggggg tccggcgcca gtccgggcgt cggggttct cggggggcgg clcaggggg 1920
cgccgtggc taagccgctg gtctcccttc tggggggcgg cggtagaat clcaggggg 1980
aagctggta gctacatgct gtctctgctg ctttctcgc ggttgcctgt cgtggatttc 2040
cagccgggc cggccgggct cctggagctg ctgctctat lctgggcttt caecgcctgt 2100
cgcgagggg tggccaggg cctgagcggg ggcggggcca gctccccag cggggggccc 2160
gggcctggcc atgctcctt ggcacagcgc ctgctcctt accctcgcca cagctggac 2220
cgttcggac tagtgctct cactctctc ctctggggc tgggctcgg gctgacccg 2280
ggtttgtacc accctgggct cctgtctct tgcctcgaet tcatggtttt caggttgccg 2340
ctgcttcaac lcttcaaggt caacaaacag ctggggccca agatcgtcat cgtgagcag 2400
atgatgaagg agtgtttct tctctcttc tctctcggg tgtggctggt agcctatgg 2460
gtggccacgg aggggtctct gaggccacgg gacatgact tcccaagtat cctgcyccgc 2520
gtcttctacc gtccctact gtagatcttc gggcagatc cccagggaga atgggacgtg 2580
gccctcatgg agcacagaa ctgctcgtg gggcccgct tctgggcaca cctcctcgg 2640
ggccagcgg gcaactgct ctccagtat gccactggg tgggtgtgct gctcctcgg 2700
atctctctg tggggccaa cttctgctg gccaacttg tcatggcat gttcagttac 2760
accctggga aagtaaggg caacagagat ctclactgga aggcgcagcg ttacggctc 2820
atccgggaat tccactctg gccagcctg gcccgagcc cccagcgct ctcccccgg 2880
cpgctcctg tccggcaatt gtgcagggc ccttctctaa ggaagcggc ggaagctgt aacgtggaa 3000
tcggtctcata aggaagactt tctgctggca cgcctaggg accagtgga ggcgactcc 3060
ggcgctctga agcgacgct ccagaggggt gacttggac tgaacagct gggacacatc 3120
cgcgagtac aacagggctt gaagtgctg ggcggggagg tccagagct tagcccgctc 3180
ctggagtggt tggccgggg cctgagcgc tctgcttgc tggcccgag tggggcgcca 3240
accctgacc tggctgggt ccaagactga gccctgctg cggacttcaa ggaagagnc 3300
cccagggga ttttctctt agagttaagg tcatctggg ctggggcccc gcaactgggt 3360
gcttctctt tgggtgagc ccatgttca tclggggcc tctcaggacc acccttggga 3420
gttcatctt tacaacccac agcatggcg gctctccca gaacagtc cagcttggga 3480
ggatcaggg cgggacccg ggcggtctc catctggagg ctgcagggtc cttggggtaa 3540
cagggaccaa agaccctca cactctcag attctcaca ctggggaaat aaacccatc 3599

```

【国際公開パンフレット（コレクトバージョン）】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
27 December 2001 (27.12.2001)

PCT

(10) International Publication Number
WO 01/098353 A3(51) International Patent Classification: C12N 15/12,
C07K 14/47, 16/18, A51K 38/17

(21) International Application Number: PCT/US01/09865

(22) International Filing Date: 20 June 2000 (20.06.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/212,890 20 June 2000 (20.06.2000) US
60/212,846 23 June 2000 (23.06.2000) US
60/212,861 27 June 2000 (27.06.2000) US
60/222,372 31 July 2000 (31.07.2000) US
60/221,432 8 September 2000 (08.09.2000) US
60/232,889 15 September 2000 (15.09.2000) US(71) Applicant (for all designated States except US): INCYTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(72) Inventors: and

(75) Inventors/Applicants (for US only): HILLMAN, Jen-
nifer, L. [US/US]; 230 Menlo Drive #17, Mountain
View, CA 94040 (US); TANG, Y., Tom [US/US]; 1730
Ranwick Court, San Jose, CA 95118 (US); YUE, Henry
[US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US);
ELLIOTT, Vicki, N. [US/US]; 3770 Polton Place Way,
San Jose, CA 95121 (US); TRIBOLET, Catherine,
M. [US/US]; 1127 Terrasone Street #5, San Francisco,
CA 94107 (US); LEE, Ernestine, A. [US/US]; 624
Kains Street, Albany, CA 94706 (US); RAMKUMAR,
Jayalaxmi [US/US]; 34359 Maybird Circle, Fremont,
CA 94555 (US); LAL, Preeti [US/US]; PO Box 5142,
Santa Clara, CA 95056 (US); XI, Yuming [US/US];
1789 Walnut Drive, Mountain View, CA 94040 (US);
WARREN, Bridget, A. [US/US]; 10130 Parwood Drive
#2, Cupertino, CA 95014 (US); HAFALLA, April, J., A.
[US/US]; 2227 Calle de Primavera, Santa Clara, CA 95064
(US); BAUGHN, Mariah, R. [US/US]; 14244 Santiago
Road, San Leandro, CA 94577 (US); AZIMZAI, Yalda[US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA
94552 (US); BATRA, Sajeev [US/US]; 555 E. Camino
Real #709, Sunnyvale, CA 94087 (US); BURFORD,
Neil [GB/US]; 105 Wildwood Circle, Durham, CT 06422
(US); VAO, Monique, G. [US/US]; 111 Frederick Court,
Mountain View, CA 94035 (US); NGUYEN, Daniel, B.
[US/US]; 1403 Ridgewood Drive, San Jose, CA 95118
(US); LI, Hyung, Alan, M. [US/US]; 233 Cey Drive, San
Jose, CA 95123 (US); WALIA, Narinder, K. [US/US];
800 Davis Street #205, San Leandro, CA 94577 (US);
GANDHI, Ameena, R. [US/US]; 837 Ruble Avenue
#1, Menlo Park, CA 94025 (US); AU-YOUNG, Janice
[US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005
(US); PATTERSON, Chandra [US/US]; 590 Sherwood
Way #1, Menlo Park, CA 94025 (US).(74) Agents: HAMLET-CON, Diana et al., Lucyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AU, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GR, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RO, RU, SD, SG, SI, SK, SL,
TH, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PL, SE, TR), OAPI patent (BF, BJ, CI,
CG, CF, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

with international search report

(88) Date of publication of the international search report:
10 October 2002For two-letter codes and other abbreviations, refer to the "Guide-
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 01/098353 A3

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/19862
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N13/12 C07K14/47 C07K16/18 A61K38/17		
According to International Patent Classification (IPC) or to both National classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used)		
BIOSIS, SEQUENCE SEARCH		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Quotation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 01 49728 A (KATO SEISHI ; KIMURA TOMOKO (JP); PROTEGENE INC (JP); SAGAMI CHEM R) 12 July 2001 (2001-07-12) page 558- page 563 page 155 -page 156; claims 1,4	1-4, 6-19,22, 25-45,89
P,X	WO 01 49466 A (STEWART TIMOTHY A ; BAKER KEVIN P (US); DEFORGE LAURA (US); DESNOYE) 7 June 2001 (2001-06-07) * see also AU 2474700 with publication date of 19.06.00 * claims 3,12; figures 195,196	1-4, 6-19,22, 25-45,89
P,X	EP 1 074 617 A (HELIX RES INST) 7 February 2001 (2001-02-07) * SEQ ID NO: 17164 and 17165 * claim 8	1-4, 6-19,22, 25-45,89
☐ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.		
* Special categories of cited documents * "A" document defining the general state of the art which is not considered to be of particular relevance "F" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claims or which is cited to establish the publication date of a prior art document or other special reason (as specified) "O" document relating to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principles or theory underlying the invention "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
18 April 2002		16.07.02
Name and mailing address of the ISA European Patent Office, P.B. 6918 Patentstrasse 2 AL - 22801 Reykjavik Tel. (+354-70) 340-2010, Ex. 31: 651 up nt, Fax: (+354-70) 340-2016		Authorized officer: Hillenbrand, G

Form PCT/ISA/210 (second sheet) July 1997

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 01/19862
Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)		
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:		
1. ☒	Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: see FURTHER INFORMATION sheet PCT/ISA/210	
2. ☒	Claims Nos.: 20-21, 23-24 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: see FURTHER INFORMATION sheet PCT/ISA/210	
3. ☐	Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).	
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)		
This International Searching Authority found multiple inventions in this International application, as follows:		
see additional 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.: 1-19, 22, 25-44, (all partially), 45, 89	
Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest. ☐ No protest accompanied the payment of additional search fees.		

International Application No. PCT/US 01/19862

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-19, 22, 25-44 (all partially), 45, 89

Invention 1

The subject-matter of this group of claims is directed to a polypeptide of SEQ ID NO: 1 encoded by the DNA of SEQ ID NO:45, a recombinant polynucleotide comprising said polynucleotide, a cell (transgenic organism) transformed with said polynucleotide, a method for producing said polypeptide, an antibody which binds to said polypeptide, a method for detecting a target polynucleotide in a sample by using said polynucleotide, a composition comprising said polypeptide, use of said composition for treating a disease, screening methods for agonists or antagonists (or compounds which bind) by using said polypeptide, the agonists or antagonists obtained by these methods, screening methods for compounds that modulate the activity of said polypeptide and screening methods for a compound for effectiveness in altering expression of a target polynucleotide, methods for assessing toxicity of a test compound by using said polynucleotide, and diagnostic tests.

Inventions 2-44

The subject-matter of the residual parts of the claims mentioned above (SEQ ID NOs: 2-44 and 46-88) and claims 46-88 (polypeptides) and claims 90-132 is directed to further inventions 2-44 directed to further polypeptides, their corresponding DNA sequences and their uses as described above.

International Application No. PCT/US 01/19862

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claim 18 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Although claims 32 and 34 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.2

Claims Nos.: 20-21, 23-24

Present claims 20-21 and 23-24, which are directed to all possible agonists or antagonists, relate to such a large number of possibly known compounds that a meaningful search was impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT				International Application No. PCT/US 01/19862	
Patent document cited in search report	Publication date	Patent family member(s)	Publication date		
WO 0149728	A	12-07-2001	AU 2404001 A	16-07-2001	
			WO 0149728 A2	12-07-2001	
WO 0140466	A	07-06-2001	AU 2055401 A	12-06-2001	
			AU 2192800 A	12-07-2000	
			AU 2474700 A	19-06-2000	
			AU 2879400 A	31-07-2001	
			AU 3514400 A	28-09-2000	
			AU 3774300 A	18-12-2000	
			AU 5152700 A	02-01-2001	
			AU 6391000 A	19-02-2001	
			EP 1173563 A1	23-01-2002	
			EP 1220905 A2	10-07-2002	
			EP 1210418 A1	05-06-2002	
			EP 1208195 A2	29-05-2002	
			EP 1141289 A2	10-10-2001	
			EP 1135495 A2	26-09-2001	
			WO 0153486 A1	26-07-2001	
			WO 0053758 A2	14-09-2000	
			WO 0073454 A1	07-12-2000	
			WO 0077037 A2	21-12-2000	
			WO 0109327 A2	08-02-2001	
			WO 0140466 A2	07-06-2001	
			WO 0037640 A2	29-06-2000	
			US 2002058309 A1	16-05-2002	
			US 2002072496 A1	13-06-2002	
			US 2002072067 A1	13-06-2002	
			US 2002072092 A1	13-06-2002	
			US 2002072497 A1	13-06-2002	
			AU 1932000 A	03-07-2000	
			AU 3107700 A	28-09-2000	
			AU 5460100 A	18-12-2000	
			AU 5591100 A	18-12-2000	
			EP 1185648 A2	13-03-2002	
			EP 1141285 A2	10-10-2001	
			EP 1159422 A1	05-12-2001	
			WO 0073348 A2	07-12-2000	
			WO 0073452 A2	07-12-2000	
			WO 0053750 A1	14-09-2000	
			AU 2224800 A	28-09-2000	
			AU 2883600 A	28-09-2000	
			AU 2883700 A	09-01-2001	
			AU 5459900 A	18-12-2000	
			WO 0053754 A1	14-09-2000	
			WO 0053756 A2	14-09-2000	
			WO 0078961 A1	28-12-2000	
			AU 1749800 A	04-10-2000	
			AU 2398700 A	05-02-2001	
			AU 2596700 A	28-09-2000	
			AU 2883900 A	30-01-2001	
			WO 0053753 A2	14-09-2000	
			WO 0104311 A1	18-01-2001	
			WO 0105836 A1	25-01-2001	
EP 1074617	A	07-02-2001	AU 6180800 A	19-02-2001	
			AU 6180900 A	19-02-2001	
			AU 6181000 A	19-02-2001	
			AU 6181100 A	19-02-2001	

Form PCT/ISA/210 (2.0001) (Annex) (July 1997)

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/19862	
Patent document cited in search report:	Publication date	Patent family member(s)	Publication date
EP 1074617 A	A	AU 6181200 A	19-02-2001
		AU 6181300 A	19-02-2001
		AU 6181400 A	19-02-2001
		AU 6181500 A	19-02-2001
		AU 6181600 A	19-02-2001
		AU 6315800 A	19-02-2001
		EP 1074617 A2	07-02-2001
		EP 1205549 A1	15-05-2002
		WO 0109315 A1	08-02-2001
		WO 0109345 A1	08-02-2001
		WO 0109316 A1	08-02-2001
		WO 0109349 A1	08-02-2001
		WO 0109317 A1	08-02-2001
		WO 0109318 A1	08-02-2001
		WO 0109319 A1	08-02-2001
		WO 0109346 A1	08-02-2001
		WO 0109320 A1	08-02-2001
		WO 0109321 A1	08-02-2001
		WO 0109322 A1	08-02-2001
		WO 0109323 A1	08-02-2001

Form PCT/ISA/210 (latest family member) (July 1992)

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 1/18
 A 6 1 P 3/00
 A 6 1 P 3/10
 A 6 1 P 3/12
 A 6 1 P 5/14
 A 6 1 P 5/38
 A 6 1 P 7/00
 A 6 1 P 7/02
 A 6 1 P 7/04
 A 6 1 P 7/06
 A 6 1 P 9/00
 A 6 1 P 9/04
 A 6 1 P 9/08
 A 6 1 P 9/10
 A 6 1 P 9/12
 A 6 1 P 11/00
 A 6 1 P 11/06
 A 6 1 P 13/00
 A 6 1 P 13/12
 A 6 1 P 17/00
 A 6 1 P 17/02
 A 6 1 P 17/06
 A 6 1 P 19/02
 A 6 1 P 19/06
 A 6 1 P 19/10
 A 6 1 P 21/00
 A 6 1 P 21/04
 A 6 1 P 25/00
 A 6 1 P 25/02
 A 6 1 P 25/14
 A 6 1 P 25/16
 A 6 1 P 25/18
 A 6 1 P 25/20
 A 6 1 P 25/22
 A 6 1 P 25/24
 A 6 1 P 25/28
 A 6 1 P 27/02
 A 6 1 P 27/06
 A 6 1 P 27/12
 A 6 1 P 27/16
 A 6 1 P 29/00
 A 6 1 P 31/04
 A 6 1 P 31/06
 A 6 1 P 31/10
 A 6 1 P 31/12
 A 6 1 P 31/18
 A 6 1 P 33/00

A 6 1 P 3/00
 A 6 1 P 3/10
 A 6 1 P 3/12
 A 6 1 P 5/14
 A 6 1 P 5/38
 A 6 1 P 7/00
 A 6 1 P 7/02
 A 6 1 P 7/04
 A 6 1 P 7/06
 A 6 1 P 9/00
 A 6 1 P 9/04
 A 6 1 P 9/08
 A 6 1 P 9/10
 A 6 1 P 9/12
 A 6 1 P 11/00
 A 6 1 P 11/06
 A 6 1 P 13/00
 A 6 1 P 13/12
 A 6 1 P 17/00
 A 6 1 P 17/02
 A 6 1 P 17/06
 A 6 1 P 19/02
 A 6 1 P 19/06
 A 6 1 P 19/10
 A 6 1 P 21/00
 A 6 1 P 21/04
 A 6 1 P 25/00
 A 6 1 P 25/02
 A 6 1 P 25/02
 A 6 1 P 25/14
 A 6 1 P 25/16
 A 6 1 P 25/18
 A 6 1 P 25/20
 A 6 1 P 25/22
 A 6 1 P 25/24
 A 6 1 P 25/28
 A 6 1 P 27/02
 A 6 1 P 27/06
 A 6 1 P 27/12
 A 6 1 P 27/16
 A 6 1 P 29/00
 A 6 1 P 31/04
 A 6 1 P 31/06
 A 6 1 P 31/10
 A 6 1 P 31/12
 A 6 1 P 31/18

1 0 3

1 0 3

4 C 0 8 4
 4 H 0 4 5

A 6 1 P	35/00	A 6 1 P	33/00	
A 6 1 P	35/02	A 6 1 P	35/00	
A 6 1 P	37/00	A 6 1 P	35/02	
A 6 1 P	37/02	A 6 1 P	37/00	
A 6 1 P	37/08	A 6 1 P	37/02	
C 0 7 K	1/22	A 6 1 P	37/08	
C 0 7 K	14/47	C 0 7 K	1/22	
C 0 7 K	16/18	C 0 7 K	14/47	
C 0 7 K	16/46	C 0 7 K	16/18	
C 0 7 K	19/00	C 0 7 K	16/46	
C 1 2 N	1/15	C 0 7 K	19/00	
C 1 2 N	1/19	C 1 2 N	1/15	
C 1 2 N	1/21	C 1 2 N	1/19	
C 1 2 N	5/10	C 1 2 N	1/21	
C 1 2 P	21/02	C 1 2 P	21/02	C
C 1 2 P	21/08	C 1 2 P	21/08	
C 1 2 Q	1/68	C 1 2 Q	1/68	A
G 0 1 N	33/15	G 0 1 N	33/15	Z
G 0 1 N	33/50	G 0 1 N	33/50	Z
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	
		C 1 2 N	5/00	A
		A 6 1 K	37/02	

- (31)優先権主張番号 60/222,372
 (32)優先日 平成12年7月31日(2000.7.31)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/231,435
 (32)優先日 平成12年9月8日(2000.9.8)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/232,889
 (32)優先日 平成12年9月15日(2000.9.15)
 (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,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

- (72)発明者 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 ユエ、ヘンリー
 アメリカ合衆国カリフォルニア州 9 4 0 8 7 ・サニーベイル・ルイスアベニュー 8 2 6
 (72)発明者 エリオット、ビッキー・エス
 アメリカ合衆国カリフォルニア州 9 5 1 2 1 ・サンノゼ・ポルトンプレイスウェイ 3 7 7 0
 (72)発明者 トリボレー、キャサリン・エム
 アメリカ合衆国カリフォルニア州 9 4 1 0 7 ・サンフランシスコ・# 5 ・テネシーストリート 1 2 1
 (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 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 ラル、ブリーティ
- アメリカ合衆国カリフォルニア州 9 5 0 5 6 ・ サンタクララ ・ ピーオーボックス 5 1 4 2
- (72)発明者 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 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 5 0 1 4 ・ クーペルティノー ・ # 2 ・ パーウッドドライブ 1 0 1 3 0
- (72)発明者 ハファリア、エープリル・ジェイ・エイ
- アメリカ合衆国カリフォルニア州 9 5 0 5 4 ・ サンタクララ ・ コーレデプリマベラ 2 2 2 7
- (72)発明者 ボーゲン、マライア・アール
- アメリカ合衆国カリフォルニア州 9 4 5 7 7 ・ サンレアンドロ ・ サンティアゴロード 1 4 2 4 4
- (72)発明者 アジムザイ、ヤルダ
- アメリカ合衆国カリフォルニア州 9 4 5 5 2 ・ カストロバレー ・ ボールダーキャニオンドライブ 5 5 1 8
- (72)発明者 バトラ、サジューブ
- 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 ・ # 7 0 9 ・ エルカミーノリアル 5 5
- (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 0 6 4 2 2 ・ ダラム ・ ワイルドウッドサークル 1 0 5
- (72)発明者 ヤオ、モニーク・ジー
- 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 ・ フレデリックコート 1 1 1
- (72)発明者 ニュエン、ダニエル・ビー
- アメリカ合衆国カリフォルニア州 9 5 1 1 8 ・ サンノゼ ・ リッジウッドドライブ 1 4 0 3
- (72)発明者 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 ワライヤ、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 7 7 ・ サンレアンドロ ・ # 2 0 5 ・ デービスストリート 8 9 0
- (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 9 4 0 2 5 ・ メンロパーク ・ # 1 ・ ローブルアベニュー 8 3 7
- (72)発明者 オウーヤング、ジャニス
- アメリカ合衆国カリフォルニア州 9 4 0 0 5 ・ ブリスベーン ・ ゴールデンイーグルレーン 2 3 3
- (72)発明者 パターソン、チャンドラ
- アメリカ合衆国カリフォルニア州 9 4 0 2 5 ・ メンロパーク ・ # 1 ・ シャーウッドウェイ 4 9 0

F ターム(参考) 2G045 AA34 AA35 BB07 BB10 BB14 BB46 BB48 CB01 DA13 DA36

FB02 FB03 FB05 FB12 GC15

4B024 AA01 AA11 BA31 BA43 CA04 CA09 FA02 GA05 HA01 HA14
HA17

4B063 QA01 QA18 QA19 QQ01 QQ43 QR55 QS25 QS34

4B064 AG01 AG27 AG31 CA19 CA20 CC24 DA01 DA13

4B065 AB01 BA02 CA24 CA44 CA46

4C084 AA17 BA01 BA08 BA20 BA21 BA22 CA53 CA59 DA39 DC22

DC25 MA13 MA52 MA56 MA59 MA60 MA63 MA66 ZA022 ZA052

ZA122 ZA162 ZA182 ZA202 ZA242 ZA332 ZA342 ZA362 ZA392 ZA422

ZA452 ZA512 ZA552 ZA592 ZA662 ZA752 ZA812 ZA942 ZA962 ZA972

ZB052 ZB132 ZB262 ZB272 ZB322 ZB332 ZB352 ZB382 ZB392 ZC062

ZC082 ZC212 ZC352

4H045 AA10 AA11 AA20 AA30 BA10 BA41 CA40 DA76 DA86 EA21
EA22 EA23 EA24 EA26 EA27 EA28 GA26

专利名称(译)	分泌蛋白質		
公开(公告)号	JP2004500870A	公开(公告)日	2004-01-15
申请号	JP2002504308	申请日	2001-06-20
[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	ヒルマンジェニファーエル タングワイトム ユエヘンリー エリオットビッキーエス トリボレーキャサリーンエム リーアーンステイーンエイ ランクマールジャヤラクシミ ラルプリーティ スーユーミング ワレンブリジットエイ ハファリアエープリルジェイエイ ボーグンマライアアール アジムザイヤルダ バトラサジープ バーフォードニール ヤオモニークジー ニュエンダニエルビー リュデュングアイナエム フライヤナリンダーケイ ガンディーアミーナアール オウヤングジャンス パターソンチャンドラ		
发明人	ヒルマン、ジェニファー・エル タング、ワイトム ユエ、ヘンリー エリオット、ビッキー・エス トリボレー、キャサリーン・エム リー、アーンステイーン・エイ ランクマール、ジャヤラクシミ ラル、プリーティ スー、ユーミング ワレン、ブリジット、エイ ハファリア、エープリル・ジェイ・エイ ボーグン、マライア・アール アジムザイ、ヤルダ バトラ、サジープ バーフォード、ニール ヤオ、モニーク・ジー ニュエン、ダニエル・ビー リュ、デュング・アイナ・エム フライヤ、ナリンダー・ケイ ガンディー、アミーナ・アール オウ・ヤング、ジャンス パターソン、チャンドラ		

IPC分类号	G01N33/50 A61K38/00 A61K45/00 A61P1/04 A61P1/16 A61P1/18 A61P3/00 A61P3/10 A61P3/12 A61P5/14 A61P5/38 A61P7/00 A61P7/02 A61P7/04 A61P7/06 A61P9/00 A61P9/04 A61P9/08 A61P9/10 A61P9/12 A61P11/00 A61P11/06 A61P13/00 A61P13/12 A61P17/00 A61P17/02 A61P17/06 A61P19/02 A61P19/06 A61P19/10 A61P21/00 A61P21/04 A61P25/00 A61P25/02 A61P25/14 A61P25/16 A61P25/18 A61P25/20 A61P25/22 A61P25/24 A61P25/28 A61P27/02 A61P27/06 A61P27/12 A61P27/16 A61P29/00 A61P31/04 A61P31/06 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37/00 A61P37/02 A61P37/08 C07K1/22 C07K14/47 C07K16/18 C07K16/46 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N5/09 C12N5/12 C12P21/02 C12P21/08 C12Q1/68 G01N33/15 G01N33/53 G01N33/566
CPC分类号	A61P1/04 A61P1/16 A61P1/18 A61P3/00 A61P3/12 A61P11/00 A61P11/06 A61P13/00 A61P13/12 A61P17/00 A61P17/02 A61P17/06 A61P19/02 A61P19/06 A61P19/10 A61P21/00 A61P21/04 A61P25/00 A61P25/02 A61P25/14 A61P25/16 A61P25/18 A61P25/20 A61P25/22 A61P25/24 A61P25/28 A61P27/02 A61P27/06 A61P27/12 A61P27/16 A61P29/00 A61P31/04 A61P31/06 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 C07K14/47
FI分类号	C12N15/00.ZNA.A A61K45/00 A61P1/04 A61P1/16 A61P1/18 A61P3/00 A61P3/10 A61P3/12 A61P5/14 A61P5/38 A61P7/00 A61P7/02 A61P7/04 A61P7/06 A61P9/00 A61P9/04 A61P9/08 A61P9/10 A61P9/10.103 A61P9/12 A61P11/00 A61P11/06 A61P13/00 A61P13/12 A61P17/00 A61P17/02 A61P17/06 A61P19/02 A61P19/06 A61P19/10 A61P21/00 A61P21/04 A61P25/00 A61P25/02 A61P25/02.103 A61P25/14 A61P25/16 A61P25/18 A61P25/20 A61P25/22 A61P25/24 A61P25/28 A61P27/02 A61P27/06 A61P27/12 A61P27/16 A61P29/00 A61P31/04 A61P31/06 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37/00 A61P37/02 A61P37/08 C07K1/22 C07K14/47 C07K16/18 C07K16/46 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.M G01N33/566 C12N5/00.A A61K37/02
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/BB07 2G045/BB10 2G045/BB14 2G045/BB46 2G045/BB48 2G045/CB01 2G045/DA13 2G045/DA36 2G045/FB02 2G045/FB03 2G045/FB05 2G045/FB12 2G045/GC15 4B024/AA01 4B024/AA11 4B024/BA31 4B024/BA43 4B024/CA04 4B024/CA09 4B024/FA02 4B024/GA05 4B024/HA01 4B024/HA14 4B024/HA17 4B063/QA01 4B063/QA18 4B063/QA19 4B063/QQ01 4B063/QQ43 4B063/QR55 4B063/QS25 4B063/QS34 4B064/AG01 4B064/AG27 4B064/AG31 4B064/CA19 4B064/CA20 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AB01 4B065/BA02 4B065/CA24 4B065/CA44 4B065/CA46 4C084/AA17 4C084/BA01 4C084/BA08 4C084/BA20 4C084/BA21 4C084/BA22 4C084/CA53 4C084/CA59 4C084/DA39 4C084/DC22 4C084/DC25 4C084/MA13 4C084/MA52 4C084/MA56 4C084/MA59 4C084/MA60 4C084/MA63 4C084/MA66 4C084/ZA022 4C084/ZA052 4C084/ZA122 4C084/ZA162 4C084/ZA182 4C084/ZA202 4C084/ZA242 4C084/ZA332 4C084/ZA342 4C084/ZA362 4C084/ZA392 4C084/ZA422 4C084/ZA452 4C084/ZA512 4C084/ZA552 4C084/ZA592 4C084/ZA662 4C084/ZA752 4C084/ZA812 4C084/ZA942 4C084/ZA962 4C084/ZA972 4C084/ZB052 4C084/ZB132 4C084/ZB262 4C084/ZB272 4C084/ZB322 4C084/ZB332 4C084/ZB352 4C084/ZB382 4C084/ZB392 4C084/ZC062 4C084/ZC082 4C084/ZC212 4C084/ZC352 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/BA41 4H045/CA40 4H045/DA76 4H045/DA86 4H045/EA21 4H045/EA22 4H045/EA23 4H045/EA24 4H045/EA26 4H045/EA27 4H045/EA28 4H045/GA26
优先权	60/212890 2000-06-20 US 60/213466 2000-06-23 US 60/214601 2000-06-27 US 60/222372 2000-07-31 US 60/231435 2000-09-08 US 60/232889 2000-09-15 US
外部链接	Espacenet

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。

[illegible]