

(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2004-518402

(P2004-518402A)

(43) 公表日 平成16年6月24日(2004.6.24)

|                           |                              |                   |                                             |                                                                                               |  |
|---------------------------|------------------------------|-------------------|---------------------------------------------|-----------------------------------------------------------------------------------------------|--|
| (51) Int.Cl. <sup>7</sup> |                              | F I               |                                             | テーマコード (参考)                                                                                   |  |
| C 1 2 N 15/09             |                              | C 1 2 N 15/00     |                                             | Z N A A 2 G O 4 5                                                                             |  |
| A 6 1 K 38/00             |                              | A 6 1 K 39/395    |                                             | D 4 B O 2 4                                                                                   |  |
| A 6 1 K 39/395            |                              | A 6 1 K 39/395    |                                             | N 4 B O 6 3                                                                                   |  |
| A 6 1 K 45/00             |                              | A 6 1 K 45/00     |                                             | 4 B O 6 4                                                                                     |  |
| A 6 1 K 49/00             |                              | A 6 1 K 49/00     |                                             | Z 4 B O 6 5                                                                                   |  |
|                           |                              | 審査請求 未請求 予備審査請求 有 |                                             | (全 218 頁) 最終頁に続く                                                                              |  |
| (21) 出願番号                 | 特願2001-576886 (P2001-576886) | (71) 出願人          | 301005050                                   | インサイト・ゲノミックス・インコーポレイテッド<br>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<br>100089266<br>弁理士 大島 陽一 |  |
| (86) (22) 出願日             | 平成13年4月11日 (2001.4.11)       | (74) 代理人          | 100089266                                   |                                                                                               |  |
| (85) 翻訳文提出日               | 平成14年10月11日 (2002.10.11)     | (72) 発明者          | グリフィン、ジェニファー・エイ                             |                                                                                               |  |
| (86) 国際出願番号               | PCT/US2001/011861            |                   | アメリカ合衆国カリフォルニア州94002・ベルモント・#11・アイリーンコート 181 |                                                                                               |  |
| (87) 国際公開番号               | W02001/079291                |                   |                                             | 最終頁に続く                                                                                        |  |
| (87) 国際公開日                | 平成13年10月25日 (2001.10.25)     |                   |                                             |                                                                                               |  |
| (31) 優先権主張番号              | 60/197,854                   |                   |                                             |                                                                                               |  |
| (32) 優先日                  | 平成12年4月14日 (2000.4.14)       |                   |                                             |                                                                                               |  |
| (33) 優先権主張国               | 米国 (US)                      |                   |                                             |                                                                                               |  |
| (31) 優先権主張番号              | 60/202,373                   |                   |                                             |                                                                                               |  |
| (32) 優先日                  | 平成12年5月4日 (2000.5.4)         |                   |                                             |                                                                                               |  |
| (33) 優先権主張国               | 米国 (US)                      |                   |                                             |                                                                                               |  |
| (31) 優先権主張番号              | 60/205,899                   |                   |                                             |                                                                                               |  |
| (32) 優先日                  | 平成12年5月18日 (2000.5.18)       |                   |                                             |                                                                                               |  |
| (33) 優先権主張国               | 米国 (US)                      |                   |                                             |                                                                                               |  |

(54) 【発明の名称】 分泌タンパク質

## (57) 【要約】

本発明は、分泌タンパク質 (S E C P) と、S E C Pを同定及びコードするポリヌクレオチドとを提供する。本発明はまた、発現ベクター及び宿主細胞、抗体、アゴニスト、アンタゴニストを提供する。更に、本発明は、S E C Pの発現に関連する疾患の診断・治療・予防方法を提供する。

## 【特許請求の範囲】

## 【請求項 1】

単離されたポリペプチドであって、

(a) SEQ ID NO: 1 乃至 SEQ ID NO: 14 (SEQ ID NO: 1-14) からなる群から選択されたアミノ酸を含むポリペプチドと、

(b) SEQ ID NO: 1-14 からなる群から選択されたアミノ酸配列と少なくとも 90% の同一性を有するアミノ酸配列を含む天然のポリペプチドと、

(c) SEQ ID NO: 1-14 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、

(d) SEQ ID NO: 1-14 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されることを特徴とする単離されたポリペプチド。 10

## 【請求項 2】

SEQ ID NO: 1-14 からなる群から選択された請求項 1 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: 15-28 からなる群から選択された請求項 4 の単離されたポリヌクレオチド。 20

## 【請求項 6】

請求項 3 の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。

## 【請求項 7】

請求項 6 の組換えポリヌクレオチドで形質転換された細胞。

## 【請求項 8】

請求項 6 の組換えポリヌクレオチドを含む遺伝子組換え生物。

## 【請求項 9】

請求項 1 のポリペプチドを産生する方法であって、 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過程と、

(b) そのように発現したポリペプチドを回収する過程とを含むことを特徴とする方法。

## 【請求項 10】

請求項 1 のポリペプチドに特異的に結合する単離された抗体。

## 【請求項 11】

単離されたポリヌクレオチドであって、

(a) SEQ ID NO: 15-28 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、 40

(b) SEQ ID NO: 15-28 からなる群から選択されたポリヌクレオチド配列と少なくとも 90% の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、

(c) 前記 (a) のポリヌクレオチドに相補的なポリヌクレオチドと、

(d) 前記 (b) のポリヌクレオチドに相補的なポリヌクレオチドと、

(e) 前記 (a) 乃至 (d) の RNA 等価物とで構成される群から選択された単離されたポリヌクレオチド。

## 【請求項 12】

請求項 11 のポリヌクレオチドの少なくとも 60 個の連続するヌクレオチドを含む単離さ 50

れたポリヌクレオチド。

【請求項 13】

サンプル中において請求項 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検出する方法であって、

(a) 前記サンプルにおいて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 個の連続するヌクレオチドを含むようなプローブであって、該プローブと前記標的ポリヌクレオチドまたはその断片との間でハイブリダイゼーション複合体が形成される条件下で、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ような前記プローブにより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が存在するか否かを検出し、存在する場合には必要に応じてその収量を測定する過程とを含む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 個の連続する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】

サンプル中において請求項 11 のポリヌクレオチド配列を有する標的ポリヌクレオチドを検出する方法であって、

(a) ポリメラーゼ連鎖反応増幅法を用いて、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) 増幅された前記標的ポリヌクレオチドまたはその断片が存在するか否かを検出し、存在する場合には必要に応じてその収量を測定する過程とを含むことを特徴とする方法。 20

【請求項 16】

請求項 1 のポリペプチドと医薬的に容認できる賦形剤とを含む組成物。

【請求項 17】

前記ポリペプチドが、SEQ ID NO: 1-14 からなる群から選択されたアミノ酸配列を含むことを特徴とする請求項 16 の組成物。

【請求項 18】

機能的な SECP (新規の分泌タンパク質) 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過程を含むことを特徴とする方法。 30

【請求項 19】

請求項 1 のポリペプチドのアゴニストとして有効な化合物をスクリーニングする方法であって、

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝露する過程と、

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

【請求項 20】

請求項 19 の方法によって同定されたアゴニスト化合物と医薬的に容認できる賦形剤とを含む組成物。

【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過程を含むことを特徴とする方法。 40

【請求項 22】

請求項 1 のポリペプチドのアンタゴニストとして有効な化合物をスクリーニングする方法であって、

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝露する過程と、

(b) 前記サンプルにおけるアンタゴニスト活性を検出する過程とを含むことを特徴とする方法。

【請求項 23】

請求項 22 の方法によって同定されたアンタゴニスト化合物と医薬的に容認できる賦形剤 50

とを含む組成物。

【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方法。

【請求項 2 5】

請求項 1 のポリペプチドに特異的に結合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 のポリペプチドを好適な条件下で少なくとも 1 つの試験化合物と結合させる過程と、

(b) 請求項 1 のポリペプチドと前記試験化合物との結合を検出して、請求項 1 のポリペプチドと特異的に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スクリーニング方法。 10

【請求項 2 6】

請求項 1 のポリペプチドの活性を変化させ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 のポリペプチドを、そのポリペプチドの活性が許容される条件下で少なくとも 1 つの試験化合物と結合させる過程と、

(b) 前記試験化合物の存在下での請求項 1 のポリペプチドの活性を評価する過程と、  
(c) 前記試験化合物の存在下での請求項 1 のポリペプチドの活性と、前記試験化合物の非存在下での請求項 1 のポリペプチドの活性とを比較する過程とを含み、

前記試験化合物の存在下での請求項 1 のポリペプチドの活性の変化が、請求項 1 のポリペプチドの活性を変化させる化合物の存在を示唆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有効な化合物をスクリーニングする方法であって、

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

(b) 前記標的ポリヌクレオチドの発現の変化を検出する過程と、

(c) 様々な量の前記化合物の存在下での前記標的ポリヌクレオチドの発現と、前記化合物の非存在下での前記標的ポリヌクレオチドの発現とを比較する過程とを含むことを特徴とする方法。

【請求項 2 8】

試験化合物の毒性を評価する方法であって、

(a) 核酸を含む生体サンプルを前記試験化合物で処理する過程と、

(b) 請求項 1 1 のポリヌクレオチドの少なくとも 2 0 個の連続するヌクレオチドを含むプローブと、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ような条件下において、前記処理した生体サンプルの核酸を前記プローブにハイブリダイズさせる過程と、

(c) ハイブリダイゼーション複合体の収量を定量する過程と、

(d) 前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処理の生体サンプルにおけるハイブリダイゼーション複合体の収量と比較する過程とを含み 40

、  
前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量の差が試験化合物の毒性を示唆することを特徴とする方法。

【請求項 2 9】

生体サンプルにおける S E C P の発現に関連する病態または疾患の診断テストであって、

(a) 請求項 1 0 の抗体と前記ポリペプチドとの結合に好適な条件下で、前記生体サンプルを請求項 1 0 の抗体と結合させて、抗体／ポリペプチド複合体を形成する過程と、

(b) 該複合体を検出する過程とを含み、

該複合体の存在が、前記生体サンプルにおける前記ポリヌクレオチドの存在と相関性を有することを特徴とする診断テスト。

## 【請求項 30】

前記抗体が、

- (a) キメラ抗体、
- (b) 一本鎖抗体、
- (c) Fab断片、
- (d) F(ab')<sub>2</sub>断片、または
- (e) ヒト化抗体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の抗体と許容される賦形剤とを含む組成物。

## 【請求項 32】

被検者のSECPの発現に関連する病態または疾患を診断する方法であって、請求項31の組成物を前記被検者に有効量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ていることを特徴とする請求項31に記載の組成物。

## 【請求項 34】

被検者のSECPの発現に関連する病態または疾患を診断する方法であって、請求項33の組成物を前記被検者に有効量投与することを含む過程を特徴とする方法。

## 【請求項 35】

請求項10の抗体の特異性を有するポリクローナル抗体を作製する方法であって、

- (a) 抗体反応が起こる条件下で、SEQ ID NO: 1-14からなる群から選択されたアミノ酸配列またはその免疫原性断片を有するポリペプチドにより動物を免疫化する過程と、
- (b) 前記動物から抗体を単離する過程と、
- (c) 前記単離された抗体を前記ポリペプチドでスクリーニングし、それによって、SEQ ID NO: 1-14からなる群から選択されたアミノ酸配列を有するポリペプチドに特異的に結合するポリクローナル抗体を同定する過程とを含むことを特徴とする方法。

## 【請求項 36】

請求項35の方法によって作製された抗体。

## 【請求項 37】

請求項36の抗体及び適切な担体を含む組成物。

## 【請求項 38】

請求項10の抗体の特異性を有するモノクローナル抗体を作製する方法であって、

- (a) 抗体反応が起こる条件下で、SEQ ID NO: 1-14からなる群から選択されたアミノ酸配列またはその免疫原性断片を有するポリペプチドにより動物を免疫化する過程と、
- (b) 前記動物から抗体を産生する細胞を単離する過程と、
- (c) 前記抗体産生細胞を不死化細胞と融合し、モノクローナル抗体産生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: 1-14からなる群から選択されたアミノ酸配列を有するポリペプチドに特異的に結合するモノクローナル抗体を培地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の方法で作製したモノクローナル抗体。

## 【請求項 40】

請求項39の抗体及び適切な担体を含む組成物。

## 【請求項 41】

Fab発現ライブラリのスクリーニングにより得られた請求項10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得られた請求項10の抗体。

10

20

30

40

50

## 【請求項 43】

サンプルにおける SEQ ID NO: 1-14 からなる群から選択されたアミノ酸配列を有するポリペプチドを検出する方法であって、

(a) 請求項 10 の抗体と前記ポリペプチドとの特異的な結合が許容される条件下で、前記抗体を前記サンプルと共にインキュベートする過程と、

(b) 特異的な結合を検出する過程とを含み、

前記特異的な結合が、前記サンプルに SEQ ID NO: 1-14 からなる群から選択されたアミノ酸配列を有するポリペプチドが存在することを示唆することを特徴とする方法。

## 【請求項 44】

10

SEQ ID NO: 1-14 からなる群から選択されたアミノ酸配列を有するポリペプチドをサンプルから精製する方法であって、

(a) 請求項 10 の抗体と前記ポリペプチドとの特異的な結合が許容される条件下で、前記抗体を前記サンプルと共に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て、SEQ ID NO: 1-14 からなる群から選択されたアミノ酸配列を有する精製されたポリペプチドを得る過程とを含むことを特徴とする方法。

## 【請求項 45】

SEQ ID NO: 1 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 46】

20

SEQ ID NO: 2 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 47】

SEQ ID NO: 3 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 48】

SEQ ID NO: 4 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 49】

SEQ ID NO: 5 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 50】

SEQ ID NO: 6 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 51】

30

SEQ ID NO: 7 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 52】

SEQ ID NO: 8 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 53】

SEQ ID NO: 9 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 54】

SEQ ID NO: 10 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 55】

SEQ ID NO: 11 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 56】

40

SEQ ID NO: 12 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 57】

SEQ ID NO: 13 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 58】

SEQ ID NO: 14 のアミノ酸配列を含む請求項 1 のポリペプチド。

## 【請求項 59】

SEQ ID NO: 15 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

## 【請求項 60】

SEQ ID NO: 16 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

## 【請求項 61】

50

SEQ ID NO: 17 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 62】

SEQ ID NO: 18 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 63】

SEQ ID NO: 19 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 64】

SEQ ID NO: 20 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 65】

SEQ ID NO: 21 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 66】

SEQ ID NO: 22 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 67】

SEQ ID NO: 23 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 68】

SEQ ID NO: 24 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 69】

SEQ ID NO: 25 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 70】

SEQ ID NO: 26 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 71】

SEQ ID NO: 27 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 72】

SEQ ID NO: 28 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【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し、また、これらの配列を利用した細胞増殖異常、自己免疫／炎症性疾患、心血管疾患、神経疾患及び発達障害の診断、治療及び予防に関する。本発明は更に、分泌タンパク質の核酸配列及びアミノ酸配列の発現に及ぼす外来性化合物の影響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、細胞機能に不可欠である。タンパク質輸送は、輸送または分泌されるタンパク質の N 末端に位置するシグナルペプチドによって仲介される。シグナルペプチドは、リボソームから小胞体 (ER) などの特定の膜結合区画まで新生タンパク質をターゲティングするような約 10～20 の疎水性アミノ酸を含む。ER にターゲティングされるタンパク質は、分泌経路を通して移動するか或いは ER やゴルジ体、リソソームなどの任意の分泌オルガネラ内に留まり得る。分泌経路を通して通過するタンパク質は、細胞外空間に分泌されるか或いは原形質膜内に留まる。原形質膜内に留まるタンパク質は、1 若しくは数個の膜貫通ドメインを有し、各々は約 20 の疎水性アミノ酸残基を含む。分泌タンパク質は、通常、分泌経路を通して通過する間に翻訳後処理イベントにより活性化される不活性前駆体として合成される。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ある。タンパク質輸送中に発生し得るその他のイベントには、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、タンパク質と受容体または孔複合体との相互作用がある。N 末端シグナルペプチドを有する分泌タンパク質の例を、細胞間シグナル伝達において重要な役割を有するタンパク質を含めて以下に示す。そのようなタンパク質には、膜貫通受容体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バソメディエータ (vasomediator)

(Reviewed in Alberts, B. ら (1994) Molecular

10

20

30

40

50

Biologv of The Cell, Garland Publishing, New York, NY, pp. 557-560, 582-592)。

#### 【0003】

細胞表面マーカーには、免疫系の白血球細胞で同定された細胞表面抗原がある。これらの抗原は、組織的な「ショットガン (shot gun)」技術に基づくモノクローナル抗体 (mAb) を用いて同定されてきた。このような技術により、結果的に、不明な細胞表面白血球抗原に対して数百のmAbが作られてきた。これらの抗原は、種々の分化または未分化白血球細胞タイプにおける通常の免疫細胞化学的局在化パターンに基づき「分化のクラスター (cluster of differentiation)」にグループ分けされた。所与のクラスターにおける抗原は、細胞表面タンパク質を同定するものと仮定されており、割り当てられた「分化のクラスター」即ち「CD」指定である。CD抗原により同定されたタンパク質をコードする遺伝子には、標準的な分子生物学技術により複製及び検証されてきたものがある。CD抗原は、グリコシルホスファチジルイノシトール (GPI) などの脂肪酸含有糖脂質への共有結合を介して原形質膜に固定された細胞表面タンパク質及び膜貫通タンパク質の両者によって特徴付けられてきた (Reviewed in Barclay, A. N. ら (1995) *The Leucocyte Antigen Facts Book*, Academic Press, San Diego, CA, pp. 17-20)。

10

#### 【0004】

基質タンパク質 (MP) は、組織の形成、成長、リモデリング及び維持において、炎症反応の重要なメディエーター及びレギュレーターとして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は、先天性疾患、後成的疾患または感染症に起因する生化学変化によって乱され得る。更に、MPは、免疫反応における白血球の移動、増殖、分化及び活性化に影響する。MPは多くの場合、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を含み得るような1若しくは数個のドメインの存在により特徴付けられる。更に、MPを高度にグリコシル化することができ、また、MPは、接着相互作用において役割を果たし得るアルギニン-グリシン-アスパラギン酸 (RGD) トリペプチドモチーフを含有し得る。MPには、フィブロネクチン、コラーゲン、ガレクチン (galectin)、ビトロネクチン及びそのタンパク分解性派生物ソマトメジンBなどの細胞外タンパク質と、細胞接着分子 (CAM)、カドヘリン及びインテグリンなどの細胞接着受容体がある (Ayad, S. ら (1994) *The Extracellular Matrix Facts Book*, Academic Press, San Diego, CA, pp. 216、Ruoslahti, E. (1997) *Kidney Int.* 51: 1413-1417、Sjaastad, M. D. and Nelson, W. J. (1997) *BioEssays* 19: 47-55を参照)。

20

30

#### 【0005】

ホルモンは、循環を介して移動し、標的細胞の表面または内部の特定の受容体に結合する分泌分子である。ホルモンは、多種多様な生化学組成及び作用メカニズムを有しているが、2つのカテゴリーにグループ分けすることができる。1つのカテゴリーには小さな親油性ホルモンが含まれ、これは標的細胞の原形質膜を介して拡散し、サイトゾル受容体または核受容体に結合し、遺伝子の発現を変化させる複合体を形成する。このような分子の例としては、レチン酸、チロキシンや、コレステロール由来のステロイドホルモン、例えばプロゲステロン、エストロゲン、テストステロン、コルチゾール、アルドステロンなどがある。第2のカテゴリーには、原形質膜全体でシグナルを伝達する細胞表面受容体に結合することによって機能する親水性ホルモンが含まれる。このようなホルモンの例としては、カテコールアミン (エピネフリン、ノルエピネフリン) 及びヒスタミンのようなアミノ酸派生物や、グルカゴン、インスリン、ガストリン、セクレチン、コレシストキニン、副腎皮質刺激ホルモン、卵胞刺激ホルモン、黄体化ホルモン、甲状腺刺激ホルモン、パソプレッシンなどのペプチドホルモンがある (例えば、Lodish ら (1995) *Mo*

40

50

molecular Cell Biology, Scientific American Books Inc., New York, NY, pp. 856-864 を参照)。

#### 【0006】

成長及び分化因子は、細胞間情報交換において機能する分泌タンパク質である。活性化のために膜タンパク質との関連またはオリゴマー化が必要な因子もある。これらの因子及びそれらの受容体の間で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を刺激または阻害するような細胞内シグナル伝達経路を誘発する。殆どの成長及び分化因子は、局所的な環境で細胞に作用する（パラクリンシグナリング）。成長及び分化因子の広範なクラスが3つある。第1のクラスには、大型ポリペプチド成長因子、例えば上皮成長因子、線維芽細胞成長因子、トランスフォーミング成長因子、インスリン様成長因子、血小板由来の成長因子が含まれる。第2のクラスには、コロニー刺激因子（CSF）などの造血成長因子が含まれる。造血成長因子は、例えば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これらの幹細胞前駆体など、血液細胞の増殖及び分化を刺激する。第3のクラスには、増殖よりも細胞機能を制御するホルモンとして機能するような、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II、血管作用性小腸ペプチド、ブラジキニンなどの小型ペプチド因子が含まれる。

#### 【0007】

成長及び分化因子は、in vitroでの細胞の腫瘍性転化及びin vivoでの腫瘍進行において重要な役割を果たす。腫瘍細胞による成長因子の不適切な発現は、腫瘍の血管新生及び転移に寄与し得る。造血中に成長因子の調整を誤ると、貧血、白血病、リンパ腫を引き起こし得る。インターフェロンなど或る種の成長因子は、in vivo及びin vitroの両方で腫瘍細胞に細胞毒性がある。更に、腫瘍性タンパク質に構造的にも機能的にも関連している成長因子及び成長因子受容体もある。それに加えて、成長因子は、プロトオンコジーン及び腫瘍抑圧遺伝子の両者の転写調節に影響する（Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor, MI, pp. 1-9 をレビューされたい）。

#### 【0008】

ニューロペプチド及びバソメディエータ（vasomediator）（NP/VM）は、内因性シグナル伝達分子の大きなファミリーを有する。このファミリーには、ニューロペプチドが含まれ、ニューロペプチドホルモン、例えばボンベシン、ニューロペプチド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IIが含まれ、アンギオテンシン、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などの循環系を介した（circulatory system-borne）シグナル伝達分子、血管作用性小腸ペプチド、バソプレッシン及び平滑筋刺激に参与している関連するペプチドが含まれる。NP/VMは、シグナルを直接伝達し、活性を調節するか或いは他の神経伝達物質及びホルモンを放出し、カスケードにおいて触媒酵素として作用することができる。NP/VMの効果は、極度に短いものから長時間持続するものまでである（Martin, C. R. (1985) Endocrine Physiology, Oxford University Press, New York, NY, pp. 57-62 をレビューされたい）。

#### 【0009】

NP/VMは、多数の神経疾患及び心血管疾患に参与している。例えば、ニューロペプチドYは、高血圧、うっ血性心不全、情動障害及び食欲制御に参与している。ソマトスタチンは、腸、膵臓腺房細胞、及び膵臓のβ細胞において分泌を阻害するのみならず、下垂体前葉における成長ホルモン及びプロラクチンの分泌も阻害する。ソマトスタチンレベルの減少は、アルツハイマー病及びパーキンソン病において報告されている。バソプレッシンは、腎臓において作用し、水及びナトリウムの吸収を増加させ、また、より高濃度で、血

管平滑筋の収縮、血小板活性化及び肝臓のグリコーゲン分解を誘発する。バソプレッシン及びその類似体は、尿崩症の治療に臨床的に用いられる。エンドセリン及びアンギオテンシンは、高血圧に関与しており、アンギオテンシンの血漿レベルを低下させるカプトプリルなどの薬剤を用いて血圧を低下させる (Watson, S. and S. Arkin stall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, p p. 194 ; 252 ; 284 ; 55 ; 111)。

#### 【0010】

ニューロペプチドは、痛覚（痛み）において役割を有することも示されている。血管作用性小腸ペプチドは、慢性の神経障害性の痛みにおいて重要な役割を果たすような外観を呈する。ノシセプチンは、オピオイド受容体様1受容体のための内因性リガンドであり、優性に抗侵害受容性効果を有すると考えられており、また、持続性または慢性の痛みの異なる動物モデルにおいて鎮痛特性を有することが示されてきた (Dickinson, T. and Fleetwood-Walker, S. M. (1998) Trends Pharmacol. Sci. 19: 346-348)。

#### 【0011】

シグナルペプチドを含有するその他のタンパク質には、酵素活性を有する分泌タンパク質がある。酵素活性には、例えばオキシドレダクターゼ／デヒドロゲナーゼ活性やトランスフェラーゼ活性、ヒドロラーゼ活性、リアーゼ活性、イソメラーゼ活性、リガーゼ活性がある。例えば、基質メタロプロテイナーゼは、細胞外基質を分解し、それによって腫瘍転移、組織形態形成及び関節炎において重要な役割を果たすような分泌加水分解酵素である (Reponen, P. ら (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、Mignatti, P. and Rifkin, D. B. (1993) Physio. Rev. 73: 161-195)。

#### 【0012】

プロサポシン (Prosaposin) は、SAP前駆体とも呼ばれるが、精漿、乳、髄液を含む幾つかの他の体液において分泌される主要な産物としてセルトリ細胞により同定されてきた。ヒトのスフィンゴリピド症は、プロサポシン遺伝子における突然変異の結果生じ得るものであり、ヒトの生理におけるプロサポシンの重要性を示す (キシモト, Y. ら (1992) J. Lipid Res. 33: 1255-1267)。細胞から分泌されるプロサポシンは、スフィンゴリピド輸送及びシグナル伝達に関与し得る (Hiesberger, T. ら (1998) EMBO J. 17: 4617-4625)。プロサポシンは、マンノース6-リン酸受容体と、マンノース受容体と、殆どの細胞に発現される多機能エンドサイトーシス受容体である低濃度リポタンパク質受容体関連タンパク質とを含む少なくとも3つの独立したメカニズムにより細胞に侵入し得る (前出のHiesberger, T. ら)。プロサポシンは、海馬ニューロン、脊髄α運動ニューロン、小脳顆粒ニューロン、ニューロblastoma細胞を含む種々のニューロン細胞において活性であり、各ニューロン細胞は神経突起の成長を誘発し、細胞死を防ぐ。プロサポシンは、ミエリン形成において役割を有し得る (Madar-Shapiro, L. ら (1999) Biochem. J. 337: 433-443)。シグナル及び輸送の役割に加えて、プロサポシンは、細胞内でタンパク質分解性に切断され、グリコスフィンゴリピド代謝に関与しているリソソーム酵素を活性化するために必要なサポシン (saposin) -A、-B、-C及び-D (スフィンゴリピド活性化タンパク質 (SAP) とも呼ばれる) を形成し得る。サポシンは、リソソーム蓄積症の患者の組織に蓄積する (前出のキシモト)。サポシンBは、セレブロシド硫酸、GM1ガングリオシド及びグロボトリアオシルセラミド (globotriaosylceramide) を含む多種多様な基質の加水分解を誘発する。ヒトのサポシンB欠乏症は、常染色体劣性形質として伝達

され、セレブロシド硫酸の組織蓄積をもたらし、また、重い神経症状に繋がる進行性脱髄に特徴付けられる遺伝したリソソーム蓄積症である変染色性白質ジストロフィに類似した臨床像を生じさせる。この疾患は、ただ1つの塩基において、即ちスレオニンからイソロイシンへのアミノ酸置換を生じさせるサポシンBの23番目のコドンにおけるC→T遷移において、通常の配列とは異なるmRNAによってマーキングされる。この塩基の変化は、結果として極性アミノ酸であるスレオニンを無極性のイソロイシンに置換する (Kretz, K. ら (1990) Proc. Natl. Acad. Sci. USA 7 : 2541-2544)。

#### 【0013】

リポカリン (Lipocalin) は、重要な輸送分子である。各リポカリンは、特定の  
10  
リガンドと関連し、そのリガンドを生物体内の適切な標的部に輸送する。最も特徴付けられたリポカリンの1つであるレチノール結合タンパク質 (RBP) は、肝臓内のストアから標的組織にレチノールを輸送する。高濃度リポタンパク質 (HDL) 及び低濃度リポタンパク質 (LDL) の成分であるアポリポタンパクD (アポD) は、体全体でのコレステロールのターゲティングされた回収及び輸送において機能する。リポカリンは、細胞調節プロセスにも関与している。アポDは、GCDFP-24 (gross-cystic-disease-fluid protein-24) と同じであり、乳房嚢胞液に高レベルで発現されたプロゲステロン/プレグネノロン結合タンパク質である。或る種のヒトの乳癌細胞株におけるアポDの分泌は、細胞増殖の減少及びより分化した表現型への細胞の進行を伴う。同様に、アポD及び別のリポカリン $\alpha_1$  酸糖タンパク質 (AGP) は、  
20  
神経細胞の再生に関与している。AGPは、抗炎症及び免疫抑制作用にも関与している。AGPは、陽性急性期タンパク質 (APP) の1つであり、AGPの循環レベルはストレス及び炎症性刺激に応じて増加する。AGPは、血小板及び好中球の活性化を阻害し、食作用を阻害するような炎症部位に蓄積する。AGPの免疫調節特性は、グリコシル化に起因する。AGPは、40%炭水化物であり、それを著しく酸性且つ可溶性にする。AGPのグリコシル化パターンは急性の応答中に変化し、脱グリコシル化したAGPには免疫抑制作用がない (Flower (1994) FEBS Lett. 354 : 7-11、前出のFlower)。

#### 【0014】

リポカリンは、様々な病状における診断及び予後マーカーとして用いられる。妊娠中及び  
30  
癌の化学療法、腎臓の機能不全、心筋梗塞、関節炎及び多発性硬化症を含む病状の診断及び予後にAGPの血漿レベルをモニターする。RBPは、腎臓における尿細管再吸収のマーカーとして臨床的に用いられ、アポDは、大きな嚢胞性乳房疾患におけるマーカーである (前出のFlower (1996))。

#### 【0015】

新たな分泌タンパク質及びそれをコードするポリヌクレオチドの発見は、細胞増殖異常、自己免疫/炎症性疾患、心血管疾患、神経疾患及び発達障害の診断、予防及び治療において、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及ぼす外来性化合物の影響についての評価において有用である新たな組成を提供することにより、当分野における  
40  
要求を満たす。

#### 【0016】

(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」および「SECP-14」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。本発明の一実施態様では、(a) SEQ ID NO: 1-14からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-14からなる群から選択されたアミノ酸と少なくとも  
50  
も90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ I

D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択された単離されたポリペプチドを提供する。別法では、S E Q I D N O : 1 - 1 4 のアミノ酸配列を含む単離されたポリペプチドを提供する。

【0017】

更に本発明は、(a) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドをコードする単離されたポリヌクレオチドを提供する。別法では、このポリヌクレオチドは、S E Q I D N O : 1 - 1 4 からなる群から選択されたポリペプチドをコードする。更なる別法では、このポリヌクレオチドは、S E Q I D N O : 1 5 - 2 8 からなる群から選択される。

10

【0018】

更に、本発明は、(a) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドをコードするポリヌクレオチドと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を提供する。別法では、本発明は、この組換えポリヌクレオチドで形質転換された細胞を提供する。更なる別法では、本発明は、この組換えポリヌクレオチドを含む遺伝子組換え生物を提供する。

20

【0019】

また、本発明は、(a) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドの産生方法を提供する。この方法は、(a) このポリペプチドの発現に好適な条件下で、この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されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養する過程と、(b) このように発現したポリペプチドを回収する過程とを含む。

30

【0020】

更に、本発明は(a) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドに特異的に結合する単離された抗体を提供する。

40

【0021】

更に、本発明は、(a) S E Q I D N O : 1 5 - 2 8 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(b) S E Q I D N O : 1 5 - 2 8 からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌク

50

レオチド配列を含む天然のポリヌクレオチドと、(c)前記(a)のポリヌクレオチドに相補的なポリヌクレオチドと、(d)前記(b)のポリヌクレオチドに相補的なポリヌクレオチドと、(e)前記(a) - (d)のRNA等価物とで構成される群から選択された単離されたポリヌクレオチドを提供する。別法では、このポリヌクレオチドは、少なくとも60個の連続するヌクレオチドを含む。

#### 【0022】

更に本発明は、(a)SEQ ID NO: 15-28からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(b)SEQ ID NO: 15-28からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(c)前記(a)のポリヌクレオチドに相補的なポリヌクレオチドと、(d)前記(b)のポリヌクレオチドに相補的なポリヌクレオチドと、(e)前記(a) - (d)のRNA等価物とで構成される群から選択されたポリヌクレオチド配列を有するサンプルにおける標的ポリヌクレオチドを検出する方法を提供する。この方法は、(a)前記サンプル内の標的ポリヌクレオチドと相補的な配列を含む少なくとも20個の連続するヌクレオチドを含むプローブと前記サンプルをハイブリダイズさせる過程であって、前記プローブと前記標的ポリヌクレオチドまたはその断片とでハイブリダイゼーション複合体が形成される条件下で、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、該過程と、(b)前記ハイブリダイゼーション複合体の存在するか否かを検出し、存在する場合には必要に応じてその収量を測定する過程とを含む。別法では、前記プローブは、少なくとも60個の連続するヌクレオチドを含む。

10

20

#### 【0023】

更に本発明は、(a)SEQ ID NO: 15-28からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(b)SEQ ID NO: 15-28からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(c)前記(a)のポリヌクレオチドに相補的なポリヌクレオチドと、(d)前記(b)のポリヌクレオチドに相補的なポリヌクレオチドと、(e)前記(a) - (d)のRNA等価物とで構成される群から選択されたポリヌクレオチド配列を有するサンプルにおける標的ポリヌクレオチドを検出する方法を提供する。この方法は、(a)ポリメラーゼ連鎖反応増幅法を用いて、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)増幅された前記標的ポリヌクレオチドまたはその断片が存在するか否かを検出し、存在する場合には必要に応じてその収量を測定する過程とを含む。

30

40

#### 【0024】

更に本発明は、(a)SEQ ID NO: 1-14からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b)SEQ ID NO: 1-14からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c)SEQ ID NO: 1-14からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d)SEQ ID NO: 1-14からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択された有効量のポリペプチド及び好適な医薬用賦形剤を含む組成物を提供する。一実施例では、SEQ ID NO: 1-14からなる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。本発明は更に、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。

#### 【0025】

更に本発明は、(a)SEQ ID NO: 1-14からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b)SEQ ID NO: 1-14からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c)SEQ ID NO: 1-14からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d)SEQ ID NO: 1-14からなる群か

50

ら選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドのアゴニストとして有効な化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを含むサンプルを化合物に曝露する過程と、(b) このサンプル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 S E C P 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。

#### 【0026】

更に、本発明は、(a) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸と少なくとも 90 % の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドのアнтаゴニストとして有効な化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを含むサンプルを化合物に曝露する過程と、(b) このサンプルのアンタ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ンタ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 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。

10

20

#### 【0027】

更に本発明は、(a) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸と少なくとも 90 % の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドに特異的に結合する化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを好適な条件下で少なくとも 1 つの試験化合物と結合させる過程と、(b) このポリペプチドとこの試験化合物との結合を検出して、このポリペプチドと特異的に結合する化合物を同定する過程とを含む。

30

#### 【0028】

更に本発明は、(a) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸と少なくとも 90 % の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) S E Q I D N O : 1 - 1 4 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチド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過程と、(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過程と、(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、この試験化合物の非存在下でのこのポリペプチドの活性とを比較する過程とを含み、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。

40

#### 【0029】

更に本発明は、S E Q I D N O : 1 5 - 2 8 からなる群から選択された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有効な化合物をスクリーニングする方法であって、(a) この標的ポリヌクレオチドを含むサンプルを化合物に曝露する過程と、(b) この標的ポリヌクレオチドの発現の変化を検出する過程とを含む、該スクリーニング方法

50

を提供する。

#### 【0030】

本発明はさらに、試験化合物の毒性を評価する方法を提供する。この方法は、(a) 核酸を含む生体サンプルを前記試験化合物で処理する過程と、(b) 処理した前記生体サンプルの核酸をプローブとハイブリダイズする過程と、(c) ハイブリダイゼーション複合体の収量を測定する過程と、(d) 前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処理の生体サンプルにおけるハイブリダイゼーション複合体の収量とを比較する過程とを含み、前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量の差異が試験化合物の毒性を示唆する。この方法における前記プローブが、(1) SEQ ID NO: 15-28 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(2) SEQ ID NO: 15-28 からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(3) 前記(1)に相補的な配列を有するポリヌクレオチドと、(4) 前記(2)のポリヌクレオチドに相補的なポリヌクレオチドと、(5) 前記(1)-(4)のRNA等価物とで構成される群から選択されたポリヌクレオチドの少なくとも20個の連続するヌクレオチドを含む。また、前記ハイブリダイゼーションは、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る。また、前記標的ポリヌクレオチドが、(1) SEQ ID NO: 15-28 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(2) SEQ ID NO: 15-28 からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(3) 前記(1)のポリヌクレオチドに相補的なポリヌクレオチドと、(4) 前記(2)のポリヌクレオチドに相補的なポリヌクレオチドと、(5) 前記(1)-(4)のRNA等価物とで構成される群から選択される。別法では、標的ポリヌクレオチドは、その標的ポリヌクレオチドの断片である。

10

20

30

#### 【0031】

(本発明の記載について)

本発明のタンパク質及び核酸配列、方法について説明する前に、本発明は、ここに開示した特定の装置及び材料、方法に限定されず、その実施形態を変更できることを理解されたい。また、ここで用いられる用語は、特定の実施例のみを説明する目的で用いられたものであり、後述の請求の範囲によってのみ限定され、本発明の範囲を限定することを意図したものではないということも理解されたい。

#### 【0032】

本明細書及び請求の範囲において単数形を表す「或る」、「その(この等)」は、文脈で明確に示していない場合は複数形を含むことに注意されたい。従って、例えば「或る宿主細胞」は複数の宿主細胞を含み、その「抗体」は複数の抗体は含まれ、当業者には周知の等価物なども含まれる。

#### 【0033】

本明細書で用いた全ての科学技術用語は、別の方法で定義されていない限り、本発明の属する技術分野の一般的な技術者が普通に解釈する意味と同じである。本明細書で記述したものと類似、或いは同等の全ての装置及び材料、方法は本発明の実施及びテストに使用できるが、好適な装置及び材料、方法をここに記す。本明細書に記載の全ての文献は、本発明に関連して使用する可能性のある文献に記載された細胞系、プロトコル、試薬、ベクターを記述し開示するために引用した。従来発明を引用したからと言って、本発明の新規性が損なわれると解釈されるものではない。

40

#### 【0034】

(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。

50

## 【0035】

用語「アゴニスト」は、S E C Pの生物学的活性を強める、或いは模倣する分子を指す。このアゴニストは、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。

## 【0036】

用語「アレル変異配列」は、S E C Pをコードする遺伝子の別の形を指す。アレル変異配列は、核酸配列における少なくとも1つの変異によって生じ、変異m R N A若しくは変異ポリペプチドになり、これらの構造や機能は変わる場合もあれば変わらない場合もある。ある遺伝子は、天然型のアレル変異配列が存在しないもの、1つ或いは多数存在するものがある。一般にアレル変異配列を生じる変異は、ヌクレオチドの自然な欠失、付加、或いは置換による。これらの各変異は、単独或いは他の変異と同時に起こり、所与の配列内で1回若しくは複数回生じる。

10

## 【0037】

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アレル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、負に荷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含まれ、正に荷電したアミノ酸にはリシン及びアルギニンが含まれ得る。類似の親水性の値をもち極性非荷電側鎖を有するアミノ酸には、アスパラギン、グルタミン、セリン、トレオニンが含まれ得る。類似の親水性の値をもち非荷電側鎖を有するアミノ酸には、ロイシン、イソロイシン、バリン、グリシン、アラニン、フェニルアラニン及びチロシンが含まれ得る。

20

## 【0038】

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

30

## 【0039】

用語「増幅」は、核酸配列の複製物を作製することに関連する。一般に増幅は、この技術分野で周知のポリメラーゼ連鎖反応（PCR）技術によって行われる。

## 【0040】

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。

40

## 【0041】

用語「抗体」は、抗原決定基と結合可能なF a b及びF（a b'）<sub>2</sub>、及びそれらの断片、F v断片などの無傷の分子を指す。S E C Pポリペプチドと結合する抗体は、抗体を免疫する小ペプチドを含む無傷の分子またはその断片を用いて作製可能である。動物（例えば、マウス、ラット、若しくはウサギ）を免疫化するのに使用されるポリペプチド或いはオリゴペプチドは、R N Aの翻訳から、或いは化学的に合成可能であり、必要に応じて担体タンパク質と結合させることも可能である。ペプチドと化学的に結合した一般に用いら

50

れる担体は、ウシ血清アルブミン、チログロブリン、及びキーホールリンペットヘモニアン（KLH）を含む。次ぎに、この結合したペプチドを用いて動物を免疫化する。

#### 【0042】

用語「抗原決定基」は、特定の抗体と接触する分子の領域（即ちエピトープ）を指す。タンパク質或いはタンパク質の断片が、宿主動物を免疫化するのに用いられるとき、このタンパク質の種々の領域は、抗原決定基（タンパク質上の特定の領域或いは三次元構造体）に特異的に結合する抗体の産生を誘発し得る。抗原決定基は、抗体と結合するために無傷の抗原（即ち、免疫応答を引き出すために用いられる免疫原）と競合し得る。

#### 【0043】

本明細書において「アンチセンス」は、特定の核酸配列のセンス（コーディング）鎖と塩基対を形成し得る任意の組成物を指す。アンチセンス成分には、DNAと、RNAと、ペプチド核酸（PNA）と、ホスホロチオネートやメチルホスホネート、ベンジルホスホネート（benzylphosphonate）などの修飾された骨格（backbone linkage）を有するオリゴヌクレオチドと、2'-メトキシエチル糖または2'-メトキシエトキシ糖などの修飾された糖を有するオリゴヌクレオチドと、5-メチルシトシンまたは2'-deoxyuracil、7-deaza-2'-deoxyguanosineなどの修飾された塩基を有するオリゴヌクレオチドを含み得る。アンチセンス分子は、化学合成や転写を含む任意の方法で作ることができる。相補的アンチセンス分子は、一度細胞に導入されると、細胞によって作られた天然の核酸配列と塩基対となって二重鎖を形成し、転写や翻訳を阻害する。「負」または「マイナス」という表現はアンチセンス鎖であり、「正」または「プラス」という表現はセンス鎖である。

#### 【0044】

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

#### 【0045】

用語「相補的」は、塩基対合によってアニールする2つの一本鎖核酸配列間の関係を指す。例えば、配列「5' AGT 3'」は相補的な配列「3' TCA 5'」と対をなす。

#### 【0046】

「所与のポリヌクレオチド配列を含む組成物」または「所与のアミノ酸配列を含む組成物」は広い意味で、所与のヌクレオチド配列若しくはアミノ酸配列を含む任意の組成物を指す。この組成物は、乾燥した製剤或いは水溶液を含み得る。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、プローブは、塩（例えば、NaCl）及び界面活性剤（例えば、SDS：ドデシル硫酸ナトリウム）、その他の物質（例えば、デンハート液、乾燥ミルク、サケ精子DNAなど）を含む水溶液に展開され得る。

#### 【0047】

「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（Applied Biosystems, Foster City CA）を用いて5'及び／または3'の方向に伸長され、再度シーケンシングされた核酸配列、またはGELVIEW断片構築システム（GCG, Madison, WI）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。伸長及び重複の両方によって構築されるコンセンサス配列もある。

#### 【0048】

用語「保存的なアミノ酸置換」は、元のタンパク質の特性を殆ど変えない置換を指す。即

ち、置換によってそのタンパク質の構造や機能が大きくは変わらず、そのタンパク質の構造、特にその機能が保存される。以下に、あるタンパク質の元のアミノ酸が別のアミノ酸に置換される保存的なアミノ酸置換を示す。

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

10

20

保存されたアミノ酸の置換では、通常、a) 置換された領域のポリペプチドの骨格構造、例えば、 $\beta$ シートや $\alpha$ ヘリックス高次構造、b) 置換された部位の分子の電荷または疎水性、及び／または、c) 側鎖の大半が維持される。

#### 【0049】

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

#### 【0050】

用語「誘導体」は、化学修飾されたポリヌクレオチドまたはポリペプチドを指す。ポリヌクレオチド配列の化学修飾には、例えば、アルキル基、アシル基、ヒドロキシル基、或いはアミノ基による水素の置換がある。誘導体ポリヌクレオチドは、自然分子（未修飾の分子）の生物学的或いは免疫学的機能の少なくとも1つを維持するポリペプチドをコードする。誘導体ポリペプチドとは、もとのポリペプチドの生物学的機能、或いは免疫学的機能の少なくとも1つを維持する、グリコシル化、ポリエチレングリコール化、或いは任意の同様のプロセスによって修飾されたポリペプチドのことである。

30

#### 【0051】

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

40

#### 【0052】

用語「断片」は、SECPまたはSECPをコードするポリヌクレオチドの固有の部分であって、その親配列（parent sequence）と同一であるがその配列より長さが短いものを指す。「断片」の最大の長さは、親配列から1つのヌクレオチド／アミノ酸残基を差し引いた長さである。例えば、ある断片は、5～1000個の連続するヌクレオチド或いはアミノ酸残基を含む。プローブ、プライマー、抗原、治療用分子、またはその他の目的に用いる断片は、少なくとも5、10、15、16、20、25、30、40、50、60、75、100、150、250若しくは500個の連続するヌクレオチド或いはアミノ酸残基の長さである。断片は、優先的に分子の特定の領域から選択される場合もある。例えば、ポリペプチド断片は、所定の配列に示された最初の250若しくは5

50

00のアミノ酸（或いは、ポリペプチドの最初の25%または50%）から選択された連続するアミノ酸の所定の長さを含み得る。これらの長さは一例であり、配列表及び表、図面を含む明細書に記載の任意の長さが、本発明の実施例に含まれ得る。

#### 【0053】

SEQ ID NO: 15-28の断片は、例えば、この断片を得たゲノム内の他の配列とは異なる、SEQ ID NO: 15-28を明確に同定する固有のポリヌクレオチド配列の領域を含む。SEQ ID NO: 15-28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: 15-28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。ある断片と一致するSEQ ID NO: 15-28の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に測定できる。 10

#### 【0054】

SEQ ID NO: 1-14のある断片は、SEQ ID NO: 15-28のある断片によってコードされる。SEQ ID NO: 1-14のある断片は、SEQ ID NO: 1-14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: 1-14のある断片は、SEQ ID NO: 1-14を特異的に認識する抗体の開発における免疫原性ペプチドとして有用である。ある断片と一致するSEQ ID NO: 1-14の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に決定できる。

#### 【0055】

「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）、それに続く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。 20

#### 【0056】

「相同性」は、2つ以上のポリヌクレオチド配列間または2つ以上のポリペプチド配列間の配列類似性である。この配列類似性は配列同一性と言い換えることができる。

#### 【0057】

ポリヌクレオチド配列についての用語「パーセントの同一性」または「%の同一性」とは、標準化されたアルゴリズムを用いてアラインメントされる、2つ以上のポリヌクレオチド配列間の一致する残基の百分率のことである。このようなアルゴリズムは、標準化され再現できる方法で、2つの配列間のアラインメントを最適化するべく、配列にギャップを挿入して、より意味をもつ2つの配列間の比較を行うことができる。 30

#### 【0058】

ポリヌクレオチド配列間の同一性のパーセントは、MEGALIGN version 3.12e配列アラインメントプログラムに組込まれるCLUSTAL Vアルゴリズムのデフォルトパラメータを用いて決定可能である。このプログラムはLASERGENEソフトウェアパッケージの一部であり、分子生物学分析プログラム一式（DNASTAR, Madison WI）である。このCLUSTAL Vは、Higgins, D. G. 及び P. M. Sharp (1989) CABIOS 5:151-153、Higgins, D. G. ら (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。同一性のパーセントは、アラインメントされたポリヌクレオチド配列間の「類似性のパーセント」としてCLUSTAL Vによって報告される。 40

#### 【0059】

或いは、よく用いられている無料の配列比較アルゴリズム一式が、National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S. F. ら (1990) J. Mol. Biol. 215:4 50

03-410)により提供されており、これは幾つかのソースから利用可能であって、例えばNCBI (メリーランド州ベセスダ)やインターネット (<http://WWW.ncbi.nlm.nih.gov/BLAST/>)上で利用可能である。このBLASTソフトウェア式には、既知のポリヌクレオチド配列と様々なデータベースの別のポリヌクレオチド配列とのアラインメントに用いられる「blastn」を含む、様々な配列分析プログラムが含まれる。「BLAST 2 Sequences」と呼ばれるツールが入手可能であり、2つのヌクレオチド配列の対を直接比較するために用いられる。「BLAST 2 Sequences」は、<http://WWW.ncbi.nlm.nih.gov/gorf/b12.html>にアクセスして、対話形式で利用ができる。「BLAST 2 Sequences」ツールは、blastn 及び blastp (以下に記載)の両方に用いることができる。BLASTプログラムは、一般的には、デフォルトを設定するギャップ及び他のパラメータと共に用いられる。例えば、2つのヌクレオチド配列を比較する場合、ある者は、デフォルトパラメータに設定された「BLAST 2 Sequences」ツールVersion 2.0.12 (April-21-2000)でblastnを使用するであろう。そのようなデフォルトパラメータは、例えば、以下のようになる。

10

#### 【0060】

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 及び Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

20

同一性のパーセントは、例えば、特定の配列番号で決められた、所定の配列の全長に対して測定してもよいし、それより短い長さに対して、例えば、ある大きな所定の配列から得られた断片、例えば、連続する少なくとも、20または30、40、50、70、100、200のヌクレオチドの断片の長さに対して測定してもよい。このような長さは単なる例であり、配列表及び表、図面を含む明細書に記載の配列の任意の長さの断片を用いて、同一性のパーセントが測定される長さを示すことができる。

30

#### 【0061】

高い同一性を示さない核酸配列でも、遺伝子コードの縮重によって類似のアミノ酸配列をコードし得る。縮重を利用して核酸配列を変え、それぞれが実質的に同じタンパク質をコードする様々な核酸配列を作製できることを理解されたい。

#### 【0062】

ポリペプチド配列に用いられる用語「パーセントの同一性」及び「%の同一性」とは、標準化されたアルゴリズムを用いてアラインメントされる2つ以上のポリペプチド配列間の一致する残基の百分率のことである。ポリペプチド配列アラインメントの方法は周知である。アラインメント方法の中には、保存的なアミノ酸置換を考慮したものもある。詳細に上述したこのような保存的な置換は、一般に、置換部位の電荷や疎水性が保存され、ポリペプチドの構造(従って機能も)が保存される。

40

#### 【0063】

ポリペプチド配列間の同一性のパーセントは、MEGALIGN バージョン3.12e 配列アラインメントプログラム(上述)に組込まれるCLUSTAL Vアルゴリズムのデフォルトパラメータを用いて決定可能である。CLUSTAL Vを用いる対方式のポリペプチド配列のアラインメントの場合、デフォルトパラメータは、Ktupple=1、gap penalty=3、window=5、及び「diagonals saved」=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、アラインメントされたポリペプチド配

50

列の対の同一性のパーセントは、「類似性のパーセント」としてCLUSTAL Vによって報告される。

#### 【0064】

別法では、NCBI BLASTソフトウェア一式が用いられる。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。そのようなデフォルトパラメータは、例えば、以下のようにする。

#### 【0065】

Matrix: BLOSUM62

Open Gap: 11 及び Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

同一性のパーセントは、例えば、特定の配列番号で決められた、所定のポリペプチド配列の全長に対して測定してもよいし、それより短い長さに対して、例えば、ある大きな所定のポリペプチド配列から得られた断片、例えば、連続する少なくとも15、20または30、40、50、70、150の残基の断片の長さに対して測定してもよい。このような長さは単なる例であり、配列表及び表、図面を含む明細書に記載の配列の任意の長さの断片を用いて、同一性のパーセントが測定される長さを示すことができる。

#### 【0066】

「ヒト人工染色体 (HAC)」は、約6 kb (キロベース) ~ 10 MbのサイズのDNA配列を含み得る、安定した有糸分裂染色体の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。

#### 【0067】

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

#### 【0068】

「ハイブリダイゼーション」とは、所定のハイブリダイゼーション条件下で、ある一本鎖ポリヌクレオチドがある相補的な一本鎖と塩基対を形成するアニーリングのプロセスである。特異的なハイブリダイゼーションとは、2つの核酸配列が高い相同性を有することを意味する。アニーリングが許容される条件下で、特異的なハイブリダイゼーション複合体が形成され、洗浄過程の後もハイブリダイズしたままである。洗浄過程は、ハイブリダイゼーションプロセスの厳密性即ちストリンジェンシー (stringency) の決定において特に重要であり、よりストリンジェントな条件では、非特異的な結合、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。核酸配列間のアニーリングが許容される条件は、当業者によって日常的に決定され、ハイブリダイゼーションの間は一定であるが、洗浄過程は、目的のストリンジェンシーにするためにその最中に条件の変更が可能であり、ハイブリダイゼーション特異性が得られる。アニーリングが許容される条件は、例えば、温度が68℃で、約6×SSC、約1% (w/v) のSDS、並びに約100 µg/mlのせん断された変性サケ精子DNAが含まれる。

#### 【0069】

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、洗浄過程を行う際の温度によっても左右される。この洗浄温度は通常、特定のイオン強度とpHにおける特定の配列の熱融点 (Tm) より約5~20℃低く選択される。このTmは、(所定のイオン強度とpHの下) 標的の配列の50%が完全に一致するプローブと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ョンの条件は、周知であり、Sambrook, J. らによる, 1989, Molecular Cloning: A Laboratory Manual, 第2版の1-3巻, Cold Spring Ha

10

20

30

40

50

rbor Press, Plainview NY; 特に2巻の9章に記載されている。

【0070】

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0.2x SSC及び約1% SDSの存在下で、約68℃で1時間の洗浄過程を含む。別法では、65℃、60℃、55℃、42℃の温度で行う。SSCの濃度は、約0.1% SDSの存在下で、約0.1~2x SSCの範囲である。通常は、ブロッキング試薬を用いて非特異的なハイブリダイゼーションを阻止する。このようなブロッキング試薬には、例えば、約100~200 µg/mlのせん断された変性サケ精子DNAが含まれる。約35~50% v/vの濃度のホルムアミドなどの有機溶剤が、例えば、RNAとDNAのハイブリダイゼーションなど特定の場合に用いることができる。これらの洗浄条件の有用な改変は、当業者には周知である。特に高いストリンジェントな条件でのハイブリダイゼーションは、ヌクレオチド間の進化における類似性を示唆し得る。このような類似性は、それらのヌクレオチド及びコードされたポリペプチドが類似の役割を果たしていることを強く示唆する。

【0071】

用語「ハイブリダイゼーション複合体」は、相補的な塩基間の水素結合によって、形成された2つの核酸配列の複合体を指す。ハイブリダイゼーション複合体は溶液中（例えば、C<sub>0</sub>tまたはR<sub>0</sub>t分析）で形成されるか、或いは溶液中の1つの核酸配列と固体の支持物（例えば、紙、膜、フィルタ、チップ、ピン、或いはスライドガラス、または細胞及びその核酸を固定する任意の適当な基板）に固定されたもう一つの核酸配列とで形成され得る。

【0072】

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

【0073】

「免疫応答」は、炎症性疾患及び外傷、免疫異常、感染症、遺伝病などに関連する症状を指す。これらの症状は、細胞系及び全身防衛系に影響を及ぼすサイトカイン及びケモカイン、別の情報伝達分子などの様々な因子の発現という特徴をもつ。

【0074】

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産生方法に有用な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。

【0075】

用語「マイクロアレイ」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。

【0076】

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

【0077】

用語「調節する」は、SECPの活性を変化させることを指す。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。

【0078】

用語「核酸」及び「核酸配列」は、ヌクレオチド、オリゴヌクレオチド、ポリヌクレオチド、或いはそれらの断片を指し、一本鎖若しくは二本鎖であって、センス鎖或いはアンチセンス鎖であるゲノム起源若しくは合成起源のDNA或いはRNA、ペプチド核酸(PNA)、任意のDNA様物質、及びRNA様物質である。

【0079】

「機能的に結合した」は、第1の核酸配列と第2の核酸配列が機能的な関係にある状態を

指す。例えば、プロモーターがコード配列の転写または発現に影響を与える場合、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。一般に、機能的に結合したDNA配列は、同じリーディングフレーム内で2つのタンパク質をコードする領域が結合する必要がある場合は、非常に近接或いは連続する。

#### 【0080】

「ペプチド核酸 (PNA)」は、末端がリシンで終わるアミノ酸残基のペプチド骨格に結合した、少なくとも約5ヌクレオチドの長さのオリゴヌクレオチドを含む、アンチセンス分子または抗遺伝子剤を指す。この末端のリシンにより、この組成物が溶解性となる。PNAは、相補的な一本鎖DNAやRNAに優先的に結合して転写物の伸長を止め、ポリエチレングリコール化して細胞における寿命を延ばし得る。

10

#### 【0081】

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

#### 【0082】

「プローブ」とは、同一配列或いはアレル核酸配列、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、検出可能な標識またはレポーター分子が結合され単離されたオリゴヌクレオチドやポリヌクレオチドである。典型的な標識には、放射性アイソトープ及びリガンド、化学発光試薬、酵素がある。「プライマー」とは、相補的な塩基対を形成して標的のポリヌクレオチドにアニーリング可能な、通常はDNAオリゴヌクレオチドである短い核酸である。プライマーがポリヌクレオチドにアニーリングした後、あるDNAポリメラーゼ酵素によって、標的のDNA一本鎖に沿って伸長される。プライマーの組は、例えば、PCR法における核酸配列の増幅 (及び同定) に用いることができる。

20

#### 【0083】

本発明に用いられるプローブ及びプライマーは、既知の配列の少なくとも15の連続するヌクレオチドを含む。特異性を高めるために、より長いプローブ及びプライマーが用いることも可能である。例えば、開示した核酸配列の連続する少なくとも20または25、30、40、50、60、70、80、90、100、150のヌクレオチドを含む。プローブ及びプライマーは、上記した例より相当長いものも用いることができ、本明細書の表及び図面、配列表に示された任意の長さのヌクレオチドも用いることができることを理解されたい。

30

#### 【0084】

プローブ及びプライマーの準備及び使用方法については、例えば、Sambrook, J. ら (1989) 「Molecular Cloning: A Laboratory Manual」第2版の1-3巻 (Cold Spring Harbor Press, Plainview NY)、またはAusubel, F. M. ら (1987) 「Current Protocols in Molecular Biology」 (Greene Pub. 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用のプライマー対は、そのような目的のためのコンピュータプログラム、例えばPrimer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA) を用いて既知の配列から導出し得る

40

プライマーとして用いるオリゴヌクレオチドは、当分野で周知のプライマー選択用のコンピュータプログラムで選択される。例えば、OLIGO 4.06ソフトウェアは、それぞれが最大100ヌクレオチドまでのPCR用のプライマー対の選択、及び32,000

50

塩基までの入力ポリヌクレオチド配列から最大5,000ヌクレオチドまでの大きなポリヌクレオチド及びオリゴヌクレオチドの分析に有用である。類似のプライマー選択用プログラムには、能力を拡大する追加の機能が含まれている。例えば、PrimOUプライマー選択プログラム（テキサス州ダラスのGenome Center at University of Texas South West Medical Centerより入手可能）は、メガベース配列から特定のプライマーを選択できるため、ゲノムワイドスコープ（genome-wide scope）におけるプライマーの設計に有用である。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 Center for Genome Researchより入手可能）によって、ユーザーは、プライマー結合部位として避けたい配列を指定できる「ミスプライミングライブラリ（mispriming library）」を入力できる。また、Primer3は、特にマイクロアレイのオリゴヌクレオチドの選択に有用である（後者2つのプライマー選択プログラムのソースコードは、それぞれのソースから得ることができ、ユーザーのニーズを満たすように変更することもできる）。PrimerGenプログラム（英国ケンブリッジのUK Human Genome Mapping Project Resource Centreより入手可能）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するため、アラインメントされた核酸配列の最も保存された領域或いは最も保存されていない領域のどちらかとハイブリダイズするプライマーを選択することができる。従って、このプログラムは、固有及び保存されたオリゴヌクレオチドやポリヌクレオチドの断片の同定に有用である。上記した任意の選択方法で同定されたオリゴヌクレオチドやポリヌクレオチドの断片は、例えば、PCR法やシーケンシングプライマー、マイクロアレイエレメント、或いはサンプルの核酸の完全或いは部分的に相補的なポリヌクレオチドを同定する特定のプローブなどの、ハイブリダイゼーション技術に有用である。オリゴヌクレオチドの選択方法は、上記した方法に制限されるものではない。

#### 【0085】

本明細書における「組換え核酸」は天然の配列ではなく、2つ以上の配列の離れたセグメントを人工的に組み合わせた配列である。この人工の組合せは、化学合成によって作られる場合も多いが、前出のSambrookに記載されたような遺伝子工学の技術を用いて核酸の離れたセグメントを人工的に操作する方がより一般的である。この「組換え核酸」には、単に核酸の一部の追加または置換、欠失によって変更された核酸も含む。組換え核酸は、あるプロモーター配列に機能的に結合した核酸配列を含む場合もある。このような組換え核酸は、例えば、ある細胞を形質転換するのに用いられるベクターの一部であり得る。

#### 【0086】

別法では、このような組換え核酸は、この組換え核酸を発現する哺乳動物のワクチン接種に用いると、その哺乳動物の防衛的な免疫応答を誘発する、ワクシニアウイルスに基づいたウイルスベクターの一部であり得る。

#### 【0087】

「調節エレメント」は、通常は遺伝子の非翻訳領域に由来する核酸配列であり、エンハンサー、プロモーター、イントロン及び5'及び3'の非翻訳領域（UTR）を含む。調節エレメントは、転写や翻訳、またはRNAの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。

#### 【0088】

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

#### 【0089】

本明細書において、DNA配列に対する「RNA等価物」とは、基準となるDNA配列と

同じ直鎖の核酸配列から構成されるが、窒素性塩基のチミンがウラシルに置換され、糖鎖の背骨がデオキシリボースではなくリボースからなる。

【0090】

用語「サンプル」は、最も広義に用いる。S E C P、S E C P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。

【0091】

用語「特異的結合」及び「特異的に結合する」は、タンパク質若しくはペプチドと、アゴニスト、抗体、アンタゴニスト、小分子、若しくは任意の天然若しくは合成の結合組成物との間の相互作用を指す。この相互作用は、結合する分子によって認識される、例えば抗原決定基つまりエピトープなどタンパク質の特定の構造の存在によって左右される。例えば、抗体がエピトープ「A」に対して特異的である場合、結合していない標識した「A」及び抗体を含む反応液に、エピトープAを含むポリペプチド或いは結合していない無標識の「A」が存在すると、抗体と結合する標識Aの量が減少する。

10

【0092】

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

20

【0093】

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

【0094】

用語「基板」は、任意の好適な固体或いは半固体の支持物を指し、膜及びフィルタ、チップ、スライド、ウェハ、ファイバー、磁気または非磁気ビード、ゲル、チューブ、プレート、ポリマー、微小粒子、毛細管が含まれる。この基板には、壁または塹壕、ピン、チャンネル、細孔などの様々な表面形態があり、そこにポリヌクレオチドやポリペプチドが結合する。

【0095】

「転写イメージ」は、所与の条件、所与の時間における特定の細胞の種類または組織による集合的遺伝子発現のパターンを指す。

30

【0096】

「形質転換」とは、外来DNAが受容細胞に導入されるプロセスのことである。形質転換は、当分野で周知の種々の方法により、自然或いは人工の条件下で起こり、原核宿主細胞若しくは真核宿主細胞の中に外来核酸配列を挿入する任意の周知の方法によって行うことができる。この形質転換の方法は、形質転換される宿主細胞のタイプによって選択される。この方法には、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）、リポフェクション、及び微粒子照射が含まれるが、これらに限定されるものではない。「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。

40

【0097】

本明細書における「遺伝子組換え生物」とは、当分野で周知の遺伝子組換え技術などを用いて、人間が生物の1つ以上の細胞に異種の核酸を導入した任意の生物であり、動物及び植物を含むが、それらに限定されるものではない。微量注入や組換えウイルスに感染させるなどの慎重な遺伝子操作によって、細胞の前駆体に直接或いは間接的に異種核酸を細胞に導入する。「遺伝子操作」とは、典型的な交雑育種や*in vitro*での受精ではなく、組換えDNA分子を導入することである。本発明に従った遺伝子組換え生物には、細

50

菌及びラン藻類、菌類、植物、動物が含まれる。本発明の単離されたDNAは、当分野で周知の、例えば、感染、形質移入、形質転換、トランス接合 (trans conjugation) などの方法によって、宿主に導入することができる。本発明のDNAをそのような生物に導入する技術は周知であり、前出の Sambrookら (1989) に記載されている。

#### 【0098】

特定の核酸配列の「変異配列」とは、デフォルトパラメータ設定の「BLAST 2 Sequences」ツール Version 2.0.9 (May-07-1999) を用いる blastn によって、ある核酸配列のある長さに対する該特定の核酸配列の同一性が、少なくとも40%と決定された核酸配列のことである。このような核酸の対は、ある長さにおいて、例えば、少なくとも50%または60%、70%、80%、85%、90%、95%、96%、97%、98%、99%或いはそれ以上の同一性を示し得る。ある変異配列は、例えば、「アレル」変異配列 (上述) または「スプライス」変異配列、「種」変異配列、「多型」変異配列と表すことができる。スプライス変異配列は基準分子と同一性が極めて高い可能性があるが、mRNA プロセッシング中のエキソンの択一的スプライシングによってポリヌクレオチドの数が多くなったり、少なくなったりする。対応するポリペプチドは、基準分子に存在する追加の機能ドメインを有したり、基準分子に存在するドメインが欠落したりし得る。種変異体は、種によって異なるポリヌクレオチド配列である。得られるポリペプチドは、互いに高いアミノ酸同一性を有する。多型変異体は、所与の種の個体間の特定の遺伝子のポリヌクレオチド配列における変異体である。多型変異体はまた、ポリヌクレオチド配列の1つのヌクレオチドが異なる「一塩基多型」(SNP) も含み得る。SNPの存在は、例えば、或る集団、病態、病態の性向を示唆し得る。

#### 【0099】

特定のポリペプチド配列の「変異体」とは、デフォルトパラメータ設定の「BLAST 2 Sequences」ツール Version 2.0.9 (May-07-1999) を用いる blastp によって、あるポリペプチド配列のある長さに対する該特定のポリペプチド配列の同一性が、少なくとも40%と決定されたポリペプチド配列のことである。このようなポリペプチドの対は、ある長さにおいて、例えば、少なくとも50%または60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或いはそれ以上の同一性を示し得る。

#### 【0100】

(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/炎症性疾患、心血管疾患、神経疾患及び発達障害の診断、治療、及び予防に関する。

#### 【0101】

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識別番号を示す。各ポリヌクレオチドおよびそれに対応するポリペプチドは、1つのインサイトプロジェクト識別番号 (IncYTE Project ID) に相関する。各ポリペプチド配列は、記載されているようにポリペプチド配列識別番号 (Polypeptide SEQ ID NO : ) およびインサイトポリペプチド配列番号 (IncYTE Polypeptide ID) の両方によって示されている。各ポリヌクレオチド配列は、記載されているようにポリヌクレオチド配列識別番号 (Polypeptide SEQ ID NO : ) およびインサイトポリヌクレオチドコンセンサス配列番号 (IncYTE Polypeptide ID) の両方によって示されている。

#### 【0102】

表2は、GenBank タンパク質 (genept) データベースに対するBLAST解析により同定された本発明のポリペプチドに相同性を有する配列を示す。列1および列2はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号 (Polypeptide SEQ ID NO : ) およびそれに対応するインサイトポリペプチド配列番

号 (I n c y t e P o l y p e p t i d e I D) を示す。列 3 は、G e n B a n k の最も近い相同体の G e n B a n k の識別番号 (G e n b a n k I D N O : ) を示す。列 4 は、各ポリペプチドとその G e n B a n k 相同体との間の一致を表す確率スコアを示す。列 5 は、G w n B a n k 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って本明細書の一部とする。

#### 【0103】

表 3 は、本発明の各ポリペプチドの様々な構造的特徴を示す。列 1 および列 2 はそれぞれ、本発明の各ポリペプチドのポリペプチド配列識別番号 (S E Q I D N O : ) およびそれに対応するインサイトポリペプチド配列番号 (I n c y t e P o l y p e p t i d e I D) 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、G C G 配列分析ソフトウェアパッケージの M O T I F S プログラム (G e n e t i c s C o m p u t e r G r o u p , M a d i s o n W I) によって決定された、潜在的なリン酸化部位および潜在的なグリコシル化部位を示す。列 6 は、サイン (s i g n a t u r e) 配列、ドメイン、およびモチーフを含むアミノ酸残基を示す。列 7 は、タンパク質の構造／機能の分析のための分析方法を示し、該当箇所には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。

10

#### 【0104】

表 2 および表 3 は共に、本発明の各ポリペプチドの特性を一覧にしたものであって、これらの特性が請求するポリペプチドが分泌タンパク質であることを立証する。例えば、S E Q I D N O : 4 はヒトスクシニル C o A : 3 オキシ酸 C o A トランスフェラーゼ前駆体 (G e n B a n k I D g 1 5 1 9 0 5 2) と 7 2 % の同一性を有することが B a s i c L o c a l A l i g n m e n t S e a r c h T o o l (B L A S T) によって示された (表 2 を参照)。B L A S T の確率スコアは  $1.8 \times 10^{-198}$  である。これは、探しているポリペプチド配列アラインメントが偶然の一致により得られる確率を示す。S E Q I D N O : 4 はまた、コエンザイム A トランスフェラーゼドメインを有する。これは、隠れマルコフモデル (H M M) 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 P F A M 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 を参照)。B L I M P S 及び B L A S T 分析から得られたデータによって、S E Q I D N O : 4 がヒトスクシニル C o A : 3 オキシ酸 C o A トランスフェラーゼ等のコエンザイム A トランスフェラーゼであることが裏付けられた。S E Q I D N O : 1 - 3 及び S E Q I D N O : 5 - 1 4 も同様の方法で解析し、注釈を付けた。S E Q I D N O : 1 - 1 4 の解析のためのアルゴリズムおよびパラメータを表 7 に記載する。

20

30

#### 【0105】

表 4 に示されているように、本発明の完全長ポリヌクレオチド配列は、c D N A 配列、またはゲノム D N A 由来のコード (エキソン) 配列、或いはこれらの 2 種類の配列のあらゆる組み合わせを用いて組み立てた。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 (P o l y n u c l e o t i d e S E Q I D N O : ) およびそれに対応するインサイトポリヌクレオチドコンセンサス配列番号 (I n c y t e P o l y n u c l e o t i d e I D) を示す。列 3 は、塩基対における各ポリヌクレオチド配列の長さを示す。列 4 は、例えば、S E Q I D N O : 1 5 - 2 8 を同定するため、或いは S E Q I D N O : 1 5 - 2 8 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、c D N A 配列、ゲノム D N A から推定されるコード配列 (エキソン)、および／または c D N A およびゲノム D N A の両方からなる群に対応する識別番号を示す。これらの配列を用いて本発明の完全長ポリヌクレオチド配列を組み立てた。表 4 の列 6 および列 7 はそれぞれ、それぞれの完全長配列に対する列 5 の c D N A 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。

40

#### 【0106】

表 4 の列 5 に示されている識別番号は、具体的には、例えばインサイト c D N A およびそれらに対応する c D N A ライブラリの識別番号を示す。例えば、2 0 8 7 2 9 3 H 1 はイ

50

ンサイト cDNA 配列の識別番号であり、P A N C N O T 0 4 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インサイト cDNA は、プールされている cDNA ライブラリ（例えば S C K A 0 1 2 7 0 V 1）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組み立てに用いた G e n B a n k の cDNA すなわち E S T の識別番号（例えば g 6 7 5 3 5 3）の場合もある。または、列 5 の識別番号は、G e n s c a n 分析によって推定されるゲノム DNA のコード領域の場合もある。例えば、G N N . g 6 4 3 7 5 1 6 \_ 0 0 0 0 0 4 \_ 0 0 2 は G e n s c a n 推定コード配列の識別番号であり、g 6 4 3 7 5 1 6 は G e n s c a n が適用された配列の G e n b a n k 識別番号である。この G e n s c a n 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（実施例 4 を参照）。または、列 5 の識別番号は、  
“ e x o n - s t i t c h i n g ” アルゴリズムによって cDNA および G e n s c a n 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）。または、列 5 の識別番号は、  
“ e x o n - s t r e t c h i n g ” アルゴリズムによって cDNA および G e n s c a n 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、列 5 に示されている配列の範囲と重複するインサイト cDNA の範囲が得られ、最終的なコンセンサス配列が決定されるが、それに相当するインサイト cDNA の識別番号は示されていない。

#### 【0107】

表 5 は、インサイト cDNA 配列を用いて組み立てられたこれらの完全長ポリヌクレオチド配列が由来する代表的な cDNA ライブラリを示す。代表的な cDNA ライブラリとは、上記ポリヌクレオチド配列の組み立ておよび決定に用いられたインサイト cDNA 配列を最も多く含むインサイト cDNA ライブラリのことである。表 5 に示されている cDNA ライブラリを作製するために用いた組織およびベクターが表 6 に示されている。

#### 【0108】

本発明はまた、S E C P の変異体も含む。好適な S E C P の変異体は、S E C P 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 S E C P 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。

#### 【0109】

本発明はまた、S E C P 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、S E C P をコードする S E Q I D N O : 1 5 - 2 8 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を提供する。配列表に示した S E Q I D N O : 1 5 - 2 8 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、糖鎖の背骨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。

#### 【0110】

本発明はまた、S E C P をコードするポリヌクレオチド配列の変異体を含む。詳細には、このようなポリヌクレオチド配列の変異体は、S E C P 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 % ものポリヌクレオチド配列同一性を有する。本発明の特定の実施形態は、S E Q I D N O : 1 5 - 2 8 からなる群から選択された核酸配列と少なくとも 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 % ものポリヌクレオチド配列同一性を有する S E Q I D N O : 1 5 - 2 8 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提供する。上記したポリヌクレオチド変異体は何れも、S E C P の機能的或いは構造的特徴の少なくとも 1 つを有するアミノ酸配列をコードする。

#### 【0111】

当業者であれば理解し得ることであるが、遺伝暗号の縮重により作り出され得る S E C P 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。したがって本発明には、可能なコドン選択に基づいた組み合わせの選択によって作り出され得る可能なポリヌクレ

オチド配列の変異の全てが含まれ得る。これらの組み合わせは、天然のS E C P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。

#### 【0112】

S E C P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S E C P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S E C P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。特定のコードが宿主によって利用される頻度に基づいてコドンを選択して、ペプチドの発現が特定の真核細胞または原核宿主に発生する割合を高め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、S E C P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

#### 【0113】

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

#### 【0114】

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: 15-28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; and 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。

20

#### 【0115】

DNAシーケンシング方法は当分野で公知であり、本発明の何れの実施例もそのような方法を用いて実行可能である。或る方法では、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)、或いはELONGASE増幅システム (Life Technologies, Gaithersburg MD) にみられるような校正エキソヌクレアーゼとポリメラーゼとの組み合わせなどの酵素を用い得る。好ましくは、MICROLAB 2200液体転移システム (Hamilton, Reno, NV)、PTC 200 Thermal Cycler 200 (MJ Research, Watertown MA) 及びABI CATALYST 800 (Applied Biosystems) などの装置を用いて配列の準備を自動化する。次に、ABI 373或いは377 DNAシーケンシングシステム (Applied Biosystems)、MEGABACE 1000 DNAシーケンシングシステム (Molecular Dynamics, Sunnyvale CA) または当分野で公知の他の方法を用いてシーケンシングを行う。得られた配列を当分野で公知の様々なアルゴリズムを用いて分析する（例えば、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. を参照）。

30

40

#### 【0116】

50

当分野で公知の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し得る。例えば制限部位PCR法を利用する1つの方法では、一般的なプライマー及び入れ子プライマー(nested primer)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(例えば、Sarkar, G. (1993) PCR Methods Applic 2:318-322を参照)。逆PCR法を用いる別法では、広範な方向に伸長して環状化した鋳型から未知の配列を増幅するプライマーを用いる。この鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限断片に由来する(例えば、Triglia, T. ら(1988) Nucleic Acids Res 16:8186を参照)。キャプチャPCR法を用いる第3の方法は、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断片のPCR増幅を含む(例えば、Lagerstrom, M. ら(1991) PCR Methods Applic 1:111-119を参照)。この方法では、多数の制限酵素による消化及びライゲーションを用いて、PCRを行う前に未知の配列の領域の中に組換え二本鎖配列を挿入することが可能である。また、当分野で公知の別の方法を用いて未知の配列を得ることも可能である。(例えば、Parker, J. D. ら(1991) Nucleic Acids Res. 19:3055-3060を参照)。更に、PCR、ネスト化プライマー、PROMOTERFINDERライブラリ(Clontech, Palo Alto CA)を用いれば、ゲノムDNA内の歩行が可能である。この方法で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/エキソン接合部を探すのに有用である。全てのPCR法をベースにした方法では、プライマーは、市販のソフトウェア、例えばOLIGO 4.06 Primer Analysis software(National Biosciences, Plymouth MN)或いは別の好適なプログラムなどを用いて、長さが22~30ヌクレオチド、GC含量が50%以上、約68℃~72℃の温度で鋳型に対してアニーリングするよう設計される。

10

20

30

40

50

#### 【0117】

完全長のcDNAをスクリーニングする場合は、大きなcDNAを含むようにサイズが選択されたライブラリを用いるのが好ましい。更に、オリゴd(T)ライブラリが完全長のcDNAを産生できない場合は、遺伝子の5'領域を有する配列を含むものが多いランダムに初回抗原刺激を受けたライブラリが有用である。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。

#### 【0118】

市販の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産物のヌクレオチド配列のサイズの分析または確認が可能である。詳しくは、キャピラリーシークエンシングには、電気泳動による分離のための流動性ポリマー、4つの異なったヌクレオチドに特異的なレーザーで活性化される蛍光色素、及び放出された波長の検出に利用するCCDカメラを使用することが可能である。出力/光強度は、適切なソフトウェア(例えば、GENOTYPER及びSEQUENCE NAVIGATOR、Applied Biosystems)を用いて電気信号に変換され、サンプルのローディングからコンピュータ分析までのプロセス及び電子データ表示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場合もあるDNAの小片のシークエンシングに特に適している。

#### 【0119】

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起因して、実質的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配列を作成し、それをSECPの発現に利用することが可能である。

#### 【0120】

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み換え、種々

の目的のためにS E C Pをコードする配列を変化させることができる。種々の目的には、限定されるものではないが、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。ランダムな断片によるDNAの混合や遺伝子断片と合成オリゴヌクレオチドのPCR再組み立てを用いて、ヌクレオチド配列の組換えが可能である。例えば、オリゴヌクレオチド媒介性定方向突然変異誘発を利用して、新しい制限部位を生成する突然変異の導入、グリコシル化パターンの変更、コドン選択の変更、スプライスバリエーションの作製等が可能である。

#### 【0121】

本発明のヌクレオチドを、カリフォルニア州サンタクララのMaxygen社のMOLECULAR BREEDING（米国特許第5,837,458号、Chang, C. - 10  
C. ら（1999）Nat. Biotechnol. 17:793-797、Christians, F. C. ら（1999）Nat. Biotechnol. 17:259-264、Cramer, A. ら（1996）Nat. Biotechnol. 14:315-319に記載されている）などのDNAシャフリング技術を用いてシャフリングして、S E C P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S E C Pの生物学的特性を変更或いは改良することができる。DNAシャフリングは、PCR法による遺伝子断片の組換えで遺伝子変異体のライブラリを作製するプロセスである。次に、このライブラリを、目的の特性を有する遺伝子変異体を同定するために選択或いはスクリーニングする。これらの好ましい変異体をプールし、DNAシャフリング及び選択／スクリーニングを繰り返す。従って、人工的な育種及び急速な分子の進化によって多様な遺伝子が作られる。例えば、ランダムな位置に変異がある1つの遺伝子の断片を、目的の特性が最適化するまで、組換え及びスクリーニング、シャフリングを実施することもできる。別法では、所与の遺伝子の断片を、同種または異種の同一遺伝子ファミリーの相同な遺伝子の断片で組換え、それによってプロトコルに従った調節可能な方法で、多数の天然遺伝子の遺伝子多様性を最大にすることができる。

#### 【0122】

別の実施例によれば、S E C P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、Horn, T. ら（1980）Nucl. Acids Res. Symp. Ser. 22 30  
5-232を参照）。別法として、化学的方法を用いてS E C P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、ペプチド合成は種々の固相技術を用いて実行可能である（例えば、Creighton, T.（1984）Proteins. Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60、Roberge, J. Y. ら（1995）Science 269:202-204を参照）。また、合成の自動化は例えばABI 431Aペプチドシンセサイザー（Applied Biosystems）を用いて達成し得る。更にS E C P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。

#### 【0123】

このペプチドは、分離用高速液体クロマトグラフィー（例えば、Chiez, R. M. 及び F. Z. Regnier（1990）Methods Enzymol. 182:392-421を参照）を用いて実質的に精製可能である。合成されたペプチドの組成は、アミノ酸分析或いはシーケンシングにより確認することができる（例えば前出のCreighton, pp 28-53を参照）。

#### 【0124】

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 50

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SECPをコードするポリヌクレオチド配列におけるエンハンサー、構成型及び発現誘導型のプロモーター、5'及び3'の非翻訳領域などの調節配列が含まれる。このようなエレメントは、その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、SECPをコードする配列のより効果的な翻訳を達成することが可能である。このようなシグナルには、ATG開始コドン及びコザック配列などの近傍の配列が含まれる。SECPをコードする配列及びその開始コドン、上流の調節配列が好適な発現ベクターに挿入された場合は、更なる転写調節シグナルや翻訳調節シグナルは必要なくなるであろう。しかしながら、コー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ATG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 10  
含まれなければならない。外来性の翻訳要素及び開始コドンは、自然及び合成の様々なものから得ることが可能である。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である（例えば、Scharf, D. ら (1994) Results Probl. Cell Differ. 201-18-162. を参照）。

#### 【0125】

当業者に公知の方法を用いて、SECPをコードする配列、好適な転写及び翻訳調節エレメントを含む発現ベクターを作製することが可能である。これらの方法には、in vitro組換えDNA技術、合成技術及びin vivo遺伝子組換え技術が含まれる（例えば、Sambrook, J. ら. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, 4章及び8章, 及び16-17章及び Ausubel, F. M. ら. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9章及び13章1-4章を参照）。 20

#### 【0126】

種々の発現ベクター／宿主系を利用して、SECPをコードする配列の保持及び発現が可能である。これらには、限定するものではないが、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れた細菌などの微生物や、酵母菌発現ベクターで形質転換された酵母菌や、ウイルス発現ベクター（例えば、バキュロウイルス 30  
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；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）で形質転換された植物細胞系や、動物細胞系などが含まれる（例えば、前出のSambrook、前出のAusubel、Van Heeke, G. 及び S. M. Schuster (1989) J. Biol. Chem. 264:5503-5509、Engelhard, E. K. ら (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227、Sandig, V. ら (1996) Hum. Gene Ther. 7:1937-1945、Takamatsu, N. (1987) EMBOJ. 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、Harrington, J. J. ら (1997) Nat. Genet. 15:345-355を参照）。レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ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. Natl. Acad. Sci. USA 90(13):6340-6344、Buller, R. M. ら (1985) Nat 40

10

20

30

40

50

ure 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。

#### 【0127】

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。ベクターの多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、これらのベクターを用いて、クローニングされた配列の in vitro での転写、ジデオキシンスクリーニング、ヘルパーファージによる一本鎖の救出、入れ子状態の欠失を作り出すことが可能である (例えば、Van Heeke, G. 及び S. M. Schuster (1989) J. Biol. Chem. 264 : 5503-5509. 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、強力に発現を誘発するT5またはT7バクテリオファージプロモーターを含むベクターを使用できる。

#### 【0128】

酵母発現系を用いてSECPを産出し得る。α因子やアルコールオキシダーゼやPGHプロモーターなどの構成型或いは誘導型のプロモーターを含む多種のベクターが、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、発現したタンパク質の分泌か細胞内への保持のどちらかを誘導し、安定した増殖のために宿主ゲノムの中に外来配列を組み込む (例えば、Ausubel, 1995, 前出、Bitter, G. A. ら (1987) Methods Enzymol. 153 : 516-544、及び Scorer, C. A. ら (1994) Bio/Technology 121-181-184. を参照)。

#### 【0129】

植物系を用いてSECPを発現することも可能である。SECPをコードする配列の転写は、例えば、CaMV由来の35S及び19Sプロモーターなどのウイルスプロモーターが単独で、或いはTMV (例えば、Coruzzi, G. ら (1984) EMBO J. 3 : 1671-1680 ; Broglie, R. ら (1984) Science 224 : 838-843 ; および Winter, J. ら (1991) Results Probl. Cell Differ. 17 : 85-105を参照) 由来のオメガリーダー配列と組み合わせて促進される。これらの作製物は、直接のDNA形質転換或いは病原体を介したトランスフェクションによって、植物細胞の中に導入可能である。(例えば、The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill NY, pp. 191-196を参照)。

#### 【0130】

哺乳動物細胞では、多種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. (1984) Proc. Natl. Acad. Sci. 81 : 3655-3659を参照)。さらに、ラウス肉腫ウイルス (RSV) エンハンサーなど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ることが可能である。タンパク質を高レベルで

発現させるために、SV40またはEBVを基にしたベクターを用いることが可能である。

#### 【0131】

ヒト人工染色体(HAC)を用いて、プラスミドで発現しそれに含まれているものより大きなDNAの断片を供給可能である。治療のために約6kb~10MbのHACs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。(例えば、Harrington, J. J. ら (1997) Nat Genet. 15:345-355. を参照)。

#### 【0132】

哺乳動物系の組換えタンパク質を長期間産生するためには、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、同一または別のベクター上の選択マーカー遺伝子を含む。ベクターの導入の後、細胞を選択培地に移す前に、強化培地で約1~2日の間増殖させる。選択マーカーの目的は選択的な媒介物に対する抵抗性を与えるとともに、その存在により導入された配列を確実に発現する細胞の増殖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、その細胞型に好適な組織培養技術を用いて増殖可能である。

#### 【0133】

任意の数の選択系を用いて、形質転換された細胞系を回収することが可能である。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<sup>-</sup>またはap<sup>-</sup>細胞において使用される。(例えば、Wigler, M. ら (1977) Cell 11:223-232; 及びLowy, I. ら (1980) Cell 22:817-823を参照)。また代謝拮抗物質、抗生物質或いは除草剤への耐性を選択のベースとして用いることができる。例えばdhfrはメトトレキセートに対する耐性を与え、neoはアミノグリコシッドネオマイシン及びG-418に対する耐性を与え、als或いはpatはクロルスルフロン(chlorosulfuron)、ホスフィノトリシンアセチルトランスフェラーゼ(phosphinotricin acetyltransferase)に対する耐性を与える(例えば、Wigler, M. ら. (1980) Proc. Natl. Acad. Sci. 77:3567-3570; Colbere-Garapin, F. ら (1981) J. Mol. Biol. 150:1-14を参照)。さらに選択に利用できる遺伝子、例えば、代謝のために細胞が必要なものを変えるtrpB及びhisDが文献に記載されている(例えば、Hartman, S. C. 及びR. C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-51を参照)。アニトシアニン、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GUS、ルシフェラーゼ及びその基質ルシフェリンなどの可視マーカーが用いられる。緑色蛍光タンパク質(GFP) (Clontech, Palo Alto, CA) も使用できる。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(例えば、Rhodes, C. A. ら (1995) Methods Mol. Biol. 55:121-131を参照)。

#### 【0134】

マーカー遺伝子の発現の存在／非存在によって目的の遺伝子の存在が示されても、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特定可能である。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一列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。

【0135】

一般に、SECPをコードする核酸配列を含み、SECPを発現する宿主細胞は、当業者に公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。

【0136】

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SECP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、酵素に結合したイムノソルベントアッセイ (ELISA)、ラジオイムノアッセイ (RIA)、蛍光標示式細胞分取器 (FACS) などがある。SECP上の2つの非干渉エピトープに反応するモノクローナル抗体を用いた、2部位のモノクローナルベースイムノアッセイ (two-site, monoclonal-based immunoassay) が好ましいが、競合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その他のアッセイは、当分野では十分に知られている。(例えば、Hampton, R. 1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Sect. IV, Coligan, J. E. 1990 Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York, NY、及びPound, J. D. (1990) Immunochemical Protocols, Humana Press, Totowa NJ)。

【0137】

種々の標識技術及び結合技術が当業者には周知であり、様々な核酸アッセイおよびアミノ酸アッセイに用いられ得る。SECPをコードするポリヌクレオチドに関連する配列を検出するための、標識されたハイブリダイゼーションプローブ或いはPCRプローブを生成する方法には、オリゴ標識化、ニックトランスレーション、末端標識化、または標識されたヌクレオチドを用いるPCR増幅が含まれる。別法として、SECPをコードする配列、またはその任意の断片を、mRNAプローブ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。当分野では周知であり市販されているこのようなベクターを、T7, T3, またはSP6などの好適なRNAポリメラーゼ及び標識されたヌクレオチドの追加によって、in vitroでのRNAプローブの合成に用いることができる。これらの方法は、例えば、Amersham Pharmacia Biotech及びPromega (ウィスコンシン州マディソン)、U. S. Biochemical Corp (オハイオ州クリーブランド) が市販する種々のキットを用いて行うことができる。容易な検出の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、コファクター、インヒビター、磁気粒子、及び放射性核種、酵素、蛍光剤、化学発光剤、色素産生剤などが含まれる。

【0138】

SECPをコードするヌクレオチド配列で形質転換された宿主細胞は、細胞培地でのこのタンパク質の発現及び回収に好適な条件下で培養される。形質転換された細胞から産生されたタンパク質が分泌されるか細胞内に留まるかは、使用されるその配列及び／またはそのベクターによる。SECPをコードするポリヌクレオチドを含む発現ベクターは、原核細胞膜及び真核細胞膜を透過するSECPの分泌を誘導するシグナル配列を含むように設計できることは、当業者であれば理解されよう。

【0139】

更に、挿入した配列の発現調節能力または発現したタンパク質を所望の形にプロセッシングする能力によって宿主細胞株が選択される。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 (lipidation)、及びアシル化が含まれるが、これらに限定されるものではない。タンパク質の「プレプロ」または「プロ」形を切断する翻訳後のプロセッシングを利用して、標的タンパク質、折りたたみ及び／または活性を特定することが可能である。翻訳後の活性のための特定の細胞装置及び特徴のある機構をもつ種類の宿主細胞 (例えば、CHO、HeLa、MDCK、MEK293、WI38) が American Type Culture Collection (バージニア州マナッサスの ATCC) より入手可能であり、外来のタンパク質の正しい修飾及びプロセッシングを確実にするために選択される。

#### 【0140】

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、或い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が、市販の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。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-mc、赤血球凝集素 (HA) が含まれるが、これらに限定されるものではない。GST及びMBP、Trx、CBP、6-Hisによって、固定されたグルタチオン、マルトース、フェニルアルシン酸化物 (phenylarsine oxide)、カルモジュリン、金属キレート樹脂のそれぞれで同族の融合タンパク質の精製が可能となる。FLAG、c-mc、及び赤血球凝集素 (HA) によって、これらのエピトープ標識を特異的に認識する市販のモノクローナル抗体及びポリクローナル抗体を用いた融合タンパク質の免疫親和性の精製ができる。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と精製の方法は、Ausubel. (1995、前出 ch 10) に記載されている。市販されている様々なキットを用いて、融合タンパク質の発現及び精製を促進できる。

#### 【0141】

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またはT3、SP6プロモーターと機能的に結合したタンパク質をコードする配列の転写と翻訳をつなげる。翻訳は、例えば、<sup>35</sup>Sメチオニンである放射能標識されたアミノ酸前駆体の存在の下で起こる。

#### 【0142】

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。

#### 【0143】

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. ら (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、大腸菌からの細胞が含まれる。SECPを発現する細胞またはSECPを含有する細胞膜断片を

試験化合物と接触させて、S E C P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。

#### 【0144】

或るアッセイは、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、フルオロフォア、放射性同位体、酵素抱合体またはその他の検出可能な標識により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S E C Pと結合させる過程と、S E C Pとこの化合物との結合を検出する過程を含み得る。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、細胞遊離剤、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。

#### 【0145】

本発明のS E C Pまたはその断片を用いて、S E C P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、S E C Pが少なくとも1つの試験化合物と結合する、S E C P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S E C Pの活性が試験化合物非存在下でのS E C Pの活性と比較する。試験化合物の存在下でのS E C Pの活性の変化は、S E C P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S E C Pの活性に適した条件下でS E C Pを含む *i n v i t r o* または細胞遊離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、S E C P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つ及び複数の試験化合物をスクリーニングすることができる。

#### 【0146】

別の実施例では、胚性幹細胞（E S細胞）における相同組換えを用いて動物モデル系内で、S E C P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）。例えば129/SvJ細胞株等のマウスE S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E S細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（*n e o* : C a p e c c h i, M. R. (1989) S c i e n c e 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、C r e - l o x P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（M a r t h, J. D. (1996) C l i n. I n v e s t. 97:1999-2002、W a g n e r, K. U. ら (1997) N u c l e i c A c i d s R e s. 25:4323-4330）。形質転換したE S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、潜在的な治療薬や毒性薬剤で検査することができる。

#### 【0147】

S E C Pをコードするポリヌクレオチドを *i n v i t r o* でヒト胚盤胞由来のE S細胞において操作することが可能である。ヒトE S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（T h o m s o n, J. A. ら (1998) S c i e n c e 282:1145-1147）。

#### 【0148】

S E C P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、潜在的な医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばSECPを乳汁内に分泌するなどSECPを過剰に発現する哺乳動物近交系は、便利なタンパク質源となり得る(Janne, J.ら(1998) Biotechnol. Annu. Rev. 4: 55-74)。

#### 【0149】

(治療)

SECPのある領域と分泌タンパク質のある領域との間には、例えば配列及びモチーフと関連して化学的及び構造的類似性が存在する。更に、SECPの発現は、内分泌、生殖、筋、腫瘍、大動脈平滑筋、脳及び精巣組織、並びに成長及び発達に関与する組織に密接に関与している。従って、SECPは、細胞増殖異常、自己免疫／炎症性疾患、心血管疾患、神経疾患及び発達障害においてある役割を果たすと考えられる。SECPの発現若しくは活性の増大に関連する疾患の治療においては、SECPの発現または活性を低下させることが望ましい。また、SECPの発現または活性の低下に関連する疾患の治療においては、SECPの発現または活性を増大させることが望ましい。

#### 【0150】

従って、一実施例において、SECPの発現または活性の低下に関連した疾患の治療または予防のために、患者にSECPまたはその断片や誘導体を投与することが可能である。限定するものではないが、このような疾患には細胞増殖異常、自己免疫／炎症性疾患、心血管疾患、神経疾患及び発達障害が含まれ、細胞増殖異常の中には日光性角化症及び動脈硬化、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、脾臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌などが含まれ、自己免疫／炎症性疾患の中には、後天性免疫不全症候群(AIDS)及びアジソン病、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(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的リンパ球減少症、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、癌の合併症、血液透析及び体外循環による炎症、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、外傷による炎症が含まれ、心血管疾患の中には、動静脈瘤、アテローム性硬化症、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)、冠状動脈バイパス術移植手術(coronary artery bypass graft surgery)、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、全身性紅斑性狼瘡の心内膜炎、カルチノイド心疾患、心筋症、心筋炎、心膜炎、腫瘍性心疾患、先天性心臓疾患、心臓移植の合併症、先天性肺異常(congenital lung anomalies)、

肺拡張不全、肺うっ血及び肺水腫、肺動脈塞栓症、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 (restrictive pulmonary disease)、慢性閉塞性肺疾患、肺気腫、慢性気管支炎、気管支喘息、細気管支拡張症、細菌性肺炎、ウイルス性肺炎及びマイコプラズマ肺炎、肺膿瘍、肺結核、びまん性間質性疾患 (diffuse interstitial diseases)、塵肺症、サルコイド症、特発性肺繊維症 (idiopathic pulmonary fibrosis)、剥離性間質性肺炎、過敏症肺炎 (hypersensitivity pneumonitis)、肺好酸球増加閉塞性細気管支炎－器質性肺炎 (pulmonary eosinophilia bronchiolitis obliterans－organizing pneumonia)、びまん性肺出血症候群 (diffuse pulmonary hemorrhage syndromes)、グッドパスチャー症候群、特発性肺血鉄症、肺併発膠原血管病 (pulmonary involvement in collagen－vascular disorders)、肺胞たんぱく症、肺腫瘍、炎症性及び非炎症性胸水 (inflammatory and noninflammatory pleural effusions)、気胸症、胸膜腫瘍、薬物による肺疾患 (drug－induced lung disease)、放射線による肺疾患 (radiation－induced lung disease) 及び肺移植の合併症が含まれ、神経疾患の中には、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、クロイツフェルト－ヤコブ病及びゲルストマン－ストロイスラー－シャインカー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 (cerebellar retinal hemangioblastomatosis)、脳3叉神経血管症候群、ダウン症候群を含む精神遅滞及びその他の中枢神経系の発達障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄疾患、筋ジストロフィ及びその他の神経筋障害、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、不安性の障害や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、遅発性ジスキネジア、ジストニー、分裂病性精神障害、帯状疱疹後神経痛、トゥレット病、進行性核上麻痺、皮質基部変性 (corticobasal degeneration) 及び家族性前頭側頭性健忘症が含まれ、発達障害の中には、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、癲癇、性腺形成異常、WAGR症候群 (ウィルムス腫瘍、無虹彩症、尿生殖器異常及び精神薄弱)、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－トゥス病や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 舞蹈病や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。

#### 【0151】

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。

#### 【0152】

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。

#### 【0153】

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 S E C 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、S E C 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。

#### 【0154】

更なる実施例では、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、患者に S E C 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／炎症性疾患、心血管疾患、神経疾患及び発達障害が含まれる。一実施態様では、S E C 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 S E C P を発現する細胞または組織に薬剤を運ぶターゲッティング或いは運搬機構として間接的に用いられ得る。

10

#### 【0155】

別の実施例では、限定するものではないが上記した疾患を含む 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。

#### 【0156】

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、相補的な配列、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。当業者は、従来の医薬原理にしたがって併用療法で用いる好適な治療薬を選択可能である。治療薬との組み合わせにより、上に列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。この方法を用いて少ない量の各薬剤で医薬効果をあげることが可能であり、広範囲な副作用の可能性を低減し得る。

20

#### 【0157】

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 を用いて抗体を産生したり、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を同定したりすることが可能である。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 フラグメント、及び F a b 発現ライブラリによって作られたフラグメントが含まれる。但し、これらに限定されるものではない。治療用には、中和抗体（即ち、二量体の形成を阻害するもの）が特に好ましい。

30

#### 【0158】

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。このようなアジュバントにはフロイントアジュバント、水酸化アルミニウムなどのミネラルゲルアジュバント、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、及びジニトロフェノールなどの界面活性剤が含まれるが、これらに限定されるものではない。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－バルブムが特に好ましい。

40

#### 【0159】

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或いはペプチド、またはそれらの断片は、天然のタンパク質のアミノ酸配列の一部と同一であることが望ましい。S E C P アミノ酸の短いストレッチは、K L H など他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。

#### 【0160】

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。これらの技術には、ハイブリド

50

ーマ技術、ヒトB細胞ハイブリドーマ技術、及びEBV-ハイブリドーマ技術が含まれるが、これらに限定されるものではない（Kohler, G. ら. (1975) *Nature* 256:495-497、Kozbor, D. ら. (1985) *J. Immunol. Methods* 81:31-42、Cote, R. J. ら. (1983) *Proc. Natl. Acad. Sci.* 80:2026-2030、Cole, S. P. ら. (1984) *Mol. Cell Biol.* 62:109-120等を参照）。

#### 【0161】

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（Morrison, S. L. ら. (1984) *Proc. Natl. Acad. Sci.* 81:6851-6855、Neuberger, M. S. ら. (1984) *Nature* 312:604-608、Takeda, S. ら. (1985) *Nature* 314:452-454等を参照）。別法では、当分野で公知の方法を用いて、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。関連する特異性を備えるが別のイディオタイプの組成の抗体は、ランダムな組み合わせの免疫グロブリンライブラリから鎖混合によって生成することもできる（Burton D. R. (1991) *Proc. Natl. Acad. Sci.* 88:10134-10137等を参照）。

#### 【0162】

抗体は、リンパ球集団の中の in vivo 産生を誘発することによって、また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示されているような、高度に特異的な結合試薬のパネルをスクリーニングすることによって、得ることもできる（Orlandi, R. ら. (1989) *Proc. Natl. Acad. Sci.* 86:3833-3837、Winter, G. ら. (1991) *Nature* 349:293-299等を参照）。

#### 【0163】

SECPに対する特異的な結合部位を含む抗体も得ることができる。例えば、限定するものではないが、抗体分子のペプシン消化によって生成される  $F(ab')_2$  断片、及び  $F(ab')_2$  断片のジスルフィド架橋を減少させて生成される Fab 断片が含まれる。別法では、所望の特異性を有するモノクローナル Fab 断片の迅速且つ容易な同定が可能となるように、Fab 発現ライブラリを作製することもできる（Huse, W. D. ら. (1989) *Science* 254:1275-1281等を参照）。

#### 【0164】

種々のイムノアッセイをスクリーニングに用い、所望の特異性を有する抗体を同定する。確立された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる競合的な結合アッセイ、または免疫放射線アッセイのための様々なプロトコルは、当分野で公知である。通常このようなイムノアッセイでは、SECPとその特異的抗体との間で形成された複合体の測定が含まれる。2つの非干渉性SECPエピトープに対して反応性のモノクローナル抗体を用いる、2部位モノクローナルベースのイムノアッセイが一般に利用されるが、競合的結合アッセイも利用することができる（前出の Pound）。

#### 【0165】

ラジオイムノアッセイ技術と共にスキッチャード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医薬の  $K_a$  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医薬の  $K_a$  は、親和性の真の測定値を表す。 $K_a$  値が  $10^9 \sim 10^{12}$  L/mol の高親和性抗体医薬は、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1

10

20

30

40

50

$0.6 \sim 10^7$  L/mol の低親和性抗体医薬は、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)。

#### 【0166】

ある下流での適用におけるこのような医薬品の品質及び適性を調べるために、ポリクローナル抗体医薬の抗体価及び結合活性を更に評価する。例えば、少なくとも  $1 \sim 2$  mg/ml の特異的な抗体、好ましくは  $5 \sim 10$  mg/ml の特異的な抗体を含むポリクローナル抗体医薬は一般に、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。様々な適用例における抗体の特異性及び抗体価、結合活性、抗体の品質や使用法の指針は一般に入手可能である (例えば、前出の Catty 及び同 Coligan らを参照)。

#### 【0167】

本発明の別の実施例では、SECP をコードするポリヌクレオチド、またはその任意の断片や相補配列が、治療目的で使用する事ができる。ある実施態様では、SECP をコードする遺伝子のコーディング領域や調節領域に相補的な配列やアンチセンス分子 (DNA 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、SECP 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 (Agrawal, S. ら (1996) Antisense Therapeutics, Humana Press Inc., Totowa NJ 等を参照)。

#### 【0168】

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送達することができる (Slater, J. E. ら (1998) J. Allergy Clin. Immunol. 102 (3): 469-475、Scanlon, K. J. ら (1995) 9 (13): 1288-1296. 等を参照)。また、アンチセンス配列は、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 (Miller, A. D. (1990) Blood 76: 271、前出の Ausubel、Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63 (3): 323-347 等を参照)。その他の遺伝子送達機構には、リポソーム系、人工的なウイルスエンベロープ、及び当分野で周知のその他の系が含まれる (Rossi, J. J. (1995) Br. Med. Bull. 51 (1): 217-225、Boado, R. J. ら (1998) J. Pharm. Sci. 87 (11): 1308-1315、Morris, M. C. ら (1997) Nucleic Acids Res. 25 (14): 2730-2736. 等を参照)。

#### 【0169】

本発明の別の実施例では、SECP をコードするポリヌクレオチドを、体細胞若しくは生殖細胞の遺伝子治療に用いることが可能である。遺伝子治療は、(i) 遺伝子欠損症 (例えば、X 染色体連鎖遺伝 (Cavazzana-Calvo, M. ら (2000) Science 288: 669-672) によって特徴づけられる重度の複合型免疫欠損 (SCID) - X1)、遺伝性アデノシンデアミナーゼ (ADA) 欠損症 (Blaese, R. M. ら (1995) Science 270: 475-480、Bordignon, C. ら (1995) Science 270: 470-475) に関連する重度の複合型免疫欠損、嚢胞性繊維症 (Zabner, J. ら (1993) Ce

1175:207-216: Crystal, R. G. ら (1995) Hum. Gene Therapy 6:643-666; Crystal, R. G. ら. (1995) Hum. Gene Therapy 6:667-703)、サラセミア (thalassaemia)、家族性高コレステロール血症、第VIII因子若しくは第IX因子欠損による血友病 (Crystal, 35 R. G. (1995) Science 270:404-410; Verma, I. M. and Somia, N. (1997) Nature 389:239-242) を治療したり、(ii) 条件的致死性遺伝子産物 (例えば、細胞増殖の制御不能による癌の場合) を発現させたり、及び (iii) 細胞内の寄生虫 (例えば、ヒト免疫不全ウイルス (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. ら (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399) や、B型若しくはC型肝炎ウイルス (HBV、HCV)、Candida albicans 及び Paracoccidioides brasiliensis 等の真菌寄生虫、Plasmodium falciparum 及び Trypanosoma cruzi 等の原虫寄生体) に対する防御機能を有するタンパク質を発現させて行うことができる。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。

#### 【0170】

本発明の更なる実施例では、SECPの欠損による疾患や異常症は、SECPをコードする哺乳動物発現ベクターを作製して、これらのベクターを機械的手段によってSECP欠損細胞に導入することによって治療する。in vivo 或いは ex vivo の細胞に用いる機械的な導入技術には、(i) 個々の細胞内へのDNAのマイクロインジェクション、(ii) 金粒子の打ち込み、(iii) リポソーム仲介性トランスフェクション、(iv) 受容体仲介性遺伝子導入、及び (v) DNAトランスポソン (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. Recipon (1998) Curr. Opin. Biotechnol. 9:445-450) の使用が含まれる。

#### 【0171】

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)、PTET-OFF、PTET-ON、PTRE2、PTRE2-LUC、PTK-HYG (Clontech, Palo Alto CA) が含まれる。SECPを発現させるために、(i) 恒常的に活性なプロモーター (例えば、サイトメガロウイルス (CMV)、ラウス肉腫ウイルス (RSV)、SV40ウイルス、チミジンキナーゼ (TK)、若しくはβ-アクチン遺伝子等)、(ii) 誘導性プロモーター (例えば、市販されているT-Rexプラスミド (Invitrogen) に含まれている、テトラサイクリン調節性プロモーター (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. U. S. A. 89:5547-5551; Gossen, M. ら (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。

#### 【0172】

市販のリポソーム形質転換キット（例えばInvitrogen社から入手可能なPERFECT LIPID及びTRANSFECTATION 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である。別法では、リン酸カルシウム法（Graham, F. L. and 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。初代細胞にDNAを導入するためには、これらの標準化された哺乳動物トランスフェクションプロトコルを変更する必要がある。

10

#### 【0173】

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社から入手可能であり、公表データ(Riviere, I. ら. (1995) Proc. Natl. Acad. Sci. U. S. A. 92: 6733-6737)に基づいて、上記データを引用することをもって本明細書の一部とする。このベクターは、VSVg(Armentano, D. ら (1987) J. Virol. 61: 1647-1650; Bender, M. A. ら (1987) J. Virol. 61: 1639-1646; Adam, M. A. and 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)等の乱交雑エンベロープタンパク質若しくは標的細胞上の受容体に対する親和性を有するエンベロープ遺伝子を発現する好適なベクター産生細胞系(VPCL)において増殖される。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、ある細胞集団（例えば、CD4<sup>+</sup>T細胞）の形質導入、並びに形質導入した細胞を患者に戻す方法は、遺伝子治療の分野では周知であり、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)。

20

30

40

#### 【0174】

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は当分野では周知である。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中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が、米国特許第5, 707, 618号(「Ade

50

novirus 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また、Antinozzi, P. A. ら (1999) Annu. Rev. Nutr. 19: 511-544及びVerma, I. M. and N. Somia (1997) Nature 18: 389: 239-242を参照し、引用することをもって本明細書の一部とする。

#### 【0175】

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分野では周知である。複製適格性の単純疱疹ウイルス(HSV)I型系のベクターは、霊長類の眼にレポーター遺伝子を送達するために用いられてきた(Liu, X. ら (1999) Exp. Eye Res. 169: 385-395)。HSV-1ウイルスベクターの作製は、DeLucaに付与された米国特許第5,804,413号(「Herpes simplex virus swains for gene transfer」)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。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コントロールの下で、細胞に導入される少なくとも1つの内在性遺伝子を含むゲノムからなる組換えHSV d92についての記載がある。また上記特許には、ICP4、ICP27及びICP22のために除去される組換えHSV株の作製及び使用方法が開示されている。HSVベクターについては、Goins, W. F. ら (1999) J. Virol. 73: 519-532 and Xu, H. ら (1994) Dev. Biol. 163: 152-161を参照し、引用することをもって本明細書の一部とする。クローニングされたヘルペスウイルス配列の操作や、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トランスフェクトした後の組換えウイルスの継代、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分野で周知の技術である。

#### 【0176】

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SECPをコードするポリヌクレオチドを標的細胞に送達する。プロトタイプのαウイルスであるセムリキ森林熱ウイルス(Semliki Forest Virus, SFV)の生物学的な研究が広範に行われ、遺伝子伝達ベクター(gene transfer vector)がSFVゲノムに基づいていることが分かった(Garoff, H. and K. -J. Li (1998) Cun. Opin. Biotech. 9: 464-469)。αウイルスRNAの複製中に、通常はウイルスカプシドタンパク質をコードするサブゲノムRNAが作り出される。このサブゲノムRNAが完全長のゲノムRNAより高いレベルで複製されるため、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対してカプシドタンパク質が過剰に産生される。同様に、SECP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SECPをコードするRNAが産生され、高いレベルでSECPが合成される。通常はαウイルス感染は2~3日以内の細胞溶解に関係するが、シンドビスウイルス(SIN)の変異体を有するハムスターの正常な腎細胞(BHK-21)の持続的な感染を確立する能力は、αウイルスの溶解性の複製が遺伝子治療に適用できるように好適に変更することが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SECPを導入することできる。ある集団における細胞のサブセットの特定の形質導入には、形質導入する前に細胞のソーティングを必要とする場合がある。αウイルスの感染性cDNAクローンの操作、αウイルスcDNA及びRNAのトランスフェクション、並びにαウイルスの感染方法は当分野で公知である

。

## 【0177】

例えば開始部位から約-10から約+10までの転写開始部位に由来するオリゴヌクレオチドを用いて、遺伝子の発現を阻害することが可能である。同様に、三重らせん塩基対合法を用いて阻害することができる。三重らせん構造は、二重らせんがポリメラーゼ、転写因子、または調節分子の結合のために十分に広がるのを阻止するため有用である。三重鎖DNAを用いる最近の治療の進歩は文献に記載されている(Gee, J. E. ら. (1994) In: Huber, B. E. 及び B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NY 等を参照)。相補的な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できる。

10

## 【0178】

酵索性RNA分子であるリボザイムは、RNAの特異的切断を触媒するために用いることができる。リボザイム作用の機構には、相補的な標的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が含まれ、ヌクレオチド鎖切断が続く。例えば、SECPをコードする配列のヌクレオチド鎖切断を、特異的且つ効果的に触媒する組換え型のハンマーヘッド型リボザイム分子が含まれる。

## 【0179】

任意の潜在的RNA標的内の特異的なリボザイム切断部位が、後続の配列GUA、GUU、及びGUCを含むリボザイム切断部位に対して、標的分子をスキャンニングすることによって初めに同定される。一度同定されると、切断部位を含む標的遺伝子の領域に対応する15個から20個のリボヌクレオチドの短いRNA配列を、オリゴヌクレオチドの機能を不全にする二次的な構造の特徴について評価することが可能である。候補標的の適合性も、リボヌクレアーゼ保護アッセイを用いて、相補的なオリゴヌクレオチドとのハイブリダイゼーションの容易性をテストすることによって評価することが可能である。

20

## 【0180】

本発明の相補的なリボ核酸分子及びリボザイムは、核酸分子の合成のための当分野で公知の任意の方法を用いて作製することができる。これらの方法には、固相ホスホラミダイト化合物などのオリゴヌクレオチドを化学的に合成する方法が含まれる。別法では、RNA分子が*in vitro*及び*in vivo*でSECPをコードするDNA配列の転写によって生成され得る、このようなDNA配列はT7またはSP6等の好適なRNAポリメラーゼプロモータを用いて、種々のベクターの中に組み入れることが可能である。別法では、相補的なRNAを構成的または誘導的に合成するこれらのcDNA作製物は、細胞株、細胞、または組織の中に導入することができる。

30

## 【0181】

RNA分子を修飾することによって、細胞内の安定性を高め、半減期を長くすることができる。可能な修飾には、分子の5'及び/または3'端部でのフランキング配列の追加、または分子のバックボーン内のホスホジエステル結合に代えてホスホロチオネートまたは2' Oメチルを用いる修飾が含まれるが、これらに限定されるものではない。PNAの生成に固有のこの概念は、内在性の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、及びウリジンのアセチル、メチル、チオール及び同様の修飾形態だけでなく、イノシン、キューエオシン(queosine)及びワイブトシン(wybutosine)などの従来のものでない塩基を含めることによって、これらの分子の全体に拡大することができる。

40

## 【0182】

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。特定のポリヌクレオチドの発現の変化に有効な化合物には、限定するものではないが、特定のポリヌクレオチド配列と相互作用可能な非高分子化学物質、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

50

成オリゴヌクレオチド、転写因子やその他のポリペプチド転写調節因子が含まれる。有効な化合物は、ポリヌクレオチド発現のインヒビター或いはエンハンサーとして作用し、ポリヌクレオチドの発現を変化させ得る。従って、SECPの発現または活性の増加に関連する疾患の治療においては、SECPをコードするポリヌクレオチドの発現を特異的に阻害する化合物が治療上有用であり、SECPの発現または活性の低下に関連する疾患の治療においては、SECPをコードするポリヌクレオチドの発現を特異的に促進する化合物が治療上有用であり得る。

#### 【0183】

特定のポリヌクレオチドの発現の変化の有効性を調べるために、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することができる。試験化合物は、有効な化合物の化学修飾を含む当分野で周知の任意の方法で得ることができる。このような方法は、ポリヌクレオチドの発現を変化させる場合、一般に市販されている或い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、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いて化合物を合理的にデザインする場合、更に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。SECPをコードするポリヌクレオチドを含むサンプルは、少なくとも1つの試験化合物に曝露して得る。サンプルには、例えば無傷細胞、透過化処理した細胞、in vitro細胞遊離系または再構成生化学系が含まれ得る。SECPをコードするポリヌクレオチドの発現における変化は、当分野で周知の任意の方法でアッセイする。通常、SECPをコードするポリヌクレオチドの配列に相補的なヌクレオチド配列を有するプローブを用いたハイブリダイゼーションにより、特定の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の収量を定量し、その値が1或いは複数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おける基準となり得る。試験化合物に曝露されるポリヌクレオチドの発現の変化が検出される場合は、ポリヌクレオチドの発現の変化に試験化合物が有効であることを示している。特定のポリヌクレオチドの発現の変化に有効な化合物を調べるために、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 細胞等のヒト細胞株 (Clarke, M. L. ら (2000) *Biochem. Biophys. Res. Commun.* 268: 8-13) を用いてスクリーニングする。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性を調べるための、各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、及び修飾オリゴヌクレオチド) の組み合わせライブラリのスクリーニングを含む (Bruce, T. W. ら (1997) 米国特許第5, 686, 242号、Bruce, T. W. ら (2000) 米国特許第6, 022, 691号)。

#### 【0184】

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でき、in vivo、in vitro、及び ex vivo での使用に等しく適している。ex vivo での治療の場合、患者から採取された肝細胞の中にベクターを導入して、自家移植で同じ患者に戻すためにクローニング増殖される。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運搬は、当分野で周知の方法を用いて実行することができる (例えば、Goldman, C. K. ら (1997) *Nature Biotechnology* 15: 462-466 を参照)。

#### 【0185】

上記したいかなる治療方法も、例えば、ヒト、イヌ、ネコ、ウシ、ウマ、ウサギ及びサルなどの哺乳動物を含む、治療が必要な全ての被験者に適用できる。

#### 【0186】

本発明の追加実施例は、上記した全ての治療効果のために、医学上許容される賦形剤と共に処方された活性成分を通常含むような組成物の投与に関連する。賦形剤には、砂糖、でんぶん、セルロース、ガム、タンパク質などがある。種々の処方については公知であり、

Remington's Pharmaceutical Sciences (Mack Publishing, Easton PA) の最新版に詳細な記載がある。このような組成物は、SECP、SECPの抗体、SECPの擬態、アゴニスト、アンタゴニストまたはインヒビターを含み得る。

#### 【0187】

本発明に用いられる組成物は、様々な経路を用いて投与するが可能である。この経路には、経口、静脈内、筋肉内、動脈内、骨髓内、クモ膜下、心室内、経皮、皮下、腹腔内、鼻腔内、腸内、局所、舌下、または直腸が含まれるがこれらに限定されるものではない。

#### 【0188】

肺投与用の組成物は、液状または乾燥粉末状に調製することができる。このような組成物は通常、患者が吸入する直前にエアロゾル化する。小分子（例えば従来の低分子量の有機薬剤）の場合には、速効製剤のエアロゾル輸送が当分野で周知である。高分子（例えばより大きなペプチドやタンパク質）の場合には、肺の肺泡領域を介する肺輸送の技術が近年向上したため、インスリン等の薬剤を実際に血中に輸送することが可能となった（Patton, J. S. らの米国特許第5,997,848号等を参照）。肺輸送は、針注射を用いなくて投与できるという点で優れており、潜在的に有毒な浸透エンハンサーが必要でなくなる。

#### 【0189】

本発明に用いる好適な組成物には、目的を達成するため、効果的な量の活性処方成分を含む組成物が含まれる。当業者は、十分に自身の能力で効果的な服用量を定めることができる。

#### 【0190】

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輸送を促進し得る。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。このようにして作製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されることが確認されている（Schwarze, S. R. ら（1999）Science 285:1569-1572）。

#### 【0191】

どのような組成物であっても、治療に効果的な薬用量は、例えば腫瘍細胞の腫瘍細胞アッセイで、或いは動物モデルのどちらかで先ず推定することができる。通常、動物モデルには、マウス、ウサギ、イヌ、サル、またはブタなどが用いられる。動物モデルはまた、好適な濃縮範囲及び投与の経路を決めるのに用いることができる。次にこのような情報を用いて、ヒトへの有益な薬用量及び投与経路を決定することができる。

#### 【0192】

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。薬用有効度及び毒性は、たとえば、ED<sub>50</sub>（服用に対して集団の50%に医薬的效果がある用量）またはLD<sub>50</sub>（服用に対して集団の50%に致命的である用量）統計を計算するなど、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決定することができる。毒性効果と治療効果との薬用量比は治療指数であり、LD<sub>50</sub>/ED<sub>50</sub>と示すことができる。高い治療指数を示す組成物が望ましい。細胞培養アッセイ及び動物実験から得られたデータが、ヒトへの適用のために、薬用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薬用量は、毒性を殆ど或いは全く含まず、ED<sub>50</sub>を含む血中濃度の範囲であることが望ましい。薬用量は、用いられる投与形態及び患者の感受性、投与の経路によって、この範囲内で様々である。

#### 【0193】

正確な薬用量は、治療が必要な患者に関する要素を考慮して、実務者によって決められるであろう。薬用量及び投与は、効果的なレベルの活性成分を与えるため或いは所望の効果

を維持するために調節される。薬用量の要素として考慮されるものには、疾患の重症度、被験者の一般的な健康状態、年齢、体重、及び患者の性別、投与の時間及び頻度、併用する薬剤、反応感受性、及び治療に対する応答が含まれる。作用期間が長い組成物は、特定の製剤の半減期及びクリアランス率に応じて3日か4日に1度、1週間に1度または2週間に1度の割合で投与し得る。

#### 【0194】

通常、薬用量は投与の経路によって異なるが、約0.1～100,000 µgまでの最大約1グラムまでである。特定の薬用量及び運搬の方法に関するガイダンスは文献に記載されており、一般に当分野の実務者は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とは異なったヌクレオチドの製剤を利用するであろう。同様に、ポリヌクレオチドまたはポリペプチドの運搬は、特定の細胞、状態、位置などに対して特異的であろう。

10

#### 【0195】

(診断)

別の実施例では、SECPに特異的に結合する抗体が、SECPの発現によって特徴付けられる疾患の診断、またはSECPやSECPのアゴニストまたはアンタゴニスト、インヒビターで治療を受けている患者をモニターするためのアッセイに用いられる。診断に有用な抗体は、治療のところで記載した方法と同じ方法で製剤される。SECPの診断アッセイには、抗体及び標識を用いてヒトの体液或いは細胞や組織から採取されたものからSECPを検出する方法が含まれる。これらの抗体は、修飾をして或いはしないで使用され、レポーター分子の共有結合性或いは非共有結合性の接着によって標識化され得る。当分野で周知の種々のレポーター分子が用いられるが、その内の幾つかは上記した。

20

#### 【0196】

SECPを測定するためのELISA,RIA,及びFACSを含む種々のプロトコルは、当分野では周知であり、変わった或いは異常なレベルのSECPの発現を診断する元となるものを提供する。正常或いは標準的なSECPの発現の値は、複合体の形成に適した条件の下、正常な哺乳動物、例えばヒトなどの被験者から採取した体液または細胞とSECPに対する抗体とを結合させることによって決定する。標準的な複合体形成の量は、測光法(photometric)などの種々の方法で定量され得る。被験者のSECPの発現の量、制御及び疾患、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間の偏差が、診断の指標となる。

30

#### 【0197】

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

#### 【0198】

一実施形態では、SECPまたは近縁の分子をコードする遺伝子配列を含むポリヌクレオチド配列を検出可能なPCRプローブを用いたハイブリダイゼーションによって、SECPをコードする核酸配列を同定することが可能である。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SECPをコードする自然界の配列のみを同定するかどうか、或いはアレル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。

40

#### 【0199】

プローブはまた、関連する配列の検出に利用され、SECPをコードする任意の配列と少なくとも50%の配列同一性を有し得る。目的の本発明のハイブリダイゼーションプローブには、DNAあるいはRNAが可能であり、SEQ ID NO: 15-28の配列、或

50

いはSECP遺伝子のプロモーター、エンハンサー、イントロンを含むゲノム配列に由来し得る。

#### 【0200】

SECPをコードするDNAに対して特異的なハイブリダイゼーションプローブの作製方法には、SECP及びSECP誘導体をコードするポリヌクレオチド配列をmRNAプローブの作製のためのベクターにクローニングする方法がある。このようなベクターは市販されており、当業者には周知であり、好適なRNAポリメラーゼ及び好適な標識されたヌクレオチドを加えることによって、in vitroでRNAプローブを合成するために用いられる。ハイブリダイゼーションプローブは、例えば<sup>32</sup>P或いは<sup>35</sup>Sなどの放射性核種、或いはアビジン／ビオチン（biotin）結合系によってプローブに結合されたアルカリホスファターゼなどの酵素標識等の種々のレポーターの集団によって標識され得る。

#### 【0201】

SECPをコードするポリヌクレオチド配列を用いて、SECPの発現に関連する疾患を診断することが可能である。限定するものではないが、このような疾患には細胞増殖異常、自己免疫／炎症性疾患、心血管疾患、神経疾患及び発達障害が含まれ、細胞増殖異常の中には日光性角化症及び動脈硬化、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、脾臓、精巣、胸腺、甲状腺、子宮の癌などが含まれ、自己免疫／炎症性疾患の中には、後天性免疫不全症候群（AIDS）及びアジソン病、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（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、癌の合併症、血液透析及び体外循環による炎症、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、外傷による炎症が含まれ、心血管疾患の中には、動静脈瘤、アテローム性硬化症、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、血管置換術、冠状動脈バイパス術移植手術、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、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、全身性紅斑性狼瘡の心内膜炎、カルチノイド心疾患、心筋症、心筋炎、心膜炎、腫瘍性心疾患、先天性心臓疾患、心臓移植の合併症、先天性肺異常、肺拡張不全、肺うっ血及び肺水腫、肺動脈塞栓症、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、細気管支拡張症、細菌性肺炎、ウイルス性肺炎及びマイコプラズマ肺炎、肺膿瘍、肺結核、びまん性間質性疾患、塵肺症、サルコイド症、特発性肺繊維症、剥離性間質性肺炎、過敏症肺炎、肺好酸球増加閉塞性細気管支炎—器質性肺炎、びまん性肺出血症候群、グッドパスチャー症候群、特発性肺血鉄症、肺併発膠原血管病、肺胞たんぱく症、肺腫瘍、炎症性及び非炎症性胸水、気胸症、胸膜腫瘍、薬物による肺疾患、放射線による肺疾患及び肺移植の合併症が含まれ、神経疾患の中には、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、クロイツフェルト-ヤコブ病及びゲルストマン-ストロイスラー-シャインカー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、脳3叉神経血管症候群、ダウン症候群を含む精神遅滞及びその他の中枢神経系の発達障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄疾患、筋ジストロフィ及びその他の神経筋障害、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、不安性の障害や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、遅発性ジスキネジア、ジストニー、分裂病性精神障害、帯状疱疹後神経痛、トゥーレット病、進行性核上麻痺、皮質基部変性及び家族性前頭側頭性健忘症が含まれ、発達障害の中には、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、癲癇、性腺形成異常、WAGR症候群（ウィルムス腫瘍、無虹彩症、尿生殖器異常及び精神薄弱）、スミス-マジェニス症候群、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-トゥス病や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舞蹈病や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、ドットプロット法、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。このような質的或いは量的方法は、当分野では公知である。

#### 【0202】

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の下、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。好適な培養期間の後、サンプルを洗浄し、シグナルを定量して基準値と比較する。患者のサンプルのシグナルの量が、制御サンプルと比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を用いて、動物実験、臨床試験、或いは個人の患者の治療を監視における、特定の治療効果を推定することが可能である。

#### 【0203】

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の下、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。標準的なハイブリダイゼーションは、正常な被験者から得た値と周知の量の実質的に精製されたポリヌクレオチドが用いられる実験からの値とを比較することによって定量可能である。正常なサンプルから得た標準的な値を、疾患の症状を示す被験者から得た値と比較可能である。基準値と被験者の値との偏差を用いて罹患しているかどうかを決定する。

#### 【0204】

疾患の存在が確定され、治療プロトコルが開始されると、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して、被験者における発現のレベルが正常な患者に示される値に近づき始めたかどうかを推定することが可能である。繰り返し行ったアッセイの結果を用いて、数日間から数ヶ月間の治療の効果を見ることができる。

#### 【0205】

癌では、個体からの生体組織における異常な量の転写物が、疾患の発生の素因を示し、また実際に臨床的症状が出る前に疾患を検出する方法を提供することが可能である。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或いは積極的な治療法を早くから利用して、癌の発生または更なる進行を防ぐことが可能となる。

#### 【0206】

SECPをコードする配列から設計されたオリゴヌクレオチドのさらなる診断への利用には、PCRの利用が含まれ得る。このようなオリゴマーは、化学的な合成、酵素を用いた生成、或いは*in vitro*で生成され得る。オリゴマーは、好ましくはSECPをコードするポリヌクレオチドの断片、或いはSECPをコードするポリヌクレオチドと相補的なポリヌクレオチドの断片を含み、最適な条件の下、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な条件の下、近縁のDNA或いはRNA配列の検出及び／または定量のため用いることが可能である。

10

#### 【0207】

或る実施態様において、SECPをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いて、一塩基多型(SNP)を検出し得る。SNPは、ヒトの先天性または後天性遺伝病の原因となる場合が多いヌクレオチドの置換、挿入及び欠失である。限定するものではないが、SNPの検出方法には、一本鎖立体構造多型(SSCP)及び蛍光SSCP(fSSCP)法が含まれる。SSCPでは、SECPをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いたポリメラーゼ連鎖反応(PCR)でDNAを増幅する。このDNAは、例えば病変或いは正常な組織、生検サンプル、体液等に由来し得る。このDNA内のSNPは、一本鎖形状のPCR産物の2次及び3次構造に差異を生じさせる。この差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。fSSCPでは、オリゴヌクレオチドプライマーを蛍光標識することによって、DNAシーケンシング装置などのハイスループット機器でアンプリマー(amp limer)の検出をすることが可能になる。更に、インシリコSNP(in silico SNP: isSNP)と呼ばれる配列データベース分析法は、共通のコンセンサス配列の構築に用いられる個々の重複するDNA断片の配列を比較することによって、多型を同定することができる。これらのコンピュータベースの方法は、DNA配列クロマトグラムの自動分析及び統計モデルを用いたシーケンシングエラーや研究室でのDNAの調整に起因する配列のばらつきを排除する。別法では、例えばハイスループットのMASSARRAYシステム(Sequenom, Inc., San Diego CA)を用いた質量分析によりSNPを検出し、特徴付ける。

20

30

#### 【0208】

SECPの発現を定量するために用いられ得る方法には、ヌクレオチドの放射標識或いはビオチン標識、調節核酸の相互増幅(coamplification)、及び標準的な曲線に結果が加えられたものが含まれる(Melby, P. C. ら(1993) J. Immunol. Methods, 159:235-44、Duplaa, C. ら(1993) Anal. Biochem. 229-236等を参照)。多数のサンプルの定量速度は、ハイスループット型のアッセイを用いることで速くなるであろう。このアッセイでは、目的のオリゴマーやポリヌクレオチドが様々な希釈液中に含まれ、分光光度法或いは非色応答によって定量が迅速である。

40

#### 【0209】

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に由来する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標的として用いることができる。マイクロアレイを、上記したように多数の遺伝子の相対的な発現レベルを同時にモニタリングする転写イメージング技術に用いることができる。マイクロアレイはまた、遺伝子変異、突然変異及び多型の同定に用いることができる。この情報を用いて、遺伝子機能を決定し、疾患の遺伝的根拠を解明し、疾患を診断し、遺伝子発現に関連する疾病の進行／後退をモニタリングし、疾患の治療における治療薬の開発や活性のモニタリングを行うことができる。特に、患者にとって最適かつ有効な治療法を選択するために、この情報

50

を用いて患者の薬理ゲノムプロファイルを作成することができる。例えば、患者の薬理ゲノムプロファイルに基づいて、患者に対して極めて効果的でありながら副作用を殆ど示さない治療薬を選択することができる。

#### 【0210】

別の実施例では、SECP、SECPの断片、SECPに特異的な抗体をマイクロアレイ上のエレメントとして用いることができる。マイクロアレイを用いて、上記のようにタンパク質間相互作用、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リング及び測定することが可能である。

#### 【0211】

特定の実施例は、或る組織または細胞型の転写イメージを生成する本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞型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、所与の条件下、所与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（Seilliamerらの米国特許第5,840,484号「Comparative Gene Transcription Analysis」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型の転写物または逆転写物の全てに本発明のポリヌクレオチドまたはその相補配列をハイブリダイズすることにより、転写イメージが生成され得る。或る実施例では、本発明のポリヌクレオチドまたはその相補配列がマイクロアレイ上に複数のエレメントのサブセットを構成するハイスループット型でハイブリダイゼーションさせる。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となり得る。

10

20

#### 【0212】

転写イメージは、組織、細胞株、生検サンプル、またはその他の生体サンプルから単離した転写物を用いて生成し得る。従って、転写イメージは、組織または生検サンプルの場合には in vivo、または細胞株の場合には in vitro における遺伝子発現を反映する。

#### 【0213】

本発明のポリヌクレオチドの発現プロファイルを示す転写イメージはまた、合成化合物または天然化合物の毒性試験のみならず、in vitroモデル系及び薬剤の前臨床評価に関連して使用され得る。全ての化合物は、作用及び毒性の機構を示唆する、頻繁に分子フィンガープリント若しくは毒性サイン（signature）と称されるような特徴的な遺伝子発現パターンを引き起こす（Nuwaysir, E. F. ら（1999）Mol. Carcinog. 24:153-159、Steiner, S. and N. L. Anderson（2000）Toxicol. Lett. 112-113:467-471、また言及することを以って本明細書の一部とする）。試験化合物が、毒性を有する既知の化合物のサインと同一のサインを有する場合には、毒性特性を共有している可能性が高い。フィンガープリンまたはサインが、より多く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れば、より有用かつ正確になる。理想としては、発現のゲノム全域にわたって測定し、最高品質のサインを提供することである。任意の試験化合物によっても発現が変化しない遺伝子も同様に重要である。それは、これらの遺伝子の発現レベルを用いて残りの発現データを標準化することができるためである。標準化処理は、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。毒性サインのエレメントへの遺伝子機能を割り当てることは毒性機構の解明に役立つが、毒性の予測につながるサインの統計的な一致には遺伝子機能の知識は必要ではない（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、毒性サインを用いる毒性スクリーニングにおいて、全ての発現した遺伝子配列を含めることは重要でありまた望ましいことである。

30

40

50

## 【0214】

一実施例では、試験化合物の毒性は、核酸を含有する生体サンプルをその試験化合物で処理して評価する。処理した生体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若しくは複数のプローブでハイブリダイズさせ、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することができる。処理した生体サンプル中の転写レベルを、非処理生体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が、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唆する。

## 【0215】

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型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」という用語は、或る特定の組織または細胞型におけるタンパク質発現の包括的パターンを指す。プロテオームを構成する各タンパク質は、個々に更なる分析をすることができる。プロテオーム発現パターン即ちプロファイルは、所与の条件下、所与の時間に発現したタンパク質の数及びそれらの相対的な存在量を定量することにより分析する。従って、ある細胞のプロテオームのプロファイルは、特定の組織または細胞型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、このような分離は2次元ゲル電気泳動によって行う。この2次元ゲル電気泳動法では、まず、1次元の等電点電気泳動によりサンプルからタンパク質を分離し、次に、2次元の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従って分離する（前出の Steiner and Anderson）。これらのタンパク質は、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染色剤を用いてゲルを染色して、分散した個別の位置にあるスポット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済みまたは未処理のいずれかの生体サンプルから得られる等位置にあ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調べる。スポット内のタンパク質は、例えば化学的または酵素的に切断した後、質量分析する標準的な方法を用いて部分的にシーケンシングする。スポット内のタンパク質の同一性は、好適には少なくとも5個の連続するアミノ酸残基であるその部分的な配列を、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。

## 【0216】

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、マイクロアレイ上の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露して各アレイエレメントへのタンパク質結合レベルを検出することによりタンパク質発現レベルを定量する（Lueking, A. ら. (1999) Anal. Biochem. 270:103-111、Mendoe, 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、チオール反応性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、各アレイエレメントにおける蛍光結合の量を検出し得る。

## 【0217】

プロテオームレベルでの毒性サイン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サインと並行して分析するべきである。或る組織における或るタンパク質では、転写物の存在量とタンパク質の存在量との相関性が低いことがあるため（Anderson, N. L. and J. Seilhamer (1997) Electrophoresis 18:533-537）、プロテオーム毒性サインは、転写イメージにはそれ程影響しないがプロテオームのプロファイルを変化させる化合物の分析において有用たり得る。更に、体液中での転写の分析は、mRNAが急速に分解するため困難である。したがって、このような場合にはプロテオームのプロファイル作成はより信頼でき、情報価値がある。

## 【0218】

別の実施例では、試験化合物の毒性は、タンパク質を含む生体サンプルをその試験化合物で処理して評価する。処理された生体サンプル中で発現したタンパク質を分離して、各タンパク質の量が定量できるようにする。各タンパク質の量を、未処理生体サンプル中の対応するタンパク質の量と比較する。両サンプル中のタンパク質の量の差は、処理されたサンプル中の試験化合物に対する毒性反応を示唆する。個々のタンパク質は、それらのアミノ酸残基をシーケンシングし、これらの部分配列を本発明のポリペプチドと比較することで同定する。

#### 【0219】

別の実施例では、試験化合物の毒性は、タンパク質を含む生体サンプルをその試験化合物で処理することにより評価する。生体サンプルから得たタンパク質を、本発明のポリペプチドに特異的な抗体と共にインキュベートする。その抗体により認識されたタンパク質の量を定量する。処理された生体サンプル中のタンパク質の量を、未処理生体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が、処理サンプル中の試験化合物に対する毒性反応を示唆する。

#### 【0220】

当分野で公知の方法を用いて、マイクロアレイを準備、使用及び分析し得る (Brennan, T. M. ら (1995) 米国特許第5,474,796号; Schena, M. ら (1996) Proc. Natl. Acad. Sci. 93:10614-10619、Baldeschweiler ら (1995) PCT出願番号WO95/251116; Shalon, D. ら (1995) PCT出願番号WO95/35505号、Heller, R. A. ら (1997) Proc. Natl. Acad. Sci. 94:2150-2155、及び Heller, M. J. ら (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周知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, ら (1999) Oxford University Press, Londonに記載されている。また、この文献を引用することを以て本明細書の一部とする。

#### 【0221】

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。コーディング配列または非コーディング配列の何れかを用いることができるが、或る例では、コーディング配列より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間にコーディング配列が保存されていることにより、染色体マッピング時に望ましくない交差ハイブリダイゼーションが生じる可能性がある。この配列は、特定の染色体、染色体の特定領域または人工の染色体、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、例えば病状の遺伝と特定の染色体領域やまたは制限断片長多型 (RFLP) の遺伝とが関連するような遺伝子連鎖地図を作成可能である (Lander, E. S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照)。

#### 【0222】

蛍光in situハイブリダイゼーション (FISH) は、他の物理的及び遺伝子地図データと関連し得る (前出のHeinz-Ulrichらによる (1995) in Meyers, pp. 965-968等を参照)。遺伝子地図データの例は、種々の科学誌あるいはOnline Mendelian Inheritance in Man (OMIM) のワールドワイドウェブのサイトで見つけることができる。物理的な染色体地図上

の S E C P 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 D N A 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。

#### 【0223】

染色体標本の *in situ* ハイブリダイゼーション及び確定した染色体マーカーを用いた結合分析などの物理的マッピング技術を用いて、遺伝子地図を拡張することもできる。マウスなど別の哺乳動物の染色体上に遺伝子を配置させることにより、たとえ正確なヒト染色体の位置が分かっていなくても、関連するマーカーが明らかになる場合が多い。この情報は、位置クローニング或いは別の遺伝子発見技術を用いて遺伝的疾患の研究をしている研究者にとって価値がある。疾患や症候群に關与する1つ或いは複数の遺伝子が、例えば血管拡張性失調症の11q22-23領域など特定の遺伝子領域へのゲノム結合により大まかに配置されると、その領域にマッピングされる任意の配列は、さらなる調査のための関連する遺伝子或いは調節遺伝子を表し得る (Gatti, R. A. らによる (1988) *Nature* 336: 577-580 等を参照)。また、目的の本発明のヌクレオチド配列を用いて、健常者、保菌者または感染者の間の、転位置、反転などによる染色体位置の違いを検出することもある。

10

#### 【0224】

本発明の別の実施例では、S E C P、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。このようなスクリーニングに用いる断片は、溶液に遊離、固体支持物に固定、細胞の表面上に保持、或いは細胞内に存在する。S E C P 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。

20

#### 【0225】

薬剤スクリーニングに用いる別の方法は、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のスクリーニング処理能力を高めるために用いられる (Geysen らによる (1984) PCT 出願番号 WO 84/03564 号等を参照)。この方法では、相当数の異なる小さな試験用化合物が、プラスチックピン或いは他の基板の上に合成される。試験用化合物は、S E C P、或いはその断片と反応してから洗浄される。次に、結合された S E C P が、当分野で公知の方法で検出される。精製された S E C P はまた、前記した薬剤をスクリーニングする技術に用いられるプレート上で直接被覆することもできる。別法では、非中和抗体を用いて、ペプチドを捕らえ、固体支持物に固定することもできる。

30

#### 【0226】

別の実施例では、S E C P と結合可能な中和抗体が S E C P 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、抗体が、S E C P 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する。

#### 【0227】

別の実施例では、発展途上の分子生物学技術に S E C P 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。

40

#### 【0228】

当業者であれば、これ以上説明しなくても前述の説明だけで最大限に本発明を利用できるであろう。したがって、以下に記載する実施例は、例示目的であって本発明を限定するものではない。

#### 【0229】

前述した及び以下に記載する全ての特許出願、特許、刊行物、特に米国特許出願第 60/091, 631 号に言及することをもって本明細書の一部とする。

#### 【0230】

(実施例)

50

## 1 cDNAライブラリの作製

インサイトcDNAはLIFESEQ GOLD データベース (Incyte Genomics, Palo Alto CA) に含まれているcDNAライブラリに由来し、表4の列5に示されている。組織の一部を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する一方、この組織の別の一部をホモジナイズしてフェノールに溶解するか、或いはTRIZOL (Life Technologies)、グアニジニウムイソチオシアネート及びフェノールの単相溶液などの好適な変性剤の混合液に溶解した。この溶解物を塩化セシウムにおいて遠心分離によって、或いはクロロホルムで抽出した。イソプロパノール或いは酢酸ナトリウムのどちらかとエタノール、或いは別の方法でこの溶解物からRNAを沈殿させた。

10

### 【0231】

RNAの純度を高めるためにRNAのフェノールによる抽出及び沈殿を必要な回数繰り返した。場合によっては、DNA分解酵素でRNAを処理する。殆どのライブラリでは、オリゴd(T)連結常磁性粒子 (Promega) またはOLIGOTEXラテックス粒子 (QIAGEN, Valencia CA)、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、POLY(A) PURE mRNA精製キット (Ambion, Austin TX) などの別のRNA単離キットを用いて組織溶解物から直接単離した。

### 【0232】

ある場合には、Stratagene社にRNAを提供し、Stratagene社が対応するcDNAライブラリを作製した。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) を用いて当分野で周知の推奨方法または類似の方法でcDNAを合成してcDNAライブラリを作製した (前出のAusubel, 1997, 第5.1-6.6章等を参照)。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に結合させてから、好適な1つの制限酵素或いは複数の制限酵素でcDNAを消化した。殆どのライブラリでは、SEPHACRYL S1000またはSEPHAROSE CL2B、SEPHAROSE CL4Bカラムクロマトグラフィー (Amersham Pharmacia Biotech)、アガロースゲル電気泳動法によってcDNAの大きさ (300~1000bp) を選択した。PBLUESCRIPTプラスミド (Stratagene) または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)、pINCYプラスミド (Incyte Pharmaceuticals, Palo Alto CA)、またはそれらの誘導体などの好適なプラスミドのポリリンカーの適合性制限酵素部位にcDNAを結合させた。この組換えプラスミドを、Stratagene社のXL1-Blue, XL1-Blue MRF、SOLR、またはLife Technologies社のDH5αまたはDH10B、ELECTROMAX DH10Bを含むコンピテント大腸菌細胞に導入し組み込んだ。

20

30

40

### 【0233】

## 2 cDNAクローンの単離

上記実施例1に記載したように得たプラスミドを、UNIZAPベクターシステム (Stratagene) 或いは細胞溶解を利用した*in vivo*切除によって宿主細胞から回収した。MagicまたはWIZARD Minipreps DNA精製システム (Promega)、及びAGTC Miniprep精製キット (Edge Biosystems, Gaithersburg MD)、QIAGEN社のQIAWELL 8 Plasmid、QIAWELL 8 Plus Plasmid、QIAWELL 8 Ultra Plasmid 精製システム、REAL Prep 96プラスミドキットの内の少なくとも1つを用いてプラスミドを精製した。沈殿させた後、0.1mlの蒸留水に再懸

50

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。

#### 【0234】

別法では、ハイスループットの直接結合PCR法によっ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を単一反応混合液で行った。サンプルを処理してから384-ウェルプレートに移して保管し、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)を用いて蛍光定量的に測定した。

#### 【0235】

10

### 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インサイトcDNAを、以下に示すようにシークエンシングした。cDNAのシークエンシング反応は標準的な方法で行うか、またはHYDRAマイクロディスペンサー(Robbins Scientific)或いはMICROLAB 2200(Hamilton)液体移送装置と共にABI CATALYST 800(Applied Biosystems)サーマルサイクラー或いはPTC-200 thermal cycler(MJ Research)などのハイスループット装置を用いて行った。cDNAのシークエンシング反応は、Amersham Pharmacia Biotech社の試薬、またはABI PRISM BIGDYE Terminator cycle sequencing ready react 20  
ionキット(Applied Biosystems)などのABIシークエンシングキットに含まれる試薬を用いて行った。cDNAシークエンシングの反応物の電気泳動的による分離及び標識したポリヌクレオチドの検出は、MEGABACE 1000 DNAシークエンシングシステム(Molecular Dynamics)、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)、または当分野で周知のその他の配列解析システムを用いて行った。cDNA配列内のリーディングフレームリーディングフレームは、標準的な方法(前出のAusubel, 1997, 第7.7章)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した。 30

#### 【0236】

インサイトcDNAに由来する本ポリヌクレオチド配列の確認は、BLAST、動的プログラミング、およびジヌクレオチドの分布による解析(dinucleotide nearest neighbor analysis)に基づいたプログラム及びアルゴリズムを用いて、ベクター、リンカー、およびポリA配列を取り除き、更にあいまいな塩基対をマスクすることで行った。次に、インサイトcDNA配列およびそれらの翻訳を、公共のデータベースであるGenBankの霊長類、げっ歯類、哺乳類、脊椎動物、および真核生物のデータベース、およびBLOCKS、PRINTS、DOMO、PRODOM、およびPFAMなどの隠れマルコフモデル(HMM)を基にしたタンパク質ファミリーのデータベースから選択した配列に対して問合せた(HMMは、遺伝子ファミリーのコンセンサス主構造を分析する確率的手法である。例えばEddy, S. R. (1996) Curr. Opin. Struct. Biol. 6: 361-365を参照)。こ 40  
のような問合せは、BLAST、FASTA、BLIMPS、およびHMMRに基づいたプログラムを用いて行った。インサイトcDNA配列を組み立てて、完全長ポリヌクレオチド配列を作製した。或いは、GenBank cDNAs、GenBank EST、ステッチ配列(stitched sequence)、ストレッチ配列(stretched sequences)、またはGenscan推定コード配列(実施例4及び5を参照)を用いて、インサイトcDNA群を完全長の配列に伸長した。配列の組み立ては、Phred、Phrap、およびConsedに基づいたプログラムを用いて行い、GeneMark、BLAST、およびFASTAに基づいたプログラムを用いて、オープン 50

リーディングフレームを決定するべく cDNA 群をスクリーニングした。完全長のポリヌクレオチド配列を翻訳して対応する完全長ポリペプチド配列を得た。別法では、本発明のポリヌクレオチドは、完全長翻訳ポリペプチドの任意のメチオニン残基から始まり得る。次に、完全長ポリペプチド配列を GenBank タンパク質データベース (genpept)、SwissProt、BLOCKS、PRINTS、DOMO、PRODOM、Prosit、および PFAM などの隠れマルコフモデル (HMM) に基づいたタンパク質ファミリーデータベースに対して問合せて分析した。これらの完全長ポリヌクレオチド配列はまた、MACDNASIS PRO ソフトウェア (Hitachi Software Engineering, South San Francisco CA) および LASERGENE ソフトウェア (DNASTAR) を用いて分析した。ポリヌクレオチド配列およびポリペプチド配列のアラインメントを、アラインメントした配列間のパーセント同一性も計算する MEGALIGN マルチシーケンスアラインメントプログラム (DNASTAR) に組み込まれた CLUSTAL アルゴリズムによって指定されたデフォルトパラメータを用いて作成した。

10

#### 【0237】

表 7 は、インサイト cDNA および完全長配列の組み立て、および組み立てた配列の分析に利用したツール、プログラム、およびアルゴリズム、並びにそれらの説明、引用文献、閾値パラメータを簡単に示す。表 7 の列 1 は用いたツール、プログラム、およびアルゴリズム、列 2 はそれらの簡単な説明、列 3 は引用することで本明細書の一部とした引用文献、列 4 の記載されている部分は 2 つの配列の一致の程度を評価するために用いたスコア、確率値及びその他のパラメータを示す (スコアが高くなれば高くなるほど即ち確率値が低ければ低いほど、配列間の相同性が高くなる)。

20

#### 【0238】

完全長の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: 15-28 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。

#### 【0239】

##### 4 ゲノム DNA からのコード配列の同定および編集

推定上の分泌タンパク質は、公のゲノム配列データベース (例えば gbpr i や gbht g) 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、様々な生物に由来するゲノム DNA 配列を分析するための汎用遺伝子同定プログラムである (Burge, C. and S. Karlin (1997) J. Mol. Biol. 268: 78-94, Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8: 346-354 を参照)。このプログラムは推定エキソンを連結して、メチオニンから停止コドンまで伸長した組み立て cDNA 配列を構築する。Genscan により得られる配列は、FASTA データベースのポリヌクレオチド配列およびポリペプチド配列になる。Genscan によって一回で解析できる配列の最大長さは 30 kb に設定されている。これらの Genscan 推定 cDNA 配列の内、どの配列が分泌タンパク質をコードするかを決定するために、コードされたポリペプチドを PFAM モデルにおいて分泌タンパク質について問合せて分析した。潜在的な分泌タンパク質が、分泌タンパク質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。次に、これらの選択された Genscan 推定配列を、BLAST 解析を用いて genept および gbpr i 公共データベースの配列と比較した。必要に応じて、Genscan 推定 cDNA 配列を、genept において BLAST で最もヒットした配列と比較して、Genscan 推定配列における余分なエキソンや省いてしまったエキソンなどのエラーを修正し、編集した。BLAST 解析を用いて Genscan 推定 cDNA 配列を含むインサイト cDNA または公共の cDNA を見つけ出すことにより、転写の証拠が得られる。インサイト cDNA が Genscan 推定 cDNA 配列を含む場合、この情報を用いて Gensc

30

40

50

an 推定配列を修正或いは確認できる。完全長ポリヌクレオチド配列は、実施例 3 に説明した組み立て方法で Genscan 推定コード配列とインサイト cDNA 及び／または公共の cDNA 配列を組み立てて作製した。別法では、完全長ポリヌクレオチド配列は、その全てが編集した或いは未編集の Genscan 推定コード配列から作製した。

#### 【0240】

##### 5 cDNA 配列データを用いるゲノム配列データの組み立て

##### ステッチ配列 (Stitched Sequence)

部分的な cDNA 配列を、実施例 4 に記載した Genscan 遺伝子同定プログラムによって推定されたエキソンで伸長した。実施例 3 に記載されたように組み立てられた部分的な cDNA をゲノム DNA にマッピングし、関連する cDNA および 1 或いは複数のゲノム配列に由来する関連する推定 Genscan エキソンを含む複数のクラスターに入れた。各クラスターを、グラフ理論および動的計画法に基づいたアルゴリズムを用いて、cDNA およびゲノム情報を統合して分析し、後に確認される潜在的なスプライスバリエーションを生成し、編集或いは伸長して完全長の配列を作製した。或るクラスターの 2 つ以上の配列に或る区間の全長が存在する配列区間を同定し、推移 (transitivity) により同定した区間を同等と考える。例えば、或る区間が cDNA および 2 つのゲノム配列のそれぞれに存在する場合、これら 3 つ全ての区間を同等と考える。この方法によって、関連しないが連続するゲノム配列を cDNA 配列によって繋ぎ 1 つにする。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。或るタイプ (cDNA と cDNA、またはゲノム配列とゲノム配列) の親配列に沿って連結される区間と区間との繋ぎ合わせは、親配列のタイプが異なる (cDNA とゲノム配列) 連結より好ましい。得られたステッチ配列を翻訳し、BLAST 解析で genpept および gbprl 公共データベースにおける配列と比較した。Genscan によって推定された不適当なエキソンを、genpept において BLAST で最もヒットした配列と比較して修正する。このような配列を更なる cDNA 配列で伸長し、必要に応じてゲノム DNA で検査した。

#### 【0241】

##### ストレッチ配列 (Stretched Sequence)

部分的な DNA 配列を BLAST 解析に基づいたアルゴリズムで完全長に伸長した。まず、実施例 3 に記載したように組み立てた部分的な cDNA を、BLAST プログラムを用いて GenBank の霊長類、げっ歯類、哺乳類、脊椎動物、および真核生物のデータベースなどの公共のデータベースに対して問い合わせた。次に、GenBank の相同性の最も高いタンパク質を、実施例 4 に記載したインサイト cDNA 或いは Genscan エキソン推定配列の何れかと比較した。得られた複数の高スコアのセグメント対 (HSP) を用いてキメラタンパク質を作製し、GenBank の相同タンパク質上に翻訳した配列をマッピングした。元の GenBank の相同タンパク質に対して、キメラタンパク質に挿入や欠失が起こり得る。公共のヒトゲノムデータベースから相同ゲノム配列を探し出すために、GenBank の相同タンパク質およびキメラタンパク質の両方をプローブとして用いた。このようにして、部分的な DNA 配列を相同ゲノム配列の付加によりストレッチすなわち伸長した。完全な遺伝子を含んでいるか得られたストレッチ配列を検査した。

#### 【0242】

##### 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: 15-28 を組み立て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、インサイト 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: 15-28 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。Stanford Human Genome Center (SHGC)、Whitehead Institute for Genome Research (WIGR) 及び Genethon などの公

共の情報源から入手できる放射線ハイブリッド (radiation hybrid) 及び遺伝子マッピングのデータを用いて、クラスター化した配列がすでにマッピングされているかを調べる。クラスターにマッピングされた配列が含まれている場合は、そのクラスターの全ての配列 (特定の SEQ ID NOを含む) をそのマッピング位置に割り当てた。

#### 【0243】

遺伝子地図の位置は、範囲、区間、またはヒト染色体によって表される。センチモルガンで示したマッピング位置の範囲は、染色体の短腕 (p) の末端から測定した (センチモルガン (cM) は、同一染色体上の遺伝子間の乗換え率に基づいた距離を表す単位である。平均すると、1 cMはヒトの染色体の1メガベースに概ね等しいが、組換え率の高い部分と低い部分があるため、大きく変化し得る)。距離 cMは、配列がそれぞれのクラスターに含まれている放射線ハイブリッドマーカーの境界を検出できる Genethonによってマッピングされた遺伝子マーカーに基づいている。NCBI「GeneMap99」 (<http://www.ncbi.nlm.nih.gov/genemap>) などの公衆利用可能なヒト遺伝子マップ及びその他の情報源を用いて、上記し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。

10

#### 【0244】

##### 7 ポリヌクレオチド発現の分析

ノーザン分析は、遺伝子の転写物の存在を検出するために用いられる実験用技術であり、特定の細胞種或いは組織からのRNAが結合されてい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伴う (前出の Sambrook 第7章、前出の Ausubel, F. M. ら、第4章及び16章等を参照)。

20

#### 【0245】

BLASTに用いる類似のコンピュータ技術を用いて、GenBank或いはLIFESEQ (Incyte Pharmaceuticals) のようなcDNAデータベース内の同一或いは関連する分子を検索する。この分析は多くの膜系ハイブリダイゼーションより非常に速度が速い。さらにコンピュータ検索の感度を変更して、任意の特定の一致が、厳密な一致或いは相同的一致の何れかとして分類されるかを確定することができる。検索の基準は、

#### 【0246】

30

#### 【数1】

(BLAST スコア×配列一致率)

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

#### 【0247】

として定義される積スコアである。積スコアは、0～100の標準化された値であり、以下のように求める。BLASTスコアにヌクレオチド配列の一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除す。高スコアのセグメントの対 (HSP) において一致する各塩基に+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 (ギャップにより離隔される)。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、BLASTアラインメントの断片的重複と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%一致する場合にのみ得られる。積スコア70は、100%の同一性で重畳が70%であるか、或いは88%の同一性で重畳が100%であるかの何れかの場合である。積スコア50は、100%の同一性で重畳が50%であるか、或いは79%の同一性で重畳が100%であるかの何れかの場合である。

40

#### 【0248】

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

50

。例えば、ある完全長の配列は、少なくとも部分的にインサイト cDNA 配列をオーバーラップさせて組み立てられる（実施例 3 を参照）。各 cDNA 配列は、ヒト組織から作製された cDNA ライブラリに由来する。各ヒト組織は、心血管系、結合組織、消化系、胚構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液および免疫系、肺、筋骨格系、神経系、脾臓、呼吸器系、感覚器官、皮膚、顎口腔系、分類不能／混合、または尿管などの 1 つの生物／組織のカテゴリーに分類される。各カテゴリーにおけるライブラリの数をカウントし、その合計数を全カテゴリーのライブラリ数で除す。同様に、各ヒト組織は、癌、細胞系、発生、炎症、神経、外傷、心血管、プール（pooled）などの 1 つの疾患／症状のカテゴリーに分類され、各カテゴリーにおけるライブラリの数をカウントし、その合計数を全カテゴリーのライブラリ数で除す。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（Incyte Genomics, Palo Alto CA）から得ることができる。

10

#### 【0249】

#### 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は、完全長分子の好適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その完全長分子の好適な断片を伸長して作製した。一方のプライマーは既知の断片の 5' の伸長を開始するために合成し、他方のプライマーは既知の断片の 3' の伸長を開始するために合成した。開始プライマーは、OLIGO 4.06 ソフトウェア（National Biosciences）或いは他の適切なプログラムを用いて、約 22 個から約 30 個のヌクレオチドの長さで約 50% 以上の GC 含量を有し、かつ約 68～72℃の温度で標的配列にアニールするように設計した。ヘアピン構造及びプライマープライマー二量体が生じないようにヌクレオチドを伸長した。

20

#### 【0250】

選択されたヒト cDNA ライブラリを用いてこの配列を伸長した。2 段階以上の伸長が必要な場合、若しくは望ましい場合は、追加或いはネスト化プライマーの組を設計する。

#### 【0251】

当分野で既知の方法を利用した PCR 法で高い忠実度で増幅した。PCR は PTC-200 thermal cycler（MJ Research, Inc.）を用いて 96 ウェルブロックプレートで行った。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  $\beta$ -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。

30

|        |                          |
|--------|--------------------------|
| ステップ 1 | 94℃で3分間                  |
| ステップ 2 | 94℃で15秒                  |
| ステップ 3 | 60℃で1分間                  |
| ステップ 4 | 68℃で2分間                  |
| ステップ 5 | ステップ 2、3 及び 4 を 20 回繰り返す |
| ステップ 6 | 68℃で5分間                  |
| ステップ 7 | 4℃で保管                    |

40

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℃で3分間                  |
| ステップ 2 | 94℃で15秒                  |
| ステップ 3 | 57℃で1分間                  |
| ステップ 4 | 68℃で2分間                  |
| ステップ 5 | ステップ 2、3 及び 4 を 20 回繰り返す |
| ステップ 6 | 68℃で5分間                  |
| ステップ 7 | 4℃で保管。                   |

50

## 【0252】

各ウェルのDNA濃度は、1X TE及び0.5  $\mu$ lの希釈していないPCR産物に溶解した100  $\mu$ lのPICO GREEN定量試薬(0.25% (v/v) PICO GREEN; Molecular Probes, Eugene OR)を不透明な蛍光光度計プレート(Coming Costar, Acton MA)の各ウェルに分配してDNAが試薬と結合できるようにして測定する。このプレートをFluoroskan II (Labsystems Oy, Helsinki, Finland)でスキャンして、サンプルの蛍光を計測してDNAの濃度を定量化する。反応混合物の5~10  $\mu$ lのアリコットを1%のアガロースミニゲル上での電気泳動によって解析し、何れの反応物が配列を伸長することに成功したかを決定する。

10

## 【0253】

伸長したヌクレオチドを脱塩及び濃縮してから384ウェルプレートに移し、CviJIコレラウィルスエンドヌクレアーゼ(Molecular Biology Research, Madison WI)で消化し、pUC18ベクター(Amersham Pharmacia Biotech)に再連結する前に音波処理またはせん断を行った。ショットガンシーケンシングのために、消化したヌクレオチドを低濃度(0.6~0.8%)のアガロースゲル上に分離して断片を切断し、寒天をAgar ACE (Promega)で消化した。T4リガーゼ(New England Biolabs, Beverly MA)を用いて伸長したクローンをpUC18ベクター(Amersham Pharmacia Biotech)に再連結し、Pfu DNAポリメラーゼ(Stratagene)で制限部位の延び出しを処理してコンピテント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。

20

## 【0254】

細胞を溶解して、Taq DNAポリメラーゼ(Amersham Pharmacia Biotech)及びPfu DNAポリメラーゼ(Stratagene)を用いて以下の手順でDNAをPCR増幅した。

- ステップ1 94℃で3分間
- ステップ2 94℃で15秒
- ステップ3 60℃で1分間
- ステップ4 72℃で2分間
- ステップ5 ステップ2、3及び4を29回繰り返す
- ステップ6 72℃で5分間
- ステップ7 4℃で保管。

30

上記したようにPICO GREEN試薬(Molecular Probes)でDNAを定量化した。DNA回収率の悪いサンプルは、上記した条件で再び増幅した。サンプルを20%のジメチルサルホサイド(dimethylsulphoxide)(1:2, v/v)で希釈し、DYENAMIC DIRECTキット(Amersham Pharmacia Biotech)またはABI PRISM BIGDYE Terminator cycle sequencing ready reactionキット(Applied Biosystems)を用いてシーケンシングした。

40

## 【0255】

同様にして、上述の手順で完全長のポリヌクレオチド配列を検査し、或いは完全長のポリヌクレオチド配列を用いて、この伸長のために設計したオリゴヌクレオチドと好適なゲノムライブラリを用いて、5'調節配列を得た。

## 【0256】

## 9 個々のハイブリダイゼーションプローブの標識化及び使用

SEQ ID NO: 15-28から導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すが、より大きなcDNAフラグメントの場合

50

合でも基本的に同じ手順を用いる。オリゴヌクレオチドを、OLIGO 4.06ソフトウェア (National Bioscience) のような最新式のソフトウェアを用いてデザインし、50 pmol の各オリゴマーと、250  $\mu$ Ci の [ $\gamma$ - $^{32}$ P] アデノシン三リン酸 (Amersham, Chicago, IL) 及び T4 ポリヌクレオチドキナーゼ (DuPont NEN, Boston MA) とを組み合わせて用いることにより標識する。標識されたオリゴヌクレオチドを、SEPHADEX G-25 超精細排除ゲキストランビードカラム (Amersham Pharmacia Biotech) を用いて実質的に精製する。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、次のエンドヌクレアーゼ、Ase I、Bgl II、Eco RI、Pst I、Xba I 或いは Pvu II (DuPont NEN) の1つを用いて切断したヒトゲノムDNAの典型的な膜ベースのハイブリダイゼーション解析において用いる。 10

#### 【0257】

各切断物からのDNAを、0.7%アガロースゲル上で分画して、ナイロン製メンブラン (Nytran Plus, Schleicher & Schuell, Durham NH) に転写する。ハイブリダイゼーションは40℃で16時間かけて行う。非特異的シグナルを取り除くため、例えば、最大0.1x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の条件の下、ブロットを順次室温にて洗浄する。ハイブリダイゼーションパターンをオートラジオグラフィー或いは別のイメージ化手段で視覚化して比較する。

#### 【0258】

20

#### 10 マイクロアレイ

マイクロアレイ上のアレイエレメントの連結または合成は、フォトリソグラフィ、圧電プリント (インクジェットプリント、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板は、均一な非多孔性の固体とするべきである (前出の Schena (1999))。推奨する基板には、シリコン、シリカ、スライドガラス、ガラスチップ及びシリコンウエハがある。別法では、ドットブロット法またはスロットブロット法に類似のアレイを利用して、熱や紫外線、または化学的或いは機械的な結合手段で基板の表面にエレメントを配置して結合させることができる。通常のアレイは利用可能な方法や機械を用いて作製でき、任意の適正な数のエレメントを含めることができる (Schena, M. ら (1995) Science 270: 467-470、Shalon, D. ら (1996) Genome Res. 6: 639-645、Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16: 27-31. 等を参照)。

30

#### 【0259】

完全長cDNA、発現遺伝子配列断片 (EST)、或いはそれらの断片やオリゴマーが、マイクロアレイのエレメントとなり得る。ハイブリダイゼーションに好適な断片やオリゴマーを、LASERGENEソフトウェア (DNASTAR) などの当分野で周知のソフトウェアを用いて選択することが可能である。このアレイエレメントを、生体サンプル中のポリヌクレオチドとハイブリダイズさせる。生体サンプル中のポリヌクレオチドは、検出を容易にするために蛍光標識またはその他の分子タグに結合する。ハイブリダイゼーションの後、生体サンプルからハイブリダイズしなかった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けるハイブリダイゼーションを検出する。別法では、レーザー脱離及び質量スペクトロメトリー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程度及び相対的存在量は、算定することができる。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。

40

#### 【0260】

#### 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 (d 50

T) セルロース法を用いてポリ (A) <sup>+</sup> RNA を精製する。各ポリ (A) <sup>+</sup> 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 (21 mer)、1×第1鎖緩衝液、0.03 単位/μl のRNアーゼインヒビター、500 μM dATP、500 μM dGTP、500 μM dTTP、40 μM dCTP、40 μM dCTP-Cy3 (BDS) または dCTP-Cy5 (Amersham Pharmacia Biotech) を用いて逆転写する。この逆転写反応は、GEMBRIGHT キット (Incyte) を用いて、200 ng のポリ (A) <sup>+</sup> RNA を含む 25 ml 容量で行う。特異的なコントロールポリ (A) <sup>+</sup> RNA は、in vitro 転写により非コーディング酵母ゲノムDNA から合成する。370℃で2時間インキュベートした後、各反応サンプル (一方はCy3 標識、他方はCy5 標識) は、2.5 ml の0.5 M 水酸化ナトリウムで処理し、850℃で20分間インキュベートし、反応を停止させてRNA を変性する。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピナラム (CLONTECH Laboratories, Inc. (CLONTECH)), Palo Alto CA) を用いて精製する。結合後、2つの反応サンプルを、1 ml のグリコーゲン (1 mg/ml)、60 ml の酢酸ナトリウム及び300 ml の100%エタノールを用いてエタノール沈殿させる。サンプルは次に、Speed VAC (Savant Instruments Inc., Holbrook NY) を用いて乾燥して仕上げ、14 μl 5×SSC/0.2% SDS 中で再懸濁する。

10

#### 【0261】

##### マイクロアレイの準備

20

本発明の配列を用いて、アレイエレメントを作製する。各アレイエレメントは、クローン化cDNA挿入断片を含むベクターを含有する細菌性細胞から増幅する。PCR増幅は、cDNA挿入断片に隣接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によって、1~2 ngの初期量から5 μgを超える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) を用いて精製する。

#### 【0262】

精製したアレイエレメントを、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大量の蒸留水での洗浄と、0.1%のSDS及びアセトン中で超音波による洗浄を行う。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%エタノール中の0.05%アミノプロピルシラン (Sigma) でコーティングする。コーティングしたスライドガラスは、110℃の天火で硬化させる。

30

#### 【0263】

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っ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。次にこの装置が、スライド毎に約5 nlのアレイエレメントサンプルを分注する。

40

#### 【0264】

マイクロアレイには、STRATALINKER UVクロスリンカー (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回洗浄し、蒸留水で3回洗浄する。非特異的な結合部位は、リン酸緩衝生理食塩水 (PBS) (Tropix, Inc., Bedford MA) における0.2%カゼイン中で60℃で30分間マイクロアレイをインキュベートし、その後上述したように0.2% SDS及び蒸留水で洗浄することによってブロックする。

#### 【0265】

50

### ハイブリダイゼーション

ハイブリダイゼーション反応液は、 $5\times\text{SSC}$ 、 $0.2\%$  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  $0.2\mu\text{g}$  含む  $9\mu\text{l}$  のサンプル混合体を含めたものである。サンプル混合液を、 $65^\circ\text{C}$  で 5 分間加熱し、マイクロアレイ表面上に一定量分注してから  $1.8\text{cm}^2$  のカバーガラスで覆う。このアレイを、顕微鏡スライドより僅かに大きいキャビティを有する防水チェンバーに移す。チェンバーの角に  $140\mu\text{l}$  の  $5\times\text{SSC}$  を加えて、チェンバー内を湿度  $100\%$  に保持する。このアレイを含むチェンバーを、 $60^\circ\text{C}$  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 ( $1\times\text{SSC}$ ,  $0.1\%$  SDS) において  $45^\circ\text{C}$  で 10 分間、第 2 洗浄緩衝液中 ( $0.1\times\text{SSC}$ ) において  $45^\circ\text{C}$  で 10 分間それぞれ 3 回洗浄し、その後乾燥させる。

### 【0266】

#### 検出

レポーター標識されたハイブリダイゼーション複合体は、Cy 3 を励起するための  $488\text{nm}$ 、及び Cy 3 を励起するための  $632\text{nm}$  のスペクトル線を生成し得る Innova 70 混合ガス 10 W レーザー (Coherent, Inc., Santa Clara CA) を備えた顕微鏡で検出する。20 倍の顕微鏡対物レンズ (Nikon, Inc., Melville NY) を用いて、アレイ上に励起レーザー光を集中させる。このアレイを含むスライドを顕微鏡のコンピュータ制御 X-Y ステージに置き、対物レンズを通してラスタスキャンする。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する。

### 【0267】

2 つの異なるスキャンにおいて、混合ガスマルチラインレーザーは 2 つの蛍光体を連続的に励起する。放射された光は、波長に基づいて 2 つの蛍光体に対応する 2 つの光電子増倍管検出器 (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) に分割される。アレイと光電子増倍管との間に配設された好適なフィルタを用いて信号をフィルタリングする。用いる蛍光体の最大発光は、Cy 3 では  $565\text{nm}$ 、Cy 5 では  $650\text{nm}$  である。装置は両方の蛍光体からのスペクトルを同時に記録できるが、レーザー源に好適なフィルタを用いて、蛍光体 1 つにつき 1 回スキャンし、各アレイを通常 2 回スキャンする。

### 【0268】

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 コントロール種により生成されるシグナル強度を用いて較正する。アレイ上の特定の位置には相補的 DNA 配列を含め、その位置におけるシグナルの強度がハイブリダイズする種の重量比 1:100, 000 に相関するようにする。異なる試料 (例えば検査細胞及びコントロール細胞を代表する) からの 2 つのサンプルを、各々異なる蛍光体で標識し、他と異なって発現する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させる場合には、較正は 2 つの蛍光体を有する較正する cDNA のサンプルを標識して、ハイブリダイゼーション混合液に各々等量を加えて行う。

### 【0269】

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、リニア 20 色変換を用いてシグナル強度が青色 (低シグナル) から赤色 (高シグナル) までの擬似カラー範囲にマッピングされるイメージとして表示される。データはまた、定量的に分析される。2 つの異なる蛍光体を同時に励起して測定する場合には、各蛍光体の発光スペクトルを用いて、先ずデータは蛍光体間の光学的漏話 (重複発光スペクトルに起因する) に対して補正される。

### 【0270】

グリッドを蛍光シグナルイメージ上に重畳して、各スポットからのシグナルがグリッドの

各エレメントに中央に位置するようにする。各エレメント内の蛍光シグナルを統合し、シグナルの平均強度に対応する数値を得る。シグナル分析に用いるソフトウェアは、G E M T O O L S 遺伝子発現分析プログラム (I n c y t e) である。

## 【0271】

### 1 1 相補的ポリヌクレオチド

S E C P 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 S E C P の発現を低下させるため即ち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、より小さな或いはより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。O l i g o 4 . 0 6 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) 及び S E C P 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' 配列から相補的な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 S E C P をコードする転写物に結合するのを阻害する。

## 【0272】

### 1 2 S E C P の発現

S E C P 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 S E C P が発現するために、抗生物質耐性及び c D N A 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 c D N A をサブクローニングする。このようなプロモーターには、l a c オペレーター調節エレメントに関連する T 5 または T 7 バクテリオファージプロモーター及び t r p - l a c ( t a c ) ハイブリッドプロモーターが含まれるが、これらに限定されるものではない。組換えベクターを、B L 2 1 ( D E 3 ) など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 ( I P T G ) で誘発されると S E C P を発現する。真核細胞での S E C P 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 A u t o g r a p h i c a c a l i f o r n i c a 核多面性ウイルス ( A c M N P V ) 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、S E C P をコードする c D N A と置換する。ウイルスの感染力は維持され、強いポリヘドリンプロモータによって高いレベルの c D N A 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 S p o d o p t e r a f r u g i p e r d a ( S f 9 ) 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 ( E n g e l h a r d , E . K . ら ( 1 9 9 4 ) P r o c . N a t l . A c a d . S c i . U S A 9 1 : 3 2 2 4 - 3 2 2 7 , S a n d i g , V . ら ( 1 9 9 6 ) H u m . G e n e T h e r . 7 : 1 9 3 7 - 1 9 4 5 . 等を参照)。

## 【0273】

殆どの発現系では、S E C P が、例えばグルタチオン S トランスフェラーゼ ( G S T ) 、または F L A G や 6 - H i s 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。 S c h i s t o s o m a j a p o n i c u m からの 26 キロダルトンの酵素 G S T によって、タンパク質の活性及び抗原性を維持した状態で固定されたグルタチオンで融合タンパク質の精製が可能となる ( A m e r s h a m P h a r m a c i a B i o t e c h ) 。精製の後、G S T 部分を特定の操作部位で S E C P からタンパク質分解的に切断できる。アミノ酸 8 個のペプチドである F L A G で、市販のモノクローナル及びポリクローナル抗 F L A G 抗体 ( E a s t m a n K o d a k ) を用いた免疫親和性の精製が可能となる。6 個の連続するヒスチジン残基のストレッチである 6 - H i s によって、金属キレート樹脂 ( Q I A G E N ) で精製が可能となる。タンパク質の発現及び精製の方法は、A u s u b e l ( 1 9 9 5 , 前出 , 第 1 0 及び 1 6 章 ) に記載されている。これらの方法で精製した S E C P を直接用いて以下の実施例 1 6 及び 1 7 のア

ッセイを行うことができる。

#### 【0274】

##### 1.3 機能的アッセイ

S E C Pの機能は、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S E C Pをコードする配列の発現によって評価する。c D N Aを、c D N A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サブクロニングする。このようなベクターには、p C M V S P O R T<sup>T M</sup> (L i f e T e c h n o l o g i e s . ) 及びp C R 3 . 1 (I n v i t r o g e n , C a r l s b a d , C A)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含んでいる。5～10 μgの組換えベクターを、例えば内皮由来か造血由来のヒト細胞株にリポソーム製剤或いは電気穿孔法によって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、形質移入された細胞と形質移入されていない細胞とを区別できる。また、標識タンパク質の発現によって、c D N Aの組換えベクターからの発現を正確に予想できる。このような標識タンパク質には、緑色蛍光タンパク質(G F P ; C l o n t e c h)、及びC D 6 4またはC D 6 4-G F P融合タンパク質が含まれる。レーザー光学に基づいた技術を利用した自動流動細胞計測法(F C M)を用いて、G F PまたはC D 6 4-G F P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。また、F C Mで、先行した或いは同時の細胞死の現象を診断する蛍光分子の取り込みを検出して計量する。これらの現象には、プロピジウムヨウ化物でのD N Aの染色によって計測される核D N A内容物の変化と、プロモデオキシウリジンの取り込み量の低下によって計測されるD N A合成の下方調節と、特異的な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のタンパク質の発現の変化と、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含まれる。流動細胞計測法は、O r m e r o d , M . G . による(1994) F l o w C y t o m e t r y O x f o r d , N e w Y o r k , N Y . に記載されている。

#### 【0275】

遺伝子発現におけるS E C Pの影響は、S E C Pをコードする配列とC D 6 4またはC D 6 4-G F P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。C D 6 4またはC D 6 4-G F P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 (I g G)の保存された領域と結合する。形質転換された細胞と形質転換されない細胞とは、ヒトI g GかC D 6 4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(D Y N A L . L a k e S u c c e s s . N Y)。m R N A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。S E C P及び目的の他の遺伝子をコードするm R N A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。

#### 【0276】

##### 1.4 S E C P特異抗体の作製

ポリアクリルアミドゲル電気泳動法(P A G E ; 例えばH a r r i n g t o n , M . G . (1990) M e t h o d s E n z y m o l . 182 : 488-495を参照)または他の精製技術で実質的に精製されたS E C P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。

#### 【0277】

別法では、S E C Pアミノ酸配列をL A S E R G E N Eソフトウェア(D N A S T A R)を用いて解析して免疫原性の高い領域を決定し、対応するオリゴペプチドを合成してこれを用いて当業者に公知の方法で抗体を産生する。C末端付近の、或いは隣接する親水性領域内のエピトープなど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A u s u b e l , 1995, 11章等を参照)。

#### 【0278】

通常、約15残基の長さのオリゴペプチドを、A p p l i e d B i o s y s t e m s の

ABI 431A ペプチドシンセサイザー (Applied Biosystems) を用いて fmoc 法のケミストリにより合成し、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) を用いた反応により KLH (Sigma-Aldrich, St. Louis MO) に結合させて、免疫原性を高める (例えば、前出の Ausubel, 1995 を参照)。フロイントの完全アジュバントにおいてオリゴペプチド-KLH 複合体を用いてウサギを免疫化する。得られた抗血清の抗ペプチド活性及び抗 SECP 活性を検査するには、ペプチドまたは SECP を基板に結合し、1% BSA を用いてブロッキング処理し、ウサギ抗血清と反応させて洗浄し、さらに放射性ヨウ素標識されたヤギ抗ウサギ IgG と反応させる。

【0279】

10

#### 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。結合の後、そのレジンを製造者の使用説明書に従ってブロッキング処理し洗浄する。

【0280】

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。

20

【0281】

#### 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なその断片を、<sup>125</sup>I Bolton-Hunter 試薬 (Bolton A. E. 及び W. M. Hunter (1973) Biochem. J. 133: 529 等を参照) で標識する。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。

30

【0282】

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

【0283】

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. ら (2000) 米国特許第 6,057,101 号)。

40

【0284】

#### 1.7 SECP の活性の実証

SECP の活性を成長刺激または阻害するためのアッセイは、Swiss マウス 3T3 細胞における DNA の合成の速度を測定する (McKay, I. および I. Leigh, eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York NY, p. 73)。このアッセイでは、放射性 DNA 前駆体分子である [<sup>3</sup>H] チミジンの存在下で様々な量の SECP を静止状態の 3T3 培養細胞に加える。このアッセイの

50

ためのSECPは、組換え手段によって或いは生化学的調製から得られる。酸沈降性DNAへの $[^3\text{H}]$ チミジンの取込みを適当な時間間隔で測定すると、取り込まれた量は新たに合成されたDNAの量に正比例する。少なくとも100倍のSECPリガンドの濃度範囲に対する一次の線量効果曲線は、成長調節活性を表す。1 ml当たりの活性1単位を、50%の反応レベルを引き起こすSECPの濃度と定義する。ここで、100%は、 $[^3\text{H}]$ チミジンの酸沈降性DNAへの最大の取込みを表す。

#### 【0285】

別法では、SECP活性のためのアッセイは、培養細胞における神経伝達の刺激または阻害を測定する。培養CHO線維芽細胞は、SECPに曝露される。SECPのエンドサイトーシスのuptakeに続いて、新鮮な培養培地で細胞を洗浄し、全細胞を電圧固定したアフリカツメガエルの筋細胞を走査して無SECP培地の線維芽細胞の1つに接触させる。筋細胞から膜電流を記録する。対照値に相対して増加または減少した電流は、SECPの神経調節性効果を示す (Morimoto, T. ら. (1995) *Neuron* 15: 689-696)。

#### 【0286】

別法では、SECP活性のためのアッセイは、分泌性の膜結合オルガネラ中のSECP量を測定する。上記したように形質移入された細胞を収集及び溶解する。溶解産物は、例えばショ糖密度勾配遠心法など当業者に公知の方法を用いて分別する。このような方法によって、ゴルジ体、ER、小さな膜結合ベシクル及びその他の分泌オルガネラなどの細胞下成分を単離することができる。分別された及び全ての細胞溶解産物からの免疫沈降は、SECP特異抗体を用いて行ない、免疫沈降したサンプルはSDS-PAGE及び免疫ブロット技術を用いて分析する。全細胞溶解産物中のSECPに関連する分泌オルガネラ中のSECP濃度は、分泌経路を通過するSECPの量に比例する。

#### 【0287】

スクシニルCoAアセトアセテートCoAトランスフェラーゼ活性などの、SECPのコエンザイムAトランスフェラーゼ活性は、アセトアセチルCoAの形成に応じた $A_{310}$ の増加を観察することにより測定できる。アッセイは、Howard, J. B. ら. (1986; *J. Biol. Chem.* 261: 60-65) 及びCorthesy-Theulaz, I. E. ら. (1997; *J. Biol. Chem.* 272: 25659-25667) に記載されているように、67 mMのリチウムアセトアセテート、300  $\mu\text{M}$ のスクシニルCoA及び15 mM  $\text{MgCl}_2$  を50 mM  $\text{Tris-HCl}$ , pH 9.1に溶解したもので行う。

#### 【0288】

別法では、リガンド蛍光強化分光蛍光分析によってリポカリン活性を測定する (Lin ら. (1997) *Molecular Vision* 3: 17)。リガンドの例としては、レチノール (Sigma, St. Louis MO) 及び16-アントリルオキシパルミチン酸 (16-AP) (Molecular Probes Inc., Eugene OR) がある。リガンドを100%エタノールに溶解し、既知の吸光係数 (レチノール: 325 nmで46,000  $\text{A}/\text{M}/\text{cm}$ 、16-AP: 361 nmで8,200  $\text{A}/\text{M}/\text{cm}$ ) を用いてその濃度を推定する。10 mM  $\text{Tris}$  (pH 7.5)、2 mM EDTA、500 mM  $\text{NaCl}$ 中に1  $\mu\text{M}$  SECPのアリコート700  $\mu\text{l}$ を経路長1 cmの石英キュベットに置き、リガンド溶液のアリコート1  $\mu\text{l}$ を添加する。読みが安定するまで各添加後100秒後に蛍光を測定する。リガンド濃度の単位変化毎の蛍光の変化は、SECP活性に比例する。

#### 【0289】

当業者であれば、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うことができるであろう。特定の好適な実施例に基づいて本発明を説明したが、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、分子生物学或いは関連する分野の専門家には明らかな、本明細書に記載の本発明の実施例の様々な改変は、特許請求の範囲に含まれる。

## 【0290】

(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に対する系統的な名称を示す。

## 【0291】

表2は、本発明のポリペプチドに最も近いGenBankの相同体のGenBankの識別番号およびアノテーションを示す。ポリペプチドとそのGenBankの相同体との間の一致を表す確率値スコアも示す。

## 【0292】

表3は、推定上のモチーフおよびドメインを含むポリペプチド配列の構造的な特徴、並びにポリペプチドの分析に用いた方法、アルゴリズム、および検索可能なデータベースを示す。

10

## 【0293】

表4は、ポリヌクレオチド配列の組み立てに用いたcDNA断片のリスト、並びにポリヌクレオチド配列の選択された断片のリストを示す。

## 【0294】

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。

## 【0295】

表6は、表5に示すcDNAライブラリの作製に用いた組織およびベクターを示す付表である。

20

## 【0296】

表7は、本発明のポリヌクレオチド及びポリペプチドの分析に用いたツール、プログラム、及びアルゴリズム、並びにその説明、引用文献、閾値パラメータを示す。

## 【表1】

表1

| インサイト<br>プロジェクトID | ポリペプチド<br>SEQ ID NO: | インサイト<br>ポリペプチドID | ポリヌクレオチド<br>SEQ ID NO: | インサイト<br>ポリヌクレオチドID |
|-------------------|----------------------|-------------------|------------------------|---------------------|
| 7473577           | 1                    | 7473577CD1        | 15                     | 7473577CB1          |
| 7474024           | 2                    | 7474024CD1        | 16                     | 7474024CB1          |
| 2480555           | 3                    | 2480555CD1        | 17                     | 2480555CB1          |
| 3187086           | 4                    | 3187086CD1        | 18                     | 3187086CB1          |
| 1274566           | 5                    | 1274566CD1        | 19                     | 1274566CB1          |
| 1349442           | 6                    | 1349442CD1        | 20                     | 1349442CB1          |
| 1400156           | 7                    | 1400156CD1        | 21                     | 1400156CB1          |
| 1610347           | 8                    | 1610347CD1        | 22                     | 1610347CB1          |
| 187209            | 9                    | 187209CD1         | 23                     | 187209CB1           |
| 2607963           | 10                   | 2607963CD1        | 24                     | 2607963CB1          |
| 412044            | 11                   | 412044CD1         | 25                     | 412044CB1           |
| 638118            | 12                   | 638118CD1         | 26                     | 638118CB1           |
| 743323            | 13                   | 743323CD1         | 27                     | 743323CB1           |
| 1691509           | 14                   | 1691509CD1        | 28                     | 1691509CB1          |

10

20

30

40

【表 2】

表2

| ポリペプチド<br>SEQ ID NO: | インサイト<br>ポリペプチド ID | GenBank<br>ID NO:    | 確率スコア                  | GenBank 相合体                                                                            |
|----------------------|--------------------|----------------------|------------------------|----------------------------------------------------------------------------------------|
| 1                    | 7473577CD1         | g337762<br>g337765   | 5.00E-119<br>5.70E-118 | プロザボン [ヒト]<br>セレプロシド硫酸活性化タンパク質 [ヒト]                                                    |
| 2                    | 7474024CD1         | g35897               | 2.30E-105              | [ヒト] 前駆体RBP<br>Colantuoni, V. ら (1983) Nucleic<br>Acids Res. 11 (22), 7769-7776 (1983) |
| 3                    | 2480555CD1         | g1890108             | 3.10E-245              | [ヒト] リンクオキシダーゼ関連タンパク質<br>Mariani, T.J. ら (1992)<br>Matrix 12:242-248                   |
| 4                    | 3187086CD1         | g3978171<br>g1519052 | 0<br>1.80E-198         | [ハツカネズミ]<br>リンクオキシダーゼ関連タンパク質 2<br>スクシニルCoA:3オキシ酸CoAトランスフェラーゼ前駆体<br>[ヒト]                |
| 6                    | 1349442CD1         | g10437669            | 0                      | [ヒト] 不特定のタンパク質産物                                                                       |
| 7                    | 1400156CD1         | g12833121            | 1.00E-158              | [ハツカネズミ] (AK002841)<br>推定上のタンパク質                                                       |
| 8                    | 1610347CD1         | g12845305            | 0                      | [ハツカネズミ] (AK010097)<br>推定上のタンパク質                                                       |
| 9                    | 187209CD1          | g4200218             | 1.00E-65               | [ヒト] 仮定上のタンパク質                                                                         |
| 10                   | 2607963CD1         | g12958148            | 1.00E-177              | [アカゲザル] (AF245204)<br>味覚芽特異タンパク質前駆体                                                    |
| 11                   | 412044CD1          | g12841382            | 1.00E-95               | [ハツカネズミ] (AK007682) 推定上                                                                |
| 14                   | 1691509CD1         | g6180013             | 3.10E-62               | [ヒト] 後期促進複合体サブユニット 5                                                                   |

【表 3】

【表 4】

| SEQ ID NO. | インサートポリペプチドID | アミノ酸残基 | 潜在的リン酸化部位                                                                                                               | 潜在的グリコシル化部位                    | サイン(signature)配列、ドメイン及びモチーフ                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 分析方法及びデータベース                                                                                                                                                                                                                                                                                                         |
|------------|---------------|--------|-------------------------------------------------------------------------------------------------------------------------|--------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1          | 7473577CD1    | 521    | S284 T83 S138<br>S154 T203 T415<br>T443 S452 T58<br>T149 S414 T415<br>S452                                              | N201 N311                      | シグナルペプチド : M1-P19<br>シグナル分割 : M1-A17<br>PD01469: C438-C469<br>グリコプロテイン タンパク質前駆体<br>PD012321: G181-E256<br>プロサボシン前駆体<br>DM020411 P07602   84-254 : D86-Q210<br>サボシンリピード<br>リボカリン : N32-M45 D34-M45<br>M1-A18 シグナルペプチド<br>M1-A18 シグナル分割<br>リボカリン / サイトリル脂酸結合タンパク質 : R37-C192<br>リボカリンタンパク質 BL00213:<br>A36-D49, I124-V134<br>リボカリンサイン PR00179:<br>A36-K48, I124-Y136, V154-I169<br>リボカリン DM00680/P18902/I-172: E19-Y191<br>スプラクト受容体リピード BL00420C:<br>C134-C144 G311-G365<br>リシロオキシダーゼ銅結合領域タンパク質<br>BL00926D: V538-I584 I584-D620<br>D620-K650 Y652-L691 Q692-A732<br>スプラクト受容体繰り返しドメインサイン<br>G292-L464<br>リシロオキシダーゼサイン PR00074A:<br>L533-L561 L561-G588 R589-T617 T617-C644<br>E646-I674 D675-E703 S704-D731<br>スプラクト受容体サイン PR00258A:<br>I417-Q433 S326-D337 A341-G351 S372-C386<br>D395-N407<br>リシロオキシダーゼ プロテインリン前駆体<br>シグナル6オキシダーゼ オキシドレタクターゼ<br>銅グリコプロテイン相同体 PD012364: D530-I729<br>シグナル分割 M1-G25<br>スカベンジャー受容体 多システインドメイン<br>SRCR: R51-K145 P183-V282 K310-N407<br>S420-S525<br>リシロオキシダーゼ : S529-A732<br>オキシダーゼ : リシン ; リシン ; 銅<br>DM04978 I005063   11-419 : D530-I729 | HMMER<br>SPSCAN<br>BLIMPS_PRODOM<br>BLAST_PRODOM<br>BLAST_DOMO<br>MOTIFS<br>SIGPEPT<br>SPSCAN<br>HMMER_PFM<br>BLIMPS_BLOCKS<br>BLIMPS_PRINTS<br>BLAST_DOMO<br>BLIMPS_BLOCKS<br>BLIMPS_BLOCKS<br>GCprofile<br>BLIMPS_PRINTS<br>BLIMPS_PRINTS<br>BLAST_PRODOM<br>SPSCAN, HMMER<br>HMMER_PFM<br>HMMER_PFM<br>BLAST_DOMO |
| 2          | 7474024CD1    | 201    | S15 S26 T74<br>S137 S196 S64<br>T94 T96 T127<br>T146                                                                    |                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                      |
| 3          | 2480555CD1    | 753    | S97 T104 S221<br>T268 S352 T510<br>S564 S649 T31<br>T68 S115 S120<br>T135 T330 S377<br>T392 T411 S424<br>T493 T527 T617 | N111 N266<br>N390 N481<br>N625 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                                                                                                                                                                                      |

表 3-2

| SEQ ID NO. | インサイト<br>ポリペプチドID | アミノ酸<br>残基 | 潜在的リン酸化部位                                                                                                               | 潜在的グリコ<br>シレ化部位                                          | サイン(signature)配列、ドメイン及びモチーフ                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 分析方法及び<br>データベース                   |
|------------|-------------------|------------|-------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------|
| 4          | 3187086CD1        | 511        | S226 S35 S423<br>T120 T140 T277<br>T346 T440 T452<br>T462 T476 T493                                                     |                                                          | シグナルペプチド : M1-G21<br>コエンザイムAトランスフェラーゼ : K41-Q267, A292-T493<br>CoAトランスフェラーゼタンパク質 :<br>BL01273A: G62-I74, BL01273B: R81-S114<br>BL01273C: E130-G169,<br>BL01273D: D206-P249<br>3-オキソアジバートCoAトランスフェラーゼβ鎖 :<br>DM02057<br>A41771 286-521: K280-V509<br>P42316 2-216: E290-L503<br>P23673 6-221: A292-N506<br>3-オキソアジバートCoAトランスフェラーゼα鎖 :<br>DM02058 A41771 39-284: K41-R279<br>スクニールCoA : 3-オキソCoAトランスフェラーゼ :<br>PD004976: K51-K265, R293-T493<br>シグナルペプチド : M1-Q21<br>シグナル分割 : M1-C16<br>シグナル分割 : M1-A28 | HMER<br>HMER-PFAM<br>BLIMPS-BLOCKS |
| 5          | 1274566CD1        | 99         | S82                                                                                                                     | N88                                                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | HMER<br>SPSCAN                     |
| 6          | 1349442CD1        | 389        | T45 T186 T195<br>S231 T261 S352<br>S110 S116 S121<br>T164 S222 T261<br>S265 T291 S331<br>S345 S364                      | N44 N259 N288<br>N313                                    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | SPSCAN                             |
| 7          | 1400156CD1        | 322        | T51 T60 T76<br>S114 T36 S108<br>T164 S282 S316                                                                          | N106 N119<br>N162 N175<br>N192 N205<br>N251 N280<br>N281 | シグナルペプチド : M1-G26<br>シグナル分割 : M1-G26                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | HMER<br>SPSCAN                     |
| 8          | 1610347CD1        | 587        | S81 S146 S230<br>T234 S281 S306<br>S348 T436 S519<br>T579 S290 S26<br>S56 S146 T167<br>S306 S392 S461<br>S516 T569 T579 | N114 N322<br>N500                                        | シグナルペプチド : M1-A32<br>タンパク質合成経路 - 硫黄システターゼ<br>システターゼオキシダーゼIIIαプロモプテリノゲン<br>ピオチントランスフェラーゼ PD000690:R100-R581                                                                                                                                                                                                                                                                                                                                                                                            | HMER<br>BLAST_PRODOR               |
| 9          | 187209CD1         | 173        | S53 T75 T114<br>T152 T86 S102<br>T152 T170                                                                              |                                                          | シグナル分割 : M1-S53                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | SPSCAN                             |

【表 5】

10

20

30

40

| SEQ ID NO. | インサイト<br>ポリペプチドID | アミノ酸<br>残基 | 潜在的な酸化部位                                                                                                                                                                 | 潜在的グリコ<br>シル化部位                | サイン(signature)配列、ドメイン及びモチーフ                                                                                                                                                                           | 分析方法及び<br>データベース                                                                    |
|------------|-------------------|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| 10         | 2607963CD1        | 325        | S175 S192 S319<br>S56 S58 S155<br>S156                                                                                                                                   |                                | シグナルペプチド : M1-S28<br>シグナル分断 : M1-A26                                                                                                                                                                  | HMER<br>SPSCAN                                                                      |
| 11         | 412044CD1         | 733        | S249 S252 S345<br>T487 T517 T607<br>T730 Y373 S231<br>S253 S475 S531<br>T575 S15 S23<br>S85 S107 T108<br>T126 S185 S190<br>S194 S258 S345<br>S483 T575 S623<br>S680 T713 | N183 N317<br>N508 N636<br>N646 | シグナル分断 : M1-G22<br>PHDフィンガー PHD:V118-R166<br>SETドメイン SET:T323-C453<br>SETドメイン タンパク質 PF00856: D129-P165,<br>V395-E416<br>ASH1 SET ドメイン タンパク質 PD140577:D111-E171<br>ウィルムス腫瘍タンパク質サイン<br>PR00049D:Y66-P80 | SPSCAN<br>HMER_PPFAM<br>HMER_PPFAM<br>BLIMPS_PPFAM<br>BLAST_PRODOR<br>BLIMPS_PRINTS |
| 12         | 638118CD1         | 242        | T18 T139 T131<br>S176 S181                                                                                                                                               |                                | シグナル分断 : M1-P42                                                                                                                                                                                       | SPSCAN                                                                              |
| 13         | 743323CD1         | 153        | T104 T123 S109<br>S67 T108                                                                                                                                               |                                | シグナル分断 : M1-G45                                                                                                                                                                                       | SPSCAN                                                                              |
| 14         | 1691509CD1        | 134        | Y94                                                                                                                                                                      |                                | シグナルペプチド : M1-S21                                                                                                                                                                                     | HMER                                                                                |

【表 6】

表4-1

| ポリヌクレオチド<br>SEQ ID NO: | インサイト ポリヌクレオチドID | 配列長さ | 選択された断片                                         | 配列断片                    | 5'位置 | 3'位置 |
|------------------------|------------------|------|-------------------------------------------------|-------------------------|------|------|
| 15                     | 7473577CB1       | 1566 | 1-273, 334-1170, 1488-1566                      | GNN.g6437516_000004_002 | 1    | 1566 |
| 16                     | 7474024CB1       | 939  | 1-22                                            | 2087293H1 (PANCNOT04)   | 1    | 138  |
| 16                     | 7474024CB1       | 939  | 1-22                                            | 6535251H1 (OVARIN02)    | 278  | 939  |
| 16                     | 7474024CB1       | 939  | 1-22                                            | 6201685H1 (PITUNON01)   | 127  | 681  |
| 16                     | 7474024CB1       | 939  | 1-22                                            | 166267H1 (LIVRNOT01)    | 74   | 429  |
| 17                     | 2480555CB1       | 2785 | 2761-2785, 1-570, 1685-2093, 1317-1480, 762-977 | 1781933R6 (PGANNON02)   | 1910 | 2356 |
| 17                     | 2480555CB1       | 2785 | 2761-2785, 1-570, 1685-2093, 1317-1480, 762-977 | stretcher.fasta         | 68   | 2329 |
| 17                     | 2480555CB1       | 2785 | 2761-2785, 1-570, 1685-2093, 1317-1480, 762-977 | 7262753H1 (UTRETM01)    | 1    | 574  |
| 17                     | 2480555CB1       | 2785 | 2761-2785, 1-570, 1685-2093, 1317-1480, 762-977 | 4204281T6 (BRAITUT29)   | 2028 | 2785 |
| 18                     | 3187086CB1       | 1733 | 1-391, 1706-1733, 816-1071                      | 4291392T6 (BRABDIR01)   | 1182 | 1718 |
| 18                     | 3187086CB1       | 1733 | 1-391, 1706-1733, 816-1071                      | 4405672F6 (PROSDIT01)   | 495  | 1010 |
| 18                     | 3187086CB1       | 1733 | 1-391, 1706-1733, 816-1071                      | 3187086R6 (THYMNON04)   | 47   | 585  |
| 18                     | 3187086CB1       | 1733 | 1-391, 1706-1733, 816-1071                      | 2005793H1 (TESTNOT03)   | 1615 | 1733 |

【表 7】

10

20

30

40

表4-2

| ポリヌクレオチド<br>SEQ ID NO: | インサイト ポリヌクレオチドID | 配列長さ | 選択された断片                          | 配列断片                  | 5'位置 | 3'位置 |
|------------------------|------------------|------|----------------------------------|-----------------------|------|------|
| 18                     | 3187086CB1       | 1733 | 1-391,<br>1706-1733,<br>816-1071 | 5100508H1 (PROSTUS20) | 1005 | 1256 |
| 18                     | 3187086CB1       | 1733 | 1-391,<br>1706-1733,<br>816-1071 | 3747662H1 (UTRSNOT18) | 1    | 297  |
| 18                     | 3187086CB1       | 1733 | 1-391,<br>1706-1733,<br>816-1071 | 6065570H1 (BRAENOT02) | 920  | 1204 |
| 19                     | 1274566CB1       | 1148 | 621-666, 1-<br>154               | 1274566F6 (TESTTUT02) | 525  | 1132 |
| 19                     | 1274566CB1       | 1148 | 621-666, 1-<br>154               | SCKA01270V1           | 345  | 1035 |
| 19                     | 1274566CB1       | 1148 | 621-666, 1-<br>154               | 1274566F6 (TESTTUT02) | 1    | 419  |
| 19                     | 1274566CB1       | 1148 | 621-666, 1-<br>154               | SCKA01551V1           | 526  | 1148 |
| 20                     | 1349442CB1       | 1213 | 1019-1213                        | SBA02792F1            | 274  | 855  |
| 20                     | 1349442CB1       | 1213 | 1019-1213                        | SBA01641F1            | 53   | 713  |
| 20                     | 1349442CB1       | 1213 | 1019-1213                        | 1349442H1 (LATRTUT02) | 1    | 253  |
| 20                     | 1349442CB1       | 1213 | 1019-1213                        | 136361R1 (SYNORAB01)  | 662  | 1213 |
| 21                     | 1400156CB1       | 2298 | 1-21, 999-<br>1018               | 70222572V1            | 1    | 488  |
| 21                     | 1400156CB1       | 2298 | 1-21, 999-<br>1018               | g675353               | 1693 | 2296 |
| 21                     | 1400156CB1       | 2298 | 1-21, 999-<br>1018               | 2496967T6 (ADRETUT05) | 936  | 1515 |
| 21                     | 1400156CB1       | 2298 | 1-21, 999-<br>1018               | g1422205              | 1713 | 2298 |
| 21                     | 1400156CB1       | 2298 | 1-21, 999-<br>1018               | SXAE05185V1           | 447  | 949  |
| 21                     | 1400156CB1       | 2298 | 1-21, 999-<br>1018               | 1400156F6 (BRAITUT08) | 1069 | 1674 |
| 21                     | 1400156CB1       | 2298 | 1-21, 999-<br>1018               | 231268F1 (SINTNOT02)  | 1505 | 2295 |
| 21                     | 1400156CB1       | 2298 | 1-21, 999-<br>1018               | SXAE01140V1           | 503  | 1020 |
| 22                     | 1610347CB1       | 2079 | 1098-1330                        | 1438856F1 (PANCNOT08) | 1464 | 2079 |
| 22                     | 1610347CB1       | 2079 | 1098-1330                        | 3204803F6 (PENCNOT03) | 1    | 575  |
| 22                     | 1610347CB1       | 2079 | 1098-1330                        | 1981287R6 (LUNGUT03)  | 1367 | 1902 |
| 22                     | 1610347CB1       | 2079 | 1098-1330                        | 744211R1 (BRAITUT01)  | 1076 | 1655 |

【表 8】

10

20

30

40

表4-3

| ポリヌクレオチド<br>SEQ ID NO: | インサイト ポリヌクレオチドID | 配列長さ | 選択された断片                           | 配列断片                  | 5'位置 | 3'位置 |
|------------------------|------------------|------|-----------------------------------|-----------------------|------|------|
| 22                     | 1610347CB1       | 2079 | 1098-1330                         | 921460H1 (RATRNOT02)  | 852  | 1142 |
| 22                     | 1610347CB1       | 2079 | 1098-1330                         | SCDA08508V1           | 470  | 1036 |
| 23                     | 187209CB1        | 846  | 1-333                             | 187209CT1 (CARDNOT01) | 1    | 846  |
| 24                     | 2607963CB1       | 1148 | 1-1148                            | SXBC00624V1           | 1    | 693  |
| 24                     | 2607963CB1       | 1148 | 1-1148                            | 4910876F6 (THYMDIT01) | 586  | 1148 |
| 25                     | 412044CB1        | 3076 | 1-157,<br>1900-3076,<br>1111-1285 | 419916T6 (BRSTNOT01)  | 2482 | 3076 |
| 25                     | 412044CB1        | 3076 | 1-157,<br>1900-3076,<br>1111-1285 | 869590R1 (LUNGAST01)  | 1281 | 1821 |
| 25                     | 412044CB1        | 3076 | 1-157,<br>1900-3076,<br>1111-1285 | 412044R1 (BRSTNOT01)  | 2168 | 2779 |
| 25                     | 412044CB1        | 3076 | 1-157,<br>1900-3076,<br>1111-1285 | 452392R7 (TLYMNOT02)  | 1908 | 2445 |
| 25                     | 412044CB1        | 3076 | 1-157,<br>1900-3076,<br>1111-1285 | 2503204T6 (CONUTUT01) | 495  | 1097 |
| 25                     | 412044CB1        | 3076 | 1-157,<br>1900-3076,<br>1111-1285 | 1678531F6 (STOMFET01) | 1095 | 1625 |
| 25                     | 412044CB1        | 3076 | 1-157,<br>1900-3076,<br>1111-1285 | 3440195H1 (PENCNOT06) | 450  | 689  |
| 25                     | 412044CB1        | 3076 | 1-157,<br>1900-3076,<br>1111-1285 | 2948371T6 (KIDNFET01) | 1689 | 2319 |
| 25                     | 412044CB1        | 3076 | 1-157,<br>1900-3076,<br>1111-1285 | 269165F1 (HNT2NOT01)  | 11   | 605  |
| 25                     | 412044CB1        | 3076 | 1-157,<br>1900-3076,<br>1111-1285 | 2503204F6 (CONUTUT01) | 1    | 416  |
| 25                     | 412044CB1        | 3076 | 1-157,<br>1900-3076,<br>1111-1285 | 412044F1 (BRSTNOT01)  | 2436 | 3075 |
| 25                     | 412044CB1        | 3076 | 1-157,<br>1900-3076,<br>1111-1285 | 2836520F6 (TLYMNOT03) | 679  | 1212 |

【表 9】

10

20

30

40

表4-4

| ポリスクレオチド<br>SEQ ID NO: | インサイト ポリスクレオチドID | 配列長さ | 選択された断片             | 配列断片                  | 5'位置 | 3'位置 |
|------------------------|------------------|------|---------------------|-----------------------|------|------|
| 26                     | 638118CB1        | 2102 | 1-288,<br>1342-1361 | 322585R6 (EOSIHET02)  | 1714 | 2102 |
| 26                     | 638118CB1        | 2102 | 1-288,<br>1342-1361 | 2476940H1 (SMCANOT01) | 1    | 224  |
| 26                     | 638118CB1        | 2102 | 1-288,<br>1342-1361 | 1440105R1 (THYRNOT03) | 1549 | 2102 |
| 26                     | 638118CB1        | 2102 | 1-288,<br>1342-1361 | 1856507F6 (PROSNOT18) | 983  | 1667 |
| 26                     | 638118CB1        | 2102 | 1-288,<br>1342-1361 | 1534331H1 (SPINNOT04) | 340  | 634  |
| 26                     | 638118CB1        | 2102 | 1-288,<br>1342-1361 | 638118R1 (BRSTNOT03)  | 607  | 1335 |
| 26                     | 638118CB1        | 2102 | 1-288,<br>1342-1361 | 1891046H1 (BLADTUT07) | 196  | 550  |
| 27                     | 743323CB1        | 807  | 1-464               | 3836037F6 (PANCNOT17) | 280  | 807  |
| 27                     | 743323CB1        | 807  | 1-464               | 950902H1 (PANCNOT05)  | 1    | 234  |
| 27                     | 743323CB1        | 807  | 1-464               | 4077778F6 (PANCNOT19) | 56   | 481  |
| 28                     | 1691509CB1       | 1049 | 1-519,<br>1021-1049 | 71149648V1            | 497  | 1049 |
| 28                     | 1691509CB1       | 1049 | 1-519,<br>1021-1049 | 1999452R6 (BRSTTUT03) | 1    | 527  |

【表 10】

10

20

30

40

表5

| ポリスクレオチド<br>SEQ ID NO: | インサイト プロジェクトID | 代表的ライブラリ  |
|------------------------|----------------|-----------|
| 16                     | 7474024CB1     | LIVRTUT04 |
| 17                     | 2480555CB1     | SMCANOT01 |
| 18                     | 3187086CB1     | TESTNOT03 |
| 19                     | 1274566CB1     | TESTTUT02 |
| 20                     | 1349442CB1     | SYNORAB01 |
| 21                     | 1400156CB1     | PANCDIT01 |
| 22                     | 1610347CB1     | COLNTUT06 |
| 23                     | 187209CB1      | CARDNOT01 |
| 24                     | 2607963CB1     | LUNGTUT07 |
| 25                     | 412044CB1      | BRSTNCT01 |
| 26                     | 638118CB1      | THYRNOT03 |
| 27                     | 743323CB1      | PANCNOT05 |
| 28                     | 1691509CB1     | PROSTUS23 |

【表 1 1】

10

20

30

【表 1 2】

| ライブラリ     | ベクター        | ライブラリの説明    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-----------|-------------|-------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|           |             | PBLUESCRIPT | ライブラリは、自動車事故で死亡した 56 歳の白人女性の乳房組織から単離した RNA を用いて作製した。                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| BRSTNOT01 | PBLUESCRIPT |             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| CARDNOT01 | PBLUESCRIPT |             | ライブラリは、銃創で死亡した 65 歳の白人男性の心筋から単離した RNA を用いて作製した。                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| COLNTUT06 | pINCY       |             | ライブラリは、45 歳の白人女性から全結腸切除術及び全腹式子宮摘出術の際に得られた結腸癌組織から単離した RNA で作製した。病理学的には、盲腸に集塊を形成している浸潤性グレード 2 結腸腺癌を示していた。患者の既往症には、直腸及び肛門の良性新生物、多発性硬化症及び僧帽弁疾患があった。家族歴には、I 型糖尿病、脳血管障害、アテローム性冠動脈疾患、悪性皮膚新生物、高血圧症、及び結腸の悪性新生物があった。                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| LIVRTUT04 | pINCY       |             | ライブラリは、50 歳の白人男性の部分肝切除の際に採取した腎腫瘍組織から単離した RNA を用いて作製した。病理学的には、グレード 3-4 のヘパトームを示し、腫瘍を形成していた。患者の既往症には、良性高血圧症、及び肝炎があった。B 型肝炎の核抗原及び B 型肝炎の表面抗原が患者に現われていた。                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| LUNGTUT07 | pINCY       |             | ライブラリは、50 歳の白人男性の部分肺切除の際に上葉から採取した肺腫瘍組織から単離した RNA を用いて作製した。病理学的には、浸潤性グレード 4 の扁平上皮癌を示していた。患者には、喫煙の習慣があった。家族歴には皮膚癌があった。                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| PANCDIT01 | PBLUESCRIPT |             | ライブラリは、銃創で死亡した 15 歳の白人男性から採取した膵臓組織から単離した RNA を用いて作製した。家族歴にはインシュリン治療された I 型糖尿病がある。                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| PANCNOT05 | PSPORT1     |             | ライブラリは、脳無酸素症で死亡した 2 歳のヒスパニックの男児から採取した膵臓組織から単離した RNA を用いて作製した。                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| PROSTUS23 | pINCY       |             | このサブトラクトされた前立腺組織は、プールされていた前立腺ライブラリよりの $1 \times 10^6$ 個のクローンで、2 ラウンドの subtractive hybridization を行った、プールされた前立腺組織ライブラリからの独立したクローン $1 \times 10^6$ 個から作製された。サブトラクションのための開始ライブラリは、リンパ節切除を伴う前立腺切除の際に 58 歳 (A)、61 歳 (B)、66 歳 (C) 及び 68 歳 (D) の白人男性から採取した前立腺腫瘍から単離した mRNA を用いた 4 つの前立腺腫瘍ライブラリから得られた同数のクローンをプールして作製した。病理学的には、すべてのドナーが腺癌を示した。ドナー A の病歴には高値の PSA、硬化、ニコチン中毒；ドナー B の病歴には高値の PSA、硬化、前立腺肥大、腎不全、変形性関節症、腎動脈狭窄、良性 HTN、血小板減少症、高脂質血症、タバコニコル依存症及び C 型肝炎（保菌者）；ドナー C の病歴には高値の PSA、硬化及びタバコ依存症；そしてドナー D の病歴には高値の PSA、硬化、高コレステロール血症及び腎臓結石がある。サブトラクションのためのハイブリダイゼーションプロローブは、3 人の異なるドナーからの前立腺組織、前立腺上皮細胞、及び前立腺の間質の線維芽細胞から得られた 3 つの前立腺組織ライブラリから同数の cDNA クローンをプールして作製した。subtractive hybridization 条件は、Swarcop ら, (1991) Nucleic Acids Res. 19:1954 及び Bonaldo ら, Genome Research (1996) 6:791 の方法論に基づいた。 |

【表 1 3】

表6-2

| ライブラリ     | ベクター        | ライブラリの説明                                                                                                                                      |
|-----------|-------------|-----------------------------------------------------------------------------------------------------------------------------------------------|
| SMCANOT01 | pINCY       | ライブラリは、ある男性の心臓移植の際に取出された心臓から採取した大動脈平滑筋細胞株から単離した RNA を用いて作製した。                                                                                 |
| SYNORAB01 | PBLUESCRIPT | ライブラリは、リウマチ様関節炎の 68 歳の白人女性から採取した滑膜組織から単離した RNA を用いて作製した。                                                                                      |
| TESTNOT03 | PBLUESCRIPT | ライブラリは、肝疾患で死亡した 37 歳の白人男性から採取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、肝硬変、黄疸及び肝不全が含まれる。                                                              |
| TESTTUT02 | pINCY       | ライブラリは、31 歳の白人男性から片側睾丸摘除の際に採取した睾丸腫瘍組織から単離した RNA を用いて作製した。病理学的には、胎児性癌を示していた。                                                                   |
| THYRNOT03 | pINCY       | ライブラリは、28 歳の白人女性の左甲状腺から完全甲状腺切除術の際に採取された甲状腺組織から単離された RNA で作製した。病変は、左甲状腺に存在する腺腫状過形成の小結節を示していた。関連腫瘍組織の病変は、左甲状腺によく被包化された腫瘍集塊を形成している優性小胞状腺腫を示していた。 |

| プログラム名                  | 説明                                                                                                                  | 引用文献                                                                                                                                                                                            | パラメーター閾値                                                                                                                 |
|-------------------------|---------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|
| ABI FACTURA             | 核酸配列においてベクター配列を除去して不明瞭な塩基をマスキングするプログラム                                                                              | PE Biosystems, Foster City, CA.                                                                                                                                                                 |                                                                                                                          |
| ABI/PARACEL FDF         | アミノ酸又は核酸配列の比較や注釈付けに有用な高速データベース                                                                                      | PE Biosystems, Foster City, CA, Paracel Inc., Pasadena, CA                                                                                                                                      | ミスマッチ<50%                                                                                                                |
| ABI AutoAssembler BLAST | 核酸配列を構築するプログラム                                                                                                      | PE Biosystems, Foster City, CA.                                                                                                                                                                 |                                                                                                                          |
|                         | アミノ酸及び核酸配列の配列類似性検索に有用な基本ローカルアライメント検索ツール。BLAST には、blastp 及び blastn、blastx、tblastn、tblastx の5つの機能がある。                 | Altschul, S.F.ら(1990) J. Mol. Biol. 215:403-410、Altschul, S.F.ら(1997) Nucleic Acids Res. 25: 3389-3402.                                                                                         | ESTs: 確率値= 1.0E-8 以下<br>完全長配列: 確率値= 1.0E-10 以下                                                                           |
| FASTA                   | 間合わせの配列と同種の配列のグループとの類似性を検索する Pearson と Lipman のアルゴリズム。FASTA は、少なくとも5つの機能 fasta、tfasta、fastx、tfastx、及び ssearch からなる。 | Pearson, W.R. and D.J. Lipman (1988) Proc.Natl. Acad. Sci. 85:2444-2448、Pearson, W.R.(1990) Methods Enzymol. 183: 63-98、Smith, T.F. and M. S. Waterman (1981) Adv.Appl. Math. 2:482-489.        | EST: fasta E 値=1.06E-6<br>アセンブル済の EST: fasta 同一性=95%以上<br>一致長=200 塩基以上<br>fastx E 値=1.0E-8 以下<br>完全長配列: fastx スコア=100 以上 |
| BLIMPS                  | 配列を、BLOCKS 及び PRINTS データベースにおける配列と一致させて、遺伝子ファミリー及び配列相同性、構造的フィンガープリント領域を検索するプロットトサチャー                                | Ittenkoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S.Henikoff (1996) Methods Enzymol. 266:88-105、Artwood, T.K.ら(1997) J. Chem. Inf. Comput. Sci. 37: 417-424. | スコア=1000 以上<br>スコア/強度=0.75 以上<br>確率値=1.0E-3 以下 (適用可能な場合)                                                                 |
| HMMER                   | PFAM 等の、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に基づいたデータベースに対する問合せ配列を検索するアルゴリズム                                                 | Krogh, A.ら(1994) J. Mol. Biol., 235:1501-1531、Sonnhammer, E.L.L.ら(1988) Nucleic Acids Res. 26:320-322.                                                                                          | 個々のタンパク質ファミリーによって、スコア=(PFAM ヒットに) 10-50 ビット                                                                              |

10

20

30

40

【表 1 4】

表7-2

| プログラム名      | 説明                                                                                                         | 引用文献                                                                                                                                                                                      | パラメーター閾値                                                                |
|-------------|------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|
| ProfileScan | Prosites で定義された配列パターンと一致するタンパク質配列において構造的モチーフ及び配列モチーフを検索するアルゴリズム                                            | Gribskov, M.ら(1988) CABIOS 4:61-66.<br>Gribskov, M.ら(1989) Methods Enzymol. 183:146-159, Bairoch, A.ら(1997) Nucleic Acids Res. 25:217-221.                                                | ノーマライズされた質のスコア≧特定の Prosites モチーフのための GCG 指定された「HIGH」値<br>通常、スコア=1.4-2.1 |
| Phred       | 高い感度及び確率で自動配列決定装置の痕跡を検査する塩基呼び出しアルゴリズム                                                                      | Ewing, B.ら(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利用に基づくプログラムである SWAT や CrossMatch を含む Phrap 改訂アセンブリプログラム。配列の相同性の検索及び DNA 配列の構築に有用。 | Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489, Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197, Green, P., University of Washington, Seattle, WA.       | スコア=120 以上<br>一致長さ=56 以上                                                |
| Consed      | Phrap アセンブリを編集及び観察するグラフィックツール                                                                              | Gordon, D.ら(1998) Genome Res. 8:195-202.                                                                                                                                                  |                                                                         |
| SPScan      | 分泌シグナルペプチドの存在するかどうかタンパク質配列をスキャンするウェイトマトリックス分析プログラム                                                         | Nielson, H.ら(1997) Protein Engineering 10: 1-6,<br>Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.                                                                                 | スコア=3.5 以上                                                              |
| TMAP        | 重み付けマトリックスを用いたタンパク質配列上の膜貫通セグメントの明確化および向きを決定するためのプログラム                                                      | Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192, Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.                                                                        |                                                                         |
| TMHMMER     | 隠れマルコフモデル (HMM) を用いたタンパク質配列上の膜貫通セグメントの明確化および向きを決定するためのプログラム                                                | Sonnhammer, E.L. ら, (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow 5, eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182. |                                                                         |
| Motifs      | Prosites で定義されたものと一致するパターンに対してアミノ酸配列を検索するプログラム                                                             | Bairoch, A.ら (1997) Nucleic Acids Res. 25:217-221, Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.                                        |                                                                         |

10

20

30

40

## 【国際公開パンフレット】

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

(19) World Intellectual Property Organization  
International Bureau(43) International Publication Date  
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number  
WO 01/79291 A2

(51) International Patent Classification: C07K 14/705

(21) International Application Number: PCT/US01/11861

(22) International Filing Date: 11 April 2001 (11.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

|            |                            |    |
|------------|----------------------------|----|
| 60/197,854 | 14 April 2000 (14.04.2000) | US |
| 60/202,373 | 4 May 2000 (04.05.2000)    | US |
| 60/205,899 | 18 May 2000 (18.05.2000)   | US |
| 60/210,155 | 1 June 2000 (01.06.2000)   | US |
| 60/209,401 | 1 June 2000 (01.06.2000)   | US |

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

|          |                            |
|----------|----------------------------|
| US       | 60/197,854 (CIP)           |
| Filed on | 14 April 2000 (14.04.2000) |
| US       | 60/202,373 (CIP)           |
| Filed on | 4 May 2000 (04.05.2000)    |
| US       | 60/205,899 (CIP)           |
| Filed on | 18 May 2000 (18.05.2000)   |
| US       | 60/210,155 (CIP)           |
| Filed on | 1 June 2000 (01.06.2000)   |
| US       | 60/209,401 (CIP)           |
| Filed on | 1 June 2000 (01.06.2000)   |

(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): GRIFFIN, Jennifer, A. [US/US]; 181 Irene Court #11, Belmont, CA 94002 (US). YAO, Monique, G. [US/US]; 111 Frederick Court, Mountain View, CA 94043 (US). BRUNS, Christopher, M. [US/US]; 575 S. Rengstorff Avenue #126, Mountain View, CA 94040 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). DELEGEANE, Angelo, M. [US/US]; 594 Angus Drive, Milpitas, CA 95035 (US). HAFALIA, April [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). PATTERSON,

Chandra [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). POLICKY, Jennifer, L. [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US). TRIBOULEY, Catherine, M. [FR/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). NGUYEN, Danna, B. [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). LAL, Preeti [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #17, Mountain View, CA 94040 (US). LU, Dyung, Aina, M. [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). BATRA, Sajeev [US/US]; 555 El Camino Real #709, Sunnyvale, CA 94087 (US). AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). REDDY, Roopa [IN/US]; 1233 W. McKinley Drive #3, Sunnyvale, CA 94086 (US). AZIMZAI, Yalda [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US).

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

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, 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.

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

## Published:

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

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

(54) Title: SECRETED PROTEINS

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

WO 01/79291 A2

WO 01/79291

PCT/US01/11861

**SECRETED PROTEINS****TECHNICAL FIELD**

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

**BACKGROUND OF THE INVENTION**

Protein transport and secretion are essential for cellular function. Protein transport is mediated by a signal peptide located at the amino terminus of the protein to be transported or secreted. The signal peptide is comprised of about ten to twenty hydrophobic amino acids which target the nascent protein from the ribosome to a particular membrane bound compartment such as the endoplasmic reticulum (ER). Proteins targeted to the ER may either proceed through the secretory pathway or remain in any of the secretory organelles such as the ER, Golgi apparatus, or lysosomes. Proteins that transit through the secretory pathway are either secreted into the extracellular space or retained in the plasma membrane. Proteins that are retained in the plasma membrane contain one or more transmembrane domains, each comprised of about 20 hydrophobic amino acid residues.

Secreted proteins are generally synthesized as inactive precursors that are activated by post-translational processing events during transit through the secretory pathway. Such events include glycosylation, proteolysis, and removal of the signal peptide by a signal peptidase. Other events that may occur during protein transport include chaperone-dependent unfolding and folding of the nascent protein and interaction of the protein with a receptor or pore complex. Examples of secreted proteins with amino terminal signal peptides are discussed below and include proteins with important roles in cell-to-cell signaling. Such proteins include transmembrane receptors and cell surface markers, extracellular matrix molecules, cytokines, hormones, growth and differentiation factors, enzymes, neuropeptides, and vasomediators. (Reviewed in Alberts, B. et al. (1994) Molecular Biology of The Cell, Garland Publishing, New York, NY, pp. 557-560, 582-592.)

Cell surface markers include cell surface antigens identified on leukocytic cells of the immune system. These antigens have been identified using systematic, monoclonal antibody (mAb)-based "shot gun" techniques. These techniques have resulted in the production of hundreds of mAbs directed against unknown cell surface leukocytic antigens. These antigens have been grouped into "clusters of differentiation" based on common immunocytochemical localization patterns in various differentiated and undifferentiated leukocytic cell types. Antigens in a given cluster are presumed to

WO 01/79291

PCT/US01/11861

identify a single cell surface protein and are assigned a "cluster of differentiation" or "CD" designation. Some of the genes encoding proteins identified by CD antigens have been cloned and verified by standard molecular biology techniques. CD antigens have been characterized as both transmembrane proteins and cell surface proteins anchored to the plasma membrane via covalent attachment to fatty acid-containing glycolipids such as glycosylphosphatidylinositol (GPI). (Reviewed in Barclay, A. N. et al. (1995) The Leucocyte Antigen Facts Book, Academic Press, San Diego, CA, pp. 17-20.)

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

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

Growth and differentiation factors are secreted proteins which function in intercellular

WO 01/79291

PCT/US01/11861

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.)

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

WO 01/79291

PCT/US01/11861

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. Arkininstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 194; 252; 284; 55; 111).

Neuropeptides have also been shown to have roles in nociception (pain). Vasoactive intestinal peptide appears to play an important role in chronic neuropathic pain. Nociceptin, an endogenous 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. (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).

Prosaposin, also called SAP precursor, has been identified as the major product secreted by Sertoli cells and in several other body fluids including seminal plasma, milk, and cerebrospinal fluid. Human sphingolipidosis, which may occur as a result of mutations in the prosaposin gene, marks the significance of prosaposin in human physiology (Kishimoto, Y. et al. (1992) J. Lipid Res. 33:1255-1267). Prosaposin secreted from the cell may participate in sphingolipid transport and signalling (Hiesberger, T. et al. (1998) EMBO J. 17:4617-4625). Prosaposin gains entry to cells by at least three independent mechanisms, including the mannose-6-phosphate receptor, the mannose receptor, and the low density lipoprotein receptor-related protein, a multifunctional endocytic receptor that is expressed on most cells (Hiesberger, T. et al., supra). Prosaposin is active in a variety of neuronal cells including hippocampal neurons, spinal cord alpha-motor neurons, cerebellar granule neurons, and neuroblastoma cells, in each of which it stimulates neurite outgrowth and prevents cell death. Prosaposin may have a role in myelin formation (Madar-Shapiro, L. et al. (1999) Biochem. J. 337:433-443). In addition to its signal and transport roles, prosaposin may be proteolytically cleaved within the cell to form saposins -A, -B, -C, and -D, (also called sphingolipid activator proteins or

WO 01/79291

PCT/US01/11861

SAP), which are required to activate lysosomal enzymes involved in glycosphingolipid metabolism. Saposins accumulate in tissues of lysosomal storage disease patients (Kishimoto, *supra*). Saposin B stimulates the hydrolysis of a wide variety of substrates including cerebroside sulfate, GM1 ganglioside, and globotriaosylceramide. Human saposin B deficiency, transmitted as an autosomal

5 recessive trait, results in tissue accumulation of cerebroside sulfate and a clinical picture resembling metachromatic leukodystrophy, an inherited lysosomal storage disease characterized by progressive demyelination leading to severe neurological symptoms. The disease is marked by mRNA that differs from the normal sequence at only one base, a C→T transition in the 23rd codon of saposin B resulting in a threonine to isoleucine amino acid substitution. This base change results in the

10 replacement of a polar amino acid, threonine, with a nonpolar isoleucine (Kretz, K. et al. (1990) Proc. Natl. Acad. Sci. USA 7:2541-2544).

Lipocalins are important transport molecules. Each lipocalin associates with a particular ligand and delivers that ligand to appropriate target sites within the organism. Retinol-binding protein (RBP), one of the best characterized lipocalins, transports retinol from stores within the liver

15 to target tissues. Apolipoprotein D (apo D), a component of high density lipoproteins (HDLs) and low density lipoproteins (LDLs), functions in the targeted collection and delivery of cholesterol throughout the body. Lipocalins also are involved in cell regulatory processes. Apo D, which is identical to gross-cystic-disease-fluid protein (GCDFP)-24, is a progesterone/pregnenolone-binding protein expressed at high levels in breast cyst fluid. Secretion of apo D in certain human breast

20 cancer cell lines is accompanied by reduced cell proliferation and progression of cells to a more differentiated phenotype. Similarly, apo D and another lipocalin,  $\alpha_1$ -acid glycoprotein (AGP), are involved in nerve cell regeneration. AGP is also involved in anti-inflammatory and immunosuppressive activities. AGP is one of the positive acute-phase proteins (APP); circulating levels of AGP increase in response to stress and inflammatory stimulation. AGP accumulates at sites

25 of inflammation where it inhibits platelet and neutrophil activation and inhibits phagocytosis. The immunomodulatory properties of AGP are due to glycosylation. AGP is 40% carbohydrate, making it unusually acidic and soluble. The glycosylation pattern of AGP changes during acute-phase response, and deglycosylated AGP has no immunosuppressive activity (Flower (1994) FEBS Lett. 354:7-11; Flower, *supra*).

30 Lipocalins are used as diagnostic and prognostic markers in a variety of disease states. The plasma level of AGP is monitored during pregnancy and in diagnosis and prognosis of conditions including cancer chemotherapy, renal disfunction, myocardial infarction, arthritis, and multiple sclerosis. RBP is used clinically as a marker of tubular reabsorption in the kidney, and apo D is a marker in gross cystic breast disease (Flower (1996) *supra*).

35 The discovery of new secreted proteins and the polynucleotides encoding them satisfies a

WO 01/79291

PCT/US01/11861

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.

5

#### 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," and "SECP-14."

10 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:1-14, 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:1-14, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group  
15 consisting of SEQ ID NO:1-14, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-14.

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  
20 group consisting of SEQ ID NO:1-14, 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:1-14, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14.  
25 In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-14. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:15-28.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group  
30 consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, 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:1-14, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, and d) an immunogenic fragment of a polypeptide having an amino  
35 acid sequence selected from the group consisting of SEQ ID NO:1-14. In one alternative, the

WO 01/79291

PCT/US01/11861

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:1-14, 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:1-14, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14. 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:1-14, 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:1-14, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14.

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:15-28, 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:15-28, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:15-28, 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:15-28, 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

WO 01/79291

PCT/US01/11861

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:15-28, 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:15-28, 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:1-14, 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:1-14, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-14. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional SECP, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, 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:1-14, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the

WO 01/79291

PCT/US01/11861

group consisting of SEQ ID NO:1-14. 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

5 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:1-14, b) a naturally occurring

10 polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14. The method comprises a) exposing a sample comprising the

15 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:1-14, 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:1-14, c) a biologically active fragment of a polypeptide having an

20 amino acid sequence selected from the group consisting of SEQ ID NO:1-14, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, 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:1-14, c) a biologically active fragment of a

25

WO 01/79291

PCT/US01/11861

polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing  
 5 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  
 10 altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:15-28, 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  
 15 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:15-28, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a  
 20 polynucleotide sequence selected from the group consisting of SEQ ID NO:15-28, 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  
 25 polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:15-28, 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:15-28, 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  
 30 polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

WO 01/79291

PCT/US01/11861

**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 genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.

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

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

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

**DESCRIPTION OF THE INVENTION**

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

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

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be

WO 01/79291

PCT/US01/11861

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 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,

WO 01/79291

PCT/US01/11861

isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

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

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

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of SECP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of SECP either by 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')<sub>2</sub>, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind SECP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

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

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having 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

WO 01/79291

PCT/US01/11861

introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

5 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.

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

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. 15 Compositions comprising polynucleotide sequences encoding SECP or fragments of SECP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl 20 sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

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

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of 30 the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

|    | Original Residue | Conservative Substitution |
|----|------------------|---------------------------|
| 35 | Ala              | Gly, Ser                  |
|    | Arg              | His, Lys                  |

WO 01/79291

PCT/US01/11861

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

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

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

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

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

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

A "fragment" is a unique portion of SECP or the polynucleotide encoding SECP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected

WO 01/79291

PCT/US01/11861

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.

5 A fragment of SEQ ID NO:15-28 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:15-28, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:15-28 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish  
10 SEQ ID NO:15-28 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:15-28 and the region of SEQ ID NO:15-28 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:1-14 is encoded by a fragment of SEQ ID NO:15-28. A fragment of SEQ ID NO:1-14 comprises a region of unique amino acid sequence that specifically identifies  
15 SEQ ID NO:1-14. For example, a fragment of SEQ ID NO:1-14 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-14. The precise length of a fragment of SEQ ID NO:1-14 and the region of SEQ ID NO:1-14 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  
20 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  
25 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  
30 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  
35 follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue

WO 01/79291

PCT/US01/11861

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/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

*Matrix: BLOSUM62*

*Reward for match: 1*

*Penalty for mismatch: -2*

*Open Gap: 5 and Extension Gap: 2 penalties*

*Gap x drop-off: 50*

*Expect: 10*

*Word Size: 11*

*Filter: on*

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

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to

WO 01/79291

PCT/US01/11861

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

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

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

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

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

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

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

WO 01/79291

PCT/US01/11861

"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*, 2<sup>nd</sup> 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 stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g.,  $C_{50}$  or  $R_{50}$  analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid

WO 01/79291

PCT/US01/11861

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 sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

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

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

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

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

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

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

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

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

35 "Post-translational modification" of an SECP may involve lipidation, glycosylation,

WO 01/79291

PCT/US01/11861

phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of SECP.

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

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

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

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

WO 01/79291

PCT/US01/11861

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, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

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

WO 01/79291

PCT/US01/11861

instead of deoxyribose.

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

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

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

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

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and 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 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 on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic

WO 01/79291

PCT/US01/11861

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 a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in  
 5 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, transformation or transjugation. Techniques for transferring the DNA of the present invention  
 10 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 the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-  
 15 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 "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have  
 20 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 another. The resulting polypeptides will generally have significant amino acid identity relative to  
 25 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 propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-  
 30 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  
 35 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence

WO 01/79291

PCT/US01/11861

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 motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are secreted proteins. For example, SEQ ID NO:4 is 72% identical to human succinyl CoA:3-oxoacid CoA transferase precursor (GenBank ID g1519052) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The

WO 01/79291

PCT/US01/11861

BLAST probability score is  $1.8e-198$ , which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:4 also contains a coenzyme A transferase 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

5 BLIMPS and BLAST analyses provide further corroborative evidence that SEQ ID NO:4 is a coenzyme A transferase, such as succinyl CoA:3-oxoacid CoA transferase. SEQ ID NO:1-3 and SEQ ID NO:5-14 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-14 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

10 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

15 of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:15-28 or that distinguish between SEQ ID NO:15-28 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

20 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 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, 2087293H1 is the

25 identification number of an Incyte cDNA sequence, and PANCNOT04 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., SCKA01270V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g675353) which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to

30 coding regions predicted by Genscan analysis of genomic DNA. For example, GNN.g6437516\_000004\_002 is the identification number of a Genscan-predicted coding sequence, with g6437516 being the GenBank identification number of the sequence to which Genscan was applied. The Genscan-predicted coding sequences may have been edited prior to assembly. (See Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both

35 cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. (See Example

WO 01/79291

PCT/US01/11861

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:15-28, which encodes SECP. The polynucleotide sequences of SEQ ID NO:15-28, 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:15-28 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:15-28. 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 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

WO 01/79291

PCT/US01/11861

polynucleotide sequence of naturally occurring SECP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode SECP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring SECP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding SECP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding SECP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

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

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:15-28 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

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

WO 01/79291

PCT/US01/11861

(1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

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

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

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

WO 01/79291

PCT/US01/11861

controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

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

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter SECP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA 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. Biotechnol. 17:259-264; and Crameri, 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 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 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 Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, SECP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques.

WO 01/79291

PCT/US01/11861

(See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of SECP, or any part thereof, may be altered during direct synthesis and/or  
 5 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 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  
 10 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 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  
 15 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 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  
 20 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 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  
 25 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 vectors containing sequences encoding SECP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in  
 30 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.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding SECP. These include, but are not limited to, microorganisms such as bacteria transformed  
 35 with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with

WO 01/79291

PCT/US01/11861

- yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, *supra*; Ausubel, *supra*; Van Heeke, G. and S.M. Schuster (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. Somia (1997) Nature 389:239-242.)
- The invention is not limited by the host cell employed.
- In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding SECP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding SECP can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding SECP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of SECP are needed, e.g. for the production of antibodies, vectors which direct high level expression of SECP may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.
- Yeast expression systems may be used for production of SECP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)
- Plant systems may also be used for expression of SECP. Transcription of sequences

WO 01/79291

PCT/US01/11861

encoding SECP may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.)

These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding SECP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses SECP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

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

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

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apv* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to

WO 01/79291

PCT/US01/11861

methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci. USA* 77:3567-3570; Colbere-Garapin, F. et al. (1981) *J. Mol. Biol.* 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech),  $\beta$  glucuronidase and its substrate  $\beta$ -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) *Methods Mol. Biol.* 55:121-131.)

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

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

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

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

WO 01/79291

PCT/US01/11861

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 cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity

WO 01/79291

PCT/US01/11861

purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the SECP encoding sequence and the heterologous protein sequence, so that SECP may be cleaved away from the heterologous moiety following purification.

- 5 Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

- In a further embodiment of the invention, synthesis of radiolabeled SECP may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These  
10 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, <sup>35</sup>S-methionine.

- SECP of the present invention or fragments thereof may be used to screen for compounds that specifically bind to SECP. At least one and up to a plurality of test compounds may be screened  
15 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):  
20 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  
25 *E. coli*. Cells expressing SECP or cell membrane fractions which contain SECP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either SECP or the compound is analyzed.

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

WO 01/79291

PCT/US01/11861

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 (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

WO 01/79291

PCT/US01/11861

integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress SECP, e.g., by secreting SECP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) *Biotechnol. Annu. Rev.* 4:55-74).

#### THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of SECP and secreted proteins. In addition, the expression of SECP is closely associated with endocrine, reproductive, muscle, tumorous, aortic smooth muscle, brain, and testicular tissue, and tissue involved with growth and development. Therefore, SECP appears to play a role in cell proliferative, autoimmune/inflammatory, cardiovascular, neurological, and developmental disorders. In the treatment of disorders associated with increased SECP expression or activity, it is desirable to decrease the expression or activity of SECP. In the treatment of disorders associated with decreased SECP expression or activity, it is desirable to increase the expression or activity of SECP.

Therefore, in one embodiment, SECP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of SECP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's

WO 01/79291

PCT/US01/11861

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 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, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation, congenital lung anomalies, atelectasis, pulmonary congestion and edema, 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 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, 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 viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental

WO 01/79291

PCT/US01/11861

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

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

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

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

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

WO 01/79291

PCT/US01/11861

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

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

An antagonist of SECP may be produced using methods which are generally known in the art. In particular, purified SECP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind SECP. Antibodies to SECP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

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, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

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

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

WO 01/79291

PCT/US01/11861

- In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce SECP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotype composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)
- Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)
- Antibody fragments which contain specific binding sites for SECP may also be generated. For example, such fragments include, but are not limited to, F(ab')<sub>2</sub> fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)
- Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between SECP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering SECP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).
- Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for SECP. Affinity is expressed as an association constant, K<sub>a</sub>, 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<sub>a</sub> determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple SECP epitopes, represents the average affinity, or avidity, of the antibodies for SECP. The K<sub>a</sub> 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<sub>a</sub> ranging from about 10<sup>9</sup> to 10<sup>12</sup> L/mole are preferred for use in immunoassays in which the

WO 01/79291

PCT/US01/11861

SECP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with  $K_a$  ranging from about  $10^6$  to  $10^7$  L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of SECP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of SECP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding SECP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding SECP. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding SECP. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

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

In another embodiment of the invention, polynucleotides encoding SECP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-

WO 01/79291

PCT/US01/11861

linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in SECP expression or regulation causes disease, the expression of SECP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in SECP are treated by constructing mammalian expression vectors encoding SECP and introducing these vectors by mechanical means into SECP-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vivo include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J.-L. and H. R  c  pon (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  $\beta$ -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V.

WO 01/79291

PCT/US01/11861

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 TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 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<sup>+</sup> 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, L. (1997) *Blood* 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding SECP to cells which have one or more genetic abnormalities with respect to the expression of SECP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas

WO 01/79291

PCT/US01/11861

(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 Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 389:239-242, both  
 5 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 SECP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with  
 10 ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is  
 15 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.  
 20 (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding SECP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid  
 30 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  
 35 alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a

WO 01/79291

PCT/US01/11861

persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of SECP into a variety of cell types. The specific transduction of a subset of  
5 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 -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.  
15 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  
20 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,  
25 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.

30 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  
35 with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs

WO 01/79291

PCT/US01/11861

that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous

endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding SECP. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased SECP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding SECP may be therapeutically useful, and in the treatment of disorders associated with decreased SECP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding SECP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding SECP is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding SECP are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding SECP. The amount of hybridization may be quantified, thus forming

WO 01/79291

PCT/US01/11861

the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

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

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

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. 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

WO 01/79291

PCT/US01/11861

and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

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

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising SECP or fragments thereof. For example, liposome preparations  
10 containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, SECP or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

15 For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

20 A therapeutically effective dose refers to that amount of active ingredient, for example SECP or fragments thereof, antibodies of SECP, and agonists, antagonists or inhibitors of SECP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population) or LD<sub>50</sub> (the dose  
25 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 LD<sub>50</sub>/ED<sub>50</sub> ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED<sub>50</sub> with little or no toxicity.  
30 The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the  
35 severity of the disease state, the general health of the subject, the age, weight, and gender of the

WO 01/79291

PCT/US01/11861

subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1  $\mu$ g to 100,000  $\mu$ 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.

#### 10 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

WO 01/79291

PCT/US01/11861

from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding SECP, allelic variants, or related sequences.

5 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:15-28 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  
10 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 <sup>32</sup>P or <sup>35</sup>S, or by enzymatic labels,  
15 such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding SECP may be used for the diagnosis of disorders associated with expression of SECP. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal  
20 hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers 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  
25 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  
30 mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome,  
35 systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura,

WO 01/79291

PCT/US01/11861

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 arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and

5 phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, and coronary artery bypass graft surgery, 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

10 endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, and complications of cardiac transplantation, congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction, pulmonary hypertension, vascular sclerosis, obstructive pulmonary disease,

15 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 interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia, diffuse pulmonary

20 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 transplantation; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms,

25 Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous

30 system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic

WO 01/79291

PCT/US01/11861

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 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. The polynucleotide

15 sequences encoding SECP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered SECP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding SECP may be useful in assays that

20 detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding SECP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to

25 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,

30 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

35 is used. Standard values obtained in this manner may be compared with values obtained from

WO 01/79291

PCT/US01/11861

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 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

WO 01/79291

PCT/US01/11861

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  
5 biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives  
10 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  
15 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  
20 effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, SECP, fragments of SECP, or antibodies specific for SECP may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein  
25 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  
30 a given time. (See Seilhamer 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  
35 invention or their complements comprise a subset of a plurality of elements on a microarray. The

WO 01/79291

PCT/US01/11861

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.

5 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 molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and  
10 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 expression information from a large number of genes and gene families. Ideally, a genome-wide  
15 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 gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms,  
20 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 toxicological screening using toxicant signatures to include all expressed gene sequences.

25 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 invention may be quantified. The transcript levels in the treated biological sample are compared with  
30 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 pattern of protein expression in a particular tissue or cell type. Each protein component of a  
35 proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles,

WO 01/79291

PCT/US01/11861

are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are  
5 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  
10 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  
15 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,  
20 and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendoz, 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  
25 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.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be  
30 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  
35 sample containing proteins with the test compound. Proteins that are expressed in the treated

WO 01/79291

PCT/US01/11861

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  
 5 the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of  
 10 protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g.,  
 15 Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed.  
 20 (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding SECP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members  
 25 of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat.  
 30 Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

35 Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic

WO 01/79291

PCT/US01/11861

map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, *supra*, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding SECP on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

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

In another embodiment of the invention, SECP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between SECP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with SECP, or fragments thereof, and washed. Bound SECP is then detected by methods well known in the art. Purified SECP can 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

WO 01/79291

PCT/US01/11861

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 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. 60/197,854, U.S. Ser. No. 60/202,373, U.S. Ser. No. 60/205,899, U.S. Ser. No. 60/210,155, and U.S. Ser. No. 60/209,401, are hereby expressly incorporated by reference.

#### EXAMPLES

##### I. Construction of cDNA Libraries

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 extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPRT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g.,

WO 01/79291

PCT/US01/11861

PBLUESCRIPT plasmid (Stratagene), PSORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 $\alpha$ , DH10B, or ElectroMAX DH10B from Life Technologies.

## II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

## III. Sequencing and Analysis

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

WO 01/79291

PCT/US01/11861

vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and

5 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.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences.

10 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 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

15 the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length

20 polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

25 Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column

30 presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ

35 ID NO:15-28. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization

WO 01/79291

PCT/US01/11861

and amplification technologies are described in Table 4, column 4.

#### IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative secreted proteins were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpr1 and gbhtg). Genscan is a general-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 FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode secreted proteins, the encoded polypeptides were analyzed by querying against PFAM models for secreted proteins. Potential secreted proteins were also identified by homology to Incyte 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

##### 25 "Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory 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

WO 01/79291

PCT/US01/11861

but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpi public databases. Incorrect exons predicted by Gensean 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.

#### 10 "Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

#### VI. Chromosomal Mapping of SECP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:15-28 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:15-28 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between

WO 01/79291

PCT/US01/11861

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 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.

In this manner, SEQ ID NO:25 was mapped to chromosome 7 within the interval from 122.30 to 126.50 centiMorgans. SEQ ID NO:28 was mapped to chromosome 12 within the interval from 137.50 to 160.90 centiMorgans.

#### VII. Analysis of Polynucleotide Expression

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

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

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

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the

WO 01/79291

PCT/US01/11861

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.

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

WO 01/79291

PCT/US01/11861

alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

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

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

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase  
 (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following  
 parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min;  
 25 Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was  
 quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA  
 recoveries were reamplified using the same conditions as described above. Samples were diluted  
 with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing  
 primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM  
 30 BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

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

#### IX. Labeling and Use of Individual Hybridization Probes

35 Hybridization probes derived from SEQ ID NO:15-28 are employed to screen cDNAs,

WO 01/79291

PCT/US01/11861

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  $\mu$ Ci of [ $\gamma$ -<sup>32</sup>P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10<sup>7</sup> 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:

10 Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate.

15 Hybridization patterns are visualized using autoradiography or an alternative imaging means and 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*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the

20 aforementioned technologies should be uniform and solid with a non-porous surface (Schema (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

25 procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schema, 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

30 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

35 fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser

WO 01/79291

PCT/US01/11861

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.

#### 5 Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)<sup>+</sup> RNA is purified using the oligo-(dT) cellulose method. Each poly(A)<sup>+</sup> RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/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  
10 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)<sup>+</sup> RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)<sup>+</sup> 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  
15 incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and  
20 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  
25 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 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  
30 washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US

WO 01/79291

PCT/US01/11861

Patent No. 5,807,522, incorporated herein by reference. 1  $\mu$ l of the array element DNA, at an average concentration of 100 ng/ $\mu$ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene).

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

10 **Hybridization**

- Hybridization reactions contain 9  $\mu$ l of sample mixture consisting of 0.2  $\mu$ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm<sup>2</sup> 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  $\mu$ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

20 **Detection**

- Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

- In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

WO 01/79291

PCT/US01/11861

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples  
 5 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 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  
 10 (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping  
 15 emission spectra) between the fluorophores using each fluorophore's emission spectrum.

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

#### 20 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 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  
 25 4.06 software (National Biosciences) and the coding sequence of SECP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the SECP-encoding transcript.

#### XII. Expression of SECP

30 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* (*lac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory  
 35 element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3).

WO 01/79291

PCT/US01/11861

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 contain the cytomegalovirus promoter. 5-10  $\mu$ 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  $\mu$ 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-

WO 01/79291

PCT/US01/11861

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 peptide synthesizer (Applied Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-SECP activity by, for example, binding the peptide or SECP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

WO 01/79291

PCT/US01/11861

**XV. Purification of Naturally Occurring SECP Using Specific Antibodies**

Naturally occurring or recombinant SECP is substantially purified by immunoaffinity chromatography using antibodies specific for SECP. An immunoaffinity column is constructed by covalently coupling anti-SECP antibody to an activated chromatographic resin, such as

5 CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing SECP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of SECP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt  
10 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 <sup>125</sup>I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules  
15 previously arrayed in the wells of a multi-well plate are incubated with the labeled SECP, washed, and any wells with labeled SECP complex are assayed. Data obtained using different concentrations of SECP are used to calculate values for the number, affinity, and association of SECP with the candidate molecules.

Alternatively, molecules interacting with SECP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989) *Nature* 340:245-246, or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

SECP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S.  
25 Patent No. 6,057,101).

**XVII. Demonstration of SECP Activity**

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  
30 added to quiescent 3T3 cultured cells in the presence of [<sup>3</sup>H]thymidine, a radioactive DNA precursor. SECP for this assay can be obtained by recombinant means or from biochemical preparations. Incorporation of [<sup>3</sup>H]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  
35 indicative of growth modulating activity. One unit of activity per milliliter is defined as the

WO 01/79291

PCT/US01/11861

concentration of SECP producing a 50% response level, where 100% represents maximal incorporation of [<sup>3</sup>H]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.

Coenzyme A transferase activity of SECP, such as succinyl CoA-acetoacetate Co-A transferase activity, can be measured by monitoring the increase in A<sub>310</sub> corresponding to the formation of acetoacetyl CoA. Assays are performed in 67 mM lithium-acetoacetate, 300 μM succinyl CoA, and 15 mM MgCl<sub>2</sub> in 50 mM Tris-HCl, pH 9.1 as described in Howard, J.B. et al. (1986; J. Biol. Chem. 261:60-65) and Cortesby-Theulaz, L.E. et al. (1997; J. Biol. Chem. 272:25659-25667).

Alternatively, lipocalin activity is measured by ligand fluorescence enhancement spectrofluorometry (Lin et al. (1997) Molecular Vision 3:17). Examples of ligands include retinol (Sigma, St. Louis MO) and 16-anthryloxy-palmitic acid (16-AP) (Molecular Probes Inc., Eugene OR). Ligand is dissolved in 100% ethanol and its concentration is estimated using known extinction coefficients (retinol: 46,000 A/M/cm at 325 nm; 16-AP: 8,200 A/M/cm at 361 nm). A 700μl aliquot of 1μM SECP in 10mM Tris (pH 7.5), 2 mM EDTA, and 500mM NaCl is placed in a 1 cm pathlength quartz cuvette and 1μl aliquots of ligand solution are added. Fluorescence is measured after 100 seconds after each addition until readings are stable. Change in fluorescence per unit change in ligand concentration is proportional to SECP activity.

35

WO 01/79291

PCT/US01/11861

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  
5 should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

WO 01/79291

PCT/US01/11861

Table 1

| Incyte<br>Product ID | Polypeptide<br>SEQ ID NO: | Incyte<br>Polypeptide ID | Polynucleotide<br>SEQ ID NO: | Incyte<br>Polynucleotide ID |
|----------------------|---------------------------|--------------------------|------------------------------|-----------------------------|
| 775077               | 1                         | 747377CD1                | 15                           | 747377CB1                   |
| 747137               | 2                         | 747377CD1                | 16                           | 747377CB1                   |
| 248055               | 3                         | 248055CD1                | 17                           | 248055CB1                   |
| 318708               | 4                         | 318708CD1                | 18                           | 318708CB1                   |
| 127456               | 5                         | 127456CD1                | 19                           | 127456CB1                   |
| 134942               | 6                         | 134942CD1                | 20                           | 134942CB1                   |
| 140015               | 7                         | 140015CD1                | 21                           | 140015CB1                   |
| 161034               | 8                         | 161034CD1                | 22                           | 161034CB1                   |
| 187209               | 9                         | 187209CD1                | 23                           | 187209CB1                   |
| 260796               | 10                        | 260796CD1                | 24                           | 260796CB1                   |
| 412044               | 11                        | 412044CD1                | 25                           | 412044CB1                   |
| 638118               | 12                        | 638118CD1                | 26                           | 638118CB1                   |
| 743323               | 13                        | 743323CD1                | 27                           | 743323CB1                   |
| 169150               | 14                        | 169150CD1                | 28                           | 169150CB1                   |

WO 01/79291

PCT/US01/11861

Table 2

| Polypeptide<br>SW ID No: | Incyte<br>Polypeptide ID | GenBank<br>Accession ID | Probability            | GenBank Homolog                                                                                              |
|--------------------------|--------------------------|-------------------------|------------------------|--------------------------------------------------------------------------------------------------------------|
| 1                        | 747337CD1                | g337762<br>g337765      | 5.00E-119<br>5.70E-118 | Proteinase [Homo sapiens]<br>Carboxypeptidase activator<br>Protein [Homo sapiens]                            |
| 2                        | 7474024CD1               | g35897                  | 2.30E-105              | [Homo sapiens] precursor RBP<br>Colantuoni V., et al. (1983) Nucleic<br>Acids Res. 11 (22), 7769-7776 (1983) |
| 3                        | 2480555CD1               | g1890108                | 3.10E-245              | [Homo sapiens] lysyl oxidase-related<br>protein<br>Mariani, A. J., et al., (1992)<br>Matrix 12:242-248       |
| 4                        | 3187086CD1               | g1519052                | 1.80E-198              | [Mus musculus] lysyl oxidase-related<br>protein<br>Kobayashi, T., et al. (1994)                              |
| 6                        | 1349442CD1               | g10437669               | 0                      | [Homo sapiens] unnamed protein<br>product                                                                    |
| 7                        | 1400156CD1               | g12833121               | 1.00E-158              | [Mus musculus] (AK002841) putative<br>protein                                                                |
| 8                        | 1610347CD1               | g12845305               | 0                      | [Mus musculus] (AK010097) putative<br>protein                                                                |
| 9                        | 187209CD1                | g4200218                | 1.00E-65               | [Homo sapiens] hypothetical protein                                                                          |
| 10                       | 2807963CD1               | g12958148               | 1.00E-177              | [Macaca mulatta] (AF24204) taste<br>receptor                                                                 |
| 11                       | 412044CD1                | g12841382               | 1.00E-95               | [Mus musculus] (AK007662) putative<br>protein                                                                |
| 14                       | 1691509CD1               | g6180013                | 3.10E-62               | [Homo sapiens] anaphase-promoting<br>complex subunit 5                                                       |

WO 01/79291

PCT/US01/11861

Table 3

| SEQ NO. | Incyte Polypeptide | Amino Acid Residues | Potential Phosphorylation Sites                                                                                          | Potential Glycosylation Sites  | Signature Sequences, Domains and Motifs                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Analytical Methods and Databases                                                                                                                                                                                                                                              |
|---------|--------------------|---------------------|--------------------------------------------------------------------------------------------------------------------------|--------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1       | 7473577CD1         | 521                 | S284 T83 S138<br>S154 T203 T415<br>T443 S452 T58<br>T149 S414 T415<br>S452                                               | N201 N311                      | signal peptide: M1-P19<br>signal cleavage: M1-A17<br>PD01469-C438-C469<br>GLYCOPROTEIN PROTEIN PRECURSOR<br>PD012321: G181-E256<br>PROSAPOSIN PRECURSOR<br>DM02041 PF07602 [84-254:D86-Q210<br>SAPPOSIN REPEAT<br>LIPOCALIN: R32-M45 D34-M45<br>M1-A18 signal peptide<br>M1-A19 signal cleavage<br>LIPOCALIN: R37-C192<br>lipocalin proteins BL00213:<br>A36-D49, I124-V134<br>LIPOCALIN SIGNATURE PR00179:<br>A36-K48, I124-Y136, V154-I169<br>LIPOCALIN DM00680/P18907/I-172: E19-Y191<br>Speract receptor repeat BL00470C:<br>C134-C144 G311-G355<br>Speract receptor-binding region<br>PD01000 BL00920D V538-I584 I584-D620<br>D620-K650 Y552-L691 Q692-A732<br>Speract receptor repeated domain signature<br>G292-L464<br>Lysyl oxidase signature PR0074A:<br>L533-L561 L561-G588 R589-T617 T617-C644<br>E646-I674 D575-E703 S704-D731<br>SPERACT RECEPTOR SIGNATURE PR00459A:<br>L487-Q443 S326-D337 A341-G351 S372-C386<br>L489-N489<br>LYSYL OXIDASE PROTEINLYSINE PRECURSOR<br>SIGNAL GOXIDASE OXIDOREDUCTASE COPPER<br>GLYCOPROTEIN HOMOLOG PD012364: D510-I729<br>signal cleavage: M1-G25<br>Scavenger receptor cysteine-rich domain<br>SRCR: R51-K145 P183-V282 K310-N407<br>S420-S525<br>LYSYL oxidase: S529-A732<br>OXIDOREDUCTASE: L181L; COPPER:<br>PD02929 [Q5063] I-412:E529-1729 | HMMER<br>SPSCAN<br>BLIMPS_PRODUM<br>BLAST_PRODUM<br>BLAST_DOMO<br>NOTES<br>SPSCAN<br>HMMER_PFBM<br>BLIMPS_BLOCKS<br>BLIMPS_BLOCKS<br>BLIMPS_BLOCKS<br>CCGprofile<br>BLIMPS_PRINTS<br>BLIMPS_PRINTS<br>BLAST_PRODUM<br>SPSCAN, HMMER<br>HMMER_PFBM<br>HMMER_PFBM<br>BLAST_DOMO |
| 2       | 7474024CD1         | 201                 | S15 S26 T74<br>S137 S136 S54<br>T94 T36 T127<br>T146                                                                     |                                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                                                                                                                                                                                                               |
| 3       | 2480555CD1         | 753                 | S97 T104 S221<br>T268 S275 T310<br>S568 S649 T310<br>T58 S115 S120<br>T135 T330 S377<br>T392 T411 S424<br>T493 T527 T617 | N111 N266<br>N310 N481<br>N623 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |                                                                                                                                                                                                                                                                               |

WO 01/79291

PCT/US01/11861

Table 3 (cont.)

| SEQ<br>NO. | Incyte<br>ID | Amino<br>Acid<br>Residues | Potential<br>Phosphorylation<br>Sites                                                                                                                  | Potential<br>Glycosylation<br>Sites              | Signature Sequences,<br>Domains and Motifs                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Analytical<br>Methods and<br>Databases                                            |
|------------|--------------|---------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| 4          | 3187086CD1   | 511                       | S226 S35 S423<br>T120 T140 T377<br>T346 T440 T452<br>T462 T476 T493                                                                                    |                                                  | Signal peptide: M1-G21<br>Coenzyme A transferase: K41-Q267, A292-T493<br>CoA transferases protein:<br>BL01273A: G62-I74, BL01273B: R61-S114<br>BL01273C: E130-G169,<br>BL01273D: D206-P249,<br>BL01273E: D206-P249<br>3-oxoadipate CoA-transferase beta chain:<br>DM02057 286-521: K280-Y509<br>P43116 2-216: E290-L503<br>P23673 6-221: A292-M506<br>3-oxoadipate CoA-transferase alpha chain:<br>DM02058 A4171 39-284: K41-R279<br>Succinyl CoA:3-oxoacid coenzyme transferase:<br>PD004976: K51-K265, R293-T493<br>Signal peptide: M1-Q21<br>signal_cleavage:M1-C16<br>signal_cleavage:M1-A28 | HMW<br>BLIMPS-BLOCKS<br>BLAST-DOMO<br>BLAST-DOMO<br>BLAST-PRODOR<br>HMW<br>SPSCAN |
| 5          | 1274566CD1   | 99                        | S82                                                                                                                                                    | N88                                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | HMW<br>SPSCAN                                                                     |
| 6          | 1349442CD1   | 389                       | T45 T86 T195<br>S21 T61 T152<br>S110 S116 S121<br>T164 S222 T361<br>S265 T291 S311<br>S345 S364                                                        | N44 N259 N286<br>N113                            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | HMW<br>SPSCAN                                                                     |
| 7          | 1400156CD1   | 322                       | T51 T60 T76<br>S114 T36 S108<br>T164 S282 S316                                                                                                         | N106 N119<br>N162 N175<br>N192 N205<br>N231 N280 | Signal peptide: M1-G26<br>signal_cleavage:M1-G26                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | HMW<br>SPSCAN                                                                     |
| 8          | 1610347CD1   | 587                       | S81 S146 S310<br>T234 S281 S306<br>S348 T436 S519<br>T579 S290 S26<br>S56 S146 T167<br>S306 S332 S461<br>S516 T569 T579<br>T152 T164 S102<br>T152 T170 | N114 N322<br>N500                                | Signal peptide: M1-A32<br>PROTEIN BIOSYNTHESIS ISONITRILASE SYNTHETASE<br>SYNTHASE OXIDASE III COPROPORPHYRINOGEN<br>BIOTIN TRANSFERASE PD000690:R100-R581                                                                                                                                                                                                                                                                                                                                                                                                                                       | HMW<br>BLAST-PRODOR                                                               |
| 9          | 187209CD1    | 173                       | S516 T569 T579<br>T152 T164 S102<br>T152 T170                                                                                                          |                                                  | signal_cleavage:M1-S53                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | SPSCAN                                                                            |

WO 01/79291

PCT/US01/11861

Table 3 (cont.)

| SEQ ID NO. | Protein ID | Amino Residues | Potential Phosphorylation Sites                                                                                             | Potential Glycosylation Sites | Signature Sequences, Domains and Motifs                                                                                                                                                          | Analytical Method Databases                                                      |
|------------|------------|----------------|-----------------------------------------------------------------------------------------------------------------------------|-------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| 10         | 2607963CD1 | 325            | S175 S192 S319 S156                                                                                                         |                               | signal_peptide:M1-S28<br>signal_cleavage:M1-A26                                                                                                                                                  | HMER<br>SPSCAN                                                                   |
| 11         | 412044CD1  | 733            | S249 S252 S345 T487 T517 T607 N508 N636 S253 S475 S531 S85 S107 T108 T126 S185 S190 S194 S258 S345 S483 T575 S623 S680 T713 | N183 N317                     | signal_cleavage:M1-G22<br>PHD-finger PHD-V118-R166<br>SET domain SET-T323-C453<br>V395-E416<br>KH1 SET domain Protein PD140577:D111-E171<br>KH1 SET domain Protein signature<br>PR00049D:Y66-F80 | SPSCAN<br>HMER PFAM<br>HMER PFAM<br>BLINPS PFAM<br>BLAST PRODOM<br>BLINPS PRINTS |
| 12         | 638118CD1  | 242            | T18 T139 T131 S176 S181                                                                                                     |                               | signal_cleavage:M1-P42                                                                                                                                                                           | SPSCAN                                                                           |
| 13         | 743323CD1  | 153            | T04 T123 S109 S21 T108                                                                                                      |                               | signal_cleavage:M1-G45                                                                                                                                                                           | SPSCAN                                                                           |
| 14         | 1691509CD1 | 134            | Y34                                                                                                                         |                               | signal_peptide:M1-S21                                                                                                                                                                            | HMER                                                                             |

WO 01/79291

PCT/US01/11861

Table 4

| Polynucleotide<br>SEQ ID NO. | Incyte<br>Polynucleotide ID | Sequence<br>Length | Selected<br>Regions                                         | Sequence Fragments      | 5' Position | 3' Position |
|------------------------------|-----------------------------|--------------------|-------------------------------------------------------------|-------------------------|-------------|-------------|
| 15                           | 7473577CB1                  | 1566               | 1-373, 334-<br>1170, 1488-<br>1566                          | GNH.96437516_000004_002 | 1           | 1566        |
| 16                           | 7474024CB1                  | 939                | 1-32                                                        | 2087293H1 (PANCNOT04)   | 1           | 138         |
| 16                           | 7474024CB1                  | 939                | 1-32                                                        | 6535251H1 (OVARDIN02)   | 278         | 939         |
| 16                           | 7474024CB1                  | 939                | 1-32                                                        | 6201685H1 (PTUNON01)    | 127         | 681         |
| 16                           | 7474024CB1                  | 939                | 1-32                                                        | 166267H1 (LIVNOT01)     | 74          | 429         |
| 17                           | 2480555CB1                  | 2785               | 2761-2785,<br>1-570,<br>1685-2093,<br>1317-1480,<br>762-977 | 1781333R6 (PGANNON02)   | 1910        | 2356        |
| 17                           | 2480555CB1                  | 2785               | 2761-2785,<br>1-570,<br>1685-2093,<br>1317-1480,<br>762-977 | stretcher.fasta         | 68          | 2329        |
| 17                           | 2480555CB1                  | 2785               | 2761-2785,<br>1-570,<br>1685-2093,<br>1317-1480,<br>762-977 | 7262753H1 (UTREMC01)    | 1           | 574         |
| 17                           | 2480555CB1                  | 2785               | 2761-2785,<br>1-570,<br>1685-2093,<br>1317-1480,<br>762-977 | 4204281T6 (BBAITUT29)   | 2028        | 2785        |
| 18                           | 3187086CB1                  | 1733               | 1-391,<br>1706-1733,<br>816-1071                            | 4291392T6 (BRABDI01)    | 1182        | 1718        |
| 18                           | 3187086CB1                  | 1733               | 1-391,<br>1706-1733,<br>816-1071                            | 4405672F6 (PROSDIT01)   | 495         | 1010        |
| 18                           | 3187086CB1                  | 1733               | 1-391,<br>1706-1733,<br>816-1071                            | 3187086R6 (THYMNON04)   | 47          | 585         |
| 18                           | 3187086CB1                  | 1733               | 1-391,<br>1706-1733,<br>816-1071                            | 2005793H1 (TESTNOT03)   | 1615        | 1733        |

WO 01/79291

PCT/US01/11861

Table 4

| Polynucleotide<br>SEQ ID NO: | Incyte<br>Polynucleotide ID | Sequence<br>Length | Selected<br>Fragments         | Sequence Fragments     | 5' Position | 3' Position |
|------------------------------|-----------------------------|--------------------|-------------------------------|------------------------|-------------|-------------|
| 18                           | 3187086CB1                  | 1733               | 1-391, 733,<br>816-1071       | 510050BH1 (PROSTVUS20) | 1005        | 1256        |
| 18                           | 3187086CB1                  | 1733               | 1-391, 1706-1733,<br>816-1071 | 3747662H1 (UTRNOT18)   | 1           | 297         |
| 18                           | 3187086CB1                  | 1733               | 1-391, 1706-1733,<br>816-1071 | 6065570H1 (BRAENOT02)  | 920         | 1204        |
| 19                           | 1274566CB1                  | 1148               | 621-666, 1-<br>154            | 1274566T6 (TESTTUT02)  | 525         | 1132        |
| 19                           | 1274566CB1                  | 1148               | 621-666, 1-<br>154            | SCXA01270V1            | 345         | 1035        |
| 19                           | 1274566CB1                  | 1148               | 621-666, 1-<br>154            | 1274566F6 (TESTTUT02)  | 1           | 419         |
| 19                           | 1274566CB1                  | 1148               | 621-666, 1-<br>154            | SCXA01551V1            | 526         | 1148        |
| 20                           | 1349442CB1                  | 1213               | 1019-1213                     | SBA02792F1             | 274         | 855         |
| 20                           | 1349442CB1                  | 1213               | 1019-1213                     | SBA01641F1             | 53          | 713         |
| 20                           | 1349442CB1                  | 1213               | 1019-1213                     | 1349442H1 (LATRTUT02)  | 1           | 253         |
| 20                           | 1349442CB1                  | 1213               | 1019-1213                     | 136161R1 (SYNORAB01)   | 662         | 1213        |
| 21                           | 1400156CB1                  | 2298               | 1-21, 999-<br>1018            | 70222572V1             | 1           | 488         |
| 21                           | 1400156CB1                  | 2298               | 1-21, 999-<br>1018            | 9675553                | 1693        | 2296        |
| 21                           | 1400156CB1                  | 2298               | 1-21, 999-<br>1018            | 2496967T6 (ADRETUT05)  | 936         | 1515        |
| 21                           | 1400156CB1                  | 2298               | 1-21, 999-<br>1018            | 91422205               | 1713        | 2298        |
| 21                           | 1400156CB1                  | 2298               | 1-21, 999-<br>1018            | SXAE05185V1            | 447         | 949         |
| 21                           | 1400156CB1                  | 2298               | 1-21, 999-<br>1018            | 1400156F6 (BRAITUT08)  | 1069        | 1674        |
| 21                           | 1400156CB1                  | 2298               | 1-21, 999-<br>1018            | 231268F1 (SINTNOT02)   | 1505        | 2295        |
| 21                           | 1400156CB1                  | 2298               | 1-21, 999-<br>1018            | SXAE01140V1            | 503         | 1020        |
| 22                           | 1610347CB1                  | 2079               | 1098-1310                     | 1438956T1 (PANCNOT08)  | 1464        | 2079        |
| 22                           | 1610347CB1                  | 2079               | 1098-1310                     | 3204803F6 (PENCNOT03)  | 1           | 575         |
| 22                           | 1610347CB1                  | 2079               | 1098-1310                     | 1981287R6 (LUNGNOT03)  | 1367        | 1902        |
| 22                           | 1610347CB1                  | 2079               | 1098-1310                     | 744211R1 (BRAITUT01)   | 1076        | 1655        |

WO 01/79291

PCT/US01/11861

Table 4

| Polynucleotide<br>SEQ ID NO: | IncYTE<br>Polynucleotide ID | Sequence<br>Length | Selected<br>Fragments              | Sequence Fragments    | 5' Position | 3' Position |
|------------------------------|-----------------------------|--------------------|------------------------------------|-----------------------|-------------|-------------|
| 22                           | 1610347CB1                  | 2079               | 1098-1330                          | 921450H1 (BARRNOT02)  | 852         | 1142        |
| 22                           | 1610347CB1                  | 2079               | 1098-1330                          | SCDA08508V1           | 470         | 1036        |
| 22                           | 187209CB1                   | 846                | 1-333                              | 187209CT1 (CARDNOT01) | 1           | 846         |
| 24                           | 260795CB1                   | 1148               | 1-1148                             | 82EC00624V1           | 593         | 693         |
| 24                           | 260795CB1                   | 1148               | 1-1148                             | 43L07826 (THHEDT01)   | 586         | 1148        |
| 25                           | 412044CB1                   | 3076               | 1-3076,<br>1900-3076,<br>1111-1285 | 4191676 (BRSTNOT01)   | 2482        | 3076        |
| 25                           | 412044CB1                   | 3076               | 1-157,<br>1900-3076,<br>1111-1285  | 869590R1 (LUNGAST01)  | 1281        | 1821        |
| 25                           | 412044CB1                   | 3076               | 1-157,<br>1900-3076,<br>1111-1285  | 412044R1 (BRSTNOT01)  | 2168        | 2779        |
| 25                           | 412044CB1                   | 3076               | 1-157,<br>1900-3076,<br>1111-1285  | 45232R7 (TLYNNOT02)   | 1908        | 2445        |
| 25                           | 412044CB1                   | 3076               | 1-157,<br>1900-3076,<br>1111-1285  | 2503204F6 (CONUTUT01) | 495         | 1097        |
| 25                           | 412044CB1                   | 3076               | 1-157,<br>1900-3076,<br>1111-1285  | 1678531F6 (STOMFET01) | 1095        | 1625        |
| 25                           | 412044CB1                   | 3076               | 1-157,<br>1900-3076,<br>1111-1285  | 3440195H1 (PENCNOT06) | 450         | 689         |
| 25                           | 412044CB1                   | 3076               | 1-157,<br>1900-3076,<br>1111-1285  | 2948371T6 (KIDNFE01)  | 1689        | 2319        |
| 25                           | 412044CB1                   | 3076               | 1-157,<br>1900-3076,<br>1111-1285  | 269165F1 (HMT2NOT01)  | 11          | 605         |
| 25                           | 412044CB1                   | 3076               | 1-157,<br>1900-3076,<br>1111-1285  | 2503204F6 (CONUTUT01) | 1           | 416         |
| 25                           | 412044CB1                   | 3076               | 1-157,<br>1900-3076,<br>1111-1285  | 412044F1 (BRSTNOT01)  | 2436        | 3075        |
| 25                           | 412044CB1                   | 3076               | 1-157,<br>1900-3076,<br>1111-1285  | 2836520F6 (TLYNNOT03) | 679         | 1212        |

WO 01/79291

PCT/US01/11861

Table 4

| Polynucleotide<br>SEQ ID NO. | Incyte<br>Polynucleotide ID | Sequence<br>Length | Selected<br>Fragments | Sequence Fragments    | 5' Position | 3' Position |
|------------------------------|-----------------------------|--------------------|-----------------------|-----------------------|-------------|-------------|
| 26                           | 638118CB1                   | 2102               | 1-288,<br>1342-1361   | 322585R6 (EOSIHF02)   | 1714        | 2102        |
| 26                           | 638118CB1                   | 2102               | 1-288,<br>1342-1361   | 2476940H1 (SMCANOT01) | 1           | 224         |
| 26                           | 638118CB1                   | 2102               | 1-288,<br>1342-1361   | 1440105R1 (THYENOT03) | 1549        | 2102        |
| 26                           | 638118CB1                   | 2102               | 1-288,<br>1342-1361   | 1856507F6 (PROSNOTI8) | 983         | 1667        |
| 26                           | 638118CB1                   | 2102               | 1-288,<br>1342-1361   | 1534131H1 (SPLINOT04) | 340         | 634         |
| 26                           | 638118CB1                   | 2102               | 1-288,<br>1342-1361   | 638118R1 (BRSTNOT03)  | 607         | 1335        |
| 26                           | 638118CB1                   | 2102               | 1-288,<br>1342-1361   | 1891046H1 (BLADTUT07) | 196         | 550         |
| 27                           | 743232CB1                   | 807                | 1-464                 | 3835037F6 (PANCNOT17) | 280         | 807         |
| 27                           | 743232CB1                   | 807                | 1-464                 | 950902H1 (PANCNOT05)  | 234         | 234         |
| 27                           | 743232CB1                   | 807                | 1-464                 | 407778F6 (PANCNOT19)  | 56          | 481         |
| 28                           | 1691509CB1                  | 1049               | 1-519,<br>1021-1049   | 71145648V1            | 497         | 1049        |
| 28                           | 1691509CB1                  | 1049               | 1-519,<br>1021-1049   | 1999452R6 (BRSTTUT03) | 1           | 527         |

WO 01/79291

PCT/US01/11861

Table 5

| Polynucleotide<br>Seq. ID No. | Incyte<br>Seq. ID | Representative Library |
|-------------------------------|-------------------|------------------------|
| 16                            | 74700CB1          | LIVERUT04              |
| 17                            | 248055CB1         | SMCANT01               |
| 18                            | 3187086CB1        | TESTUT03               |
| 19                            | 1274566CB1        | TESTUT02               |
| 20                            | 1349442CB1        | SYNOFA01               |
| 21                            | 1400156CB1        | PANCDIT01              |
| 22                            | 1610347CB1        | COLINTUT06             |
| 23                            | 187209CB1         | CARDNOT01              |
| 24                            | 2607963CB1        | LUNGUT07               |
| 25                            | 412044CB1         | BRSTNOT01              |
| 26                            | 638118CB1         | THYRNOT03              |
| 27                            | 743223CB1         | PANCNOT05              |
| 28                            | 1691593CB1        | PROSVS23               |

WO 01/79291

PCT/US01/11861

Table 6

| Library    | Vector      | Library Description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|------------|-------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| BHSTN0T01  | PBLUESCRIPT | Library was constructed using RNA isolated from the breast tissue of a 56-year-old Caucasian female who died in a motor vehicle accident.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| CARDNGT01  | PBLUESCRIPT | Library was constructed using RNA isolated from the cardiac muscle of a 65-year-old Caucasian male who died from a gunshot wound.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| COLANTUT06 | P1NCT       | Library was constructed using RNA isolated from colon tumor tissue obtained from a 45-year-old Caucasian female during a total colectomy and total abdominal hysterectomy. Pathology indicated invasive grade 2 colonic adenocarcinoma forming a cecal mass. Patient history included benign neoplasms of the rectum and anus, multiple sclerosis and mitral valve disorder. Previous surgeries included a polypectomy. Family history included type 1 diabetes, cerebrovascular disease, atherosclerotic coronary artery disease, malignant skin neoplasms, hypertension, atherosclerotic coronary artery disease and gallbladder disease.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| LIVARTUT04 | P1NCT       | Library was constructed using RNA isolated from liver tissue removed from a 50-year-old Caucasian male during a partial hepatectomy. Pathology indicated a grade 3-4 hepatoma, forming a mass. Patient history included benign hypertension and hepatitis. Hepatitis B core antigen and hepatitis B surface antigen was present in the patient.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| LUNGTFUT07 | P1NCT       | Library was constructed using RNA isolated from lung tumor tissue removed from the upper lobe of a 50-year-old Caucasian male during segmental lung resection. Pathology indicated an invasive grade 4 squamous cell adenocarcinoma. Patient history included tobacco use. Family history included skin cancer.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| PANCDIT01  | PBLUESCRIPT | Library was constructed using RNA isolated from pancreatic tissue removed from a 15-year-old Caucasian male who died from a gunshot wound. Patient history included type 1 diabetes which was being treated with insulin.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| PANNGT05   | PSP0RT1     | Library was constructed using RNA isolated from the pancreatic tissue of a 2-year-old Hispanic male who died from cerebral anoxia.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| PROSTUS23  | P1NCT       | This subtracted prostate tumor library was constructed using 10 million clones from a pooled prostate tumor library that was subjected to 2 rounds of subtractive hybridization with 10 million clones from a pooled prostate tissue library. The starting library for subtraction was constructed by pooling equal numbers of clones from 4 prostate tumor libraries isolated from Caucasian males at ages 58 (A), 61 (B), 66 (C) and 68 (D) during prostatectomy with lymph node excision. Pathology indicated adenocarcinoma in all donors. History included elevated PSA, induration and tobacco abuse in donor A; elevated PSA, induration, prostate hyperplasia, renal failure, osteoarthritis, renal artery stenosis, benign HTN, thrombocytopenia, hyperlipidemia, tobacco/alcohol abuse and hepatitis C (carrier) in donor B; elevated PSA, induration, and tobacco abuse in donor C; and elevated PSA, induration, hypercholesterolemia, and kidney calculus in donor D. The hybridization probe for subtraction was constructed by pooling equal numbers of cDNA clones from 3 prostate tumor libraries isolated from Caucasian males. The subtractive hybridization fibroblasts from prostate stroma from 3 different donors. Subtractive hybridization conditions were based on the methodologies of Swarcop et al., NAR 19 (1991):1954 and |

WO 01/79291

PCT/US01/11861

Table 6 (Cont.)

|            |             |                                                                                                                                                                                                                                                                                                                                   |
|------------|-------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| SMCNOT01   | PINCY       | Bonaldo, et al. Genome Research 6 (1996):791.<br>Library was constructed using RNA isolated from an aortic smooth muscle cell line derived from the explanted heart of a male during a heart transplant.                                                                                                                          |
| SYNORAB01  | PBLUESCRIPT | Library was constructed using RNA isolated from the synovial membrane tissue of a 68-year-old Caucasian female with rheumatoid arthritis.                                                                                                                                                                                         |
| TESTINGT03 | PBLUESCRIPT | Library was constructed using RNA isolated from the synovial membrane tissue of a 37-year-old Caucasian male, who died from liver disease. Patient history included cirrhosis, jaundice, and liver failure.                                                                                                                       |
| TESTTUT02  | PINCY       | Library was constructed using RNA isolated from testicular tumor removed from a 31-year-old Caucasian male during unilateral orchiectomy. Pathology indicated embryonal carcinoma.                                                                                                                                                |
| THYRNOT03  | PINCY       | Library was constructed using RNA isolated from thyroid tissue removed from the left thyroid of a 28-year-old Caucasian female during a complete thyroidectomy. Pathology indicated a small nodule of adenomatous hyperplasia present in the left thyroid. Pathology also indicated a well-encapsulated mass in the left thyroid. |

WO 01/79291

PCT/US01/11861

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; Paracel Inc., Pasadena, CA.                                                                                                                                                | Mismatch <50%                                                                                                                                                                                                  |
| ABI AutoAssembler | A program that assembles nucleic acid sequences.                                                                                                                                                                 | Applied Biosystems, Foster City, CA.                                                                                                                                                                            |                                                                                                                                                                                                                |
| BLAST             | A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.                 | Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.                                                                                           | ESTs: Probability value= 1.0E-8 or less<br>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, fastx, fasty, and search. | Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.             | ESTs: fasta E value=1.0E-6<br>Assembled ESTs: fasta Identity= 95% or greater and<br>Match length=200 bases or greater;<br>fastx E value=1.0E-8 or less<br>Full Length sequences:<br>fastx score=100 or greater |
| BLIMPS            | A BLOCKS IMPROVED Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.         | Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Atwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424. | Probability value= 1.0E-3 or less                                                                                                                                                                              |
| 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.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nussli, Cambridge Univ. Press, pp. 1-350.    | PFAM hits: Probability value= 1.0E-3 or less<br>Signal peptide hits: Score= 0 or greater                                                                                                                       |

WO 01/79291

PCT/US01/11861

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.                                                                             | Gribskov, M. et al. (1988) CABIOS 4:61-66;<br>Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Batroch, A. et al. (1997) Nucleic Acids Res. 25:217-221.                                    | Normalized quality scores: GCG-specified "HQQH" value for that particular Prosite motif.<br>Generally, score=1.4-2.1. |
| Phred       | A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.                                                                                                        | Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.                                                                                                |                                                                                                                       |
| Phrap       | A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences. | Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.             | Score= 120 or greater;<br>Match length= 56 or greater                                                                 |
| Consed      | A graphical tool for viewing and editing Phrap assemblies.                                                                                                                                                      | Gordon, D. et al. (1998) Genome Res. 8:195-202.                                                                                                                                                    |                                                                                                                       |
| SPScan      | A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.                                                                                                    | Nielsen, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.                                                                                        | Score=3.5 or greater                                                                                                  |
| TMAP        | A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.                                                                                         | Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.                                                                                 |                                                                                                                       |
| TMHMMER     | A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.                                                                             | Sonnhammer, E.L. et al. (1998) Proc. Sixth 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. |                                                                                                                       |
| Motifs      | A program that searches amino acid sequences for patterns that matched those defined in Prosite.                                                                                                                | Batroch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.                                           |                                                                                                                       |

WO 01/79291

PCT/US01/11861

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:1-14,
  - 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:1-14,
  - c) a biologically active fragment of a polypeptide having an amino acid sequence selected  
from the group consisting of SEQ ID NO:1-14, and
  - 10 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from  
the group consisting of SEQ ID NO:1-14.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-  
14.
- 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:15-28.
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 01/79291

PCT/US01/11861

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:15-28,

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:15-28,

10 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).

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:

20 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.

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 01/79291

PCT/US01/11861

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:1-14.

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.

10

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.

15

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.

30

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

WO 01/79291

PCT/US01/11861

1, said method comprising the steps of:

- 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.

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,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

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:

- 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.

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 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

WO 01/79291

PCT/US01/11861

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 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

10 b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

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

a) a chimeric antibody,

b) a single chain antibody,

15 c) a Fab fragment,

d) a F(ab')<sub>2</sub> fragment, or

e) a humanized antibody.

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.

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

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:1-14, or an immunogenic fragment thereof, under conditions to

WO 01/79291

PCT/US01/11861

elicit an antibody response;

b) isolating antibodies from said animal; and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14.

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:1-14, 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:1-14.

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

25

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

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:1-14 in a sample, comprising the steps of:

WO 01/79291

PCT/US01/11861

- a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
- b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14 in the sample.

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-14 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
- 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:1-14.

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

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.

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.

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.

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 01/79291

PCT/US01/11861

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.
58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
59. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:15.
60. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:16.
61. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:17.
62. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:18.
63. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:19.
64. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:20.
65. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:21.
66. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:22.
67. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:23.

WO 01/79291

PCT/US01/11861

68. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID  
NO:24.

69. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID  
5 NO:25.

70. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID  
NO:26.

71. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID  
10 NO:27.

72. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID  
NO:28.  
15

20

25

30

WO 01/79291

PCT/US01/11861

<110> INCYTE GENOMICS, INC.  
 GRIFFIN, Jennifer A.  
 YAO, Monique G.  
 BRUNS, Christopher M.  
 YUE, Henry  
 DELEGEANE, Angelo M.  
 HAFALIA, April  
 PATTERSON, Chandra  
 POLICKY, Jennifer L.  
 TRIBOULEY, Catherine M.  
 BAUGHN, Mariah R.  
 NGUYEN, Darniel B.  
 LAL, Preeti  
 TANG, Y. Tom  
 HILLMAN, Jennifer L.  
 LU, Dyrng Aina M.  
 BATRA, Sajeew  
 AU-YOUNG, Janice  
 REDDY, Roopa  
 AZIMZAI, Yalda

<120> SECRETED PROTEINS

<130> PI-0071 PCT

<140> To Be Assigned  
 <141> Herewith

<150> 60/197,854; 60/202,373; 60/205,899; 60/210,155; 60/209,401  
 <151> 2000-04-14; 2000-05-04; 2000-05-18; 2000-06-01; 2000-06-01

<160> 28  
 <170> PERL Program

<210> 1  
 <211> 521  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7473577CD1

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

WO 01/79291

PCT/US01/11861

Ala Leu Ser Leu Cys Glu Pro Leu Gln Arg His Leu Ala Thr Leu  
 140 145 150  
 Arg Pro Leu Ser Lys Glu Asp Thr Phe Glu Ala Val Ala Pro Phe  
 155 160 165  
 Met Ala Asn Gly Pro Leu Thr Phe His Pro Arg Gln Ala Pro Glu  
 170 175 180  
 Gly Ala Leu Cys Gln Asp Cys Val Arg Gln Val Ser Arg Leu Gln  
 185 190 195  
 Glu Ala Val Arg Ser Asn Leu Thr Leu Ala Asp Leu Asn Ile Gln  
 200 205 210  
 Glu Gln Cys Glu Ser Leu Gly Pro Gly Leu Ala Val Leu Cys Lys  
 215 220 225  
 Asn Tyr Leu Phe Gln Phe Phe Val Pro Ala Asp Gln Ala Leu Arg  
 230 235 240  
 Leu Leu Pro Pro Gln Glu Leu Cys Arg Lys Gly Gly Phe Cys Glu  
 245 250 255  
 Glu Leu Gly Ala Pro Ala Arg Leu Thr Gln Val Val Ala Met Asp  
 260 265 270  
 Gly Val Pro Ser Leu Glu Leu Gly Leu Pro Arg Lys Gln Ser Glu  
 275 280 285  
 Met Gln Met Lys Ala Gly Val Thr Cys Glu Val Cys Met Asn Val  
 290 295 300  
 Val Gln Lys Leu Asp His Trp Leu Met Ser Asn Ser Ser Glu Leu  
 305 310 315  
 Met Ile Thr His Ala Leu Glu Arg Val Cys Ser Val Met Pro Ala  
 320 325 330  
 Ser Ile Thr Lys Glu Cys Ile Ile Leu Val Asp Thr Tyr Ser Pro  
 335 340 345  
 Ser Leu Val Gln Leu Val Ala Lys Ile Thr Pro Glu Lys Val Cys  
 350 355 360  
 Lys Phe Ile Arg Leu Cys Gly Asn Arg Arg Arg Ala Arg Ala Val  
 365 370 375  
 His Asp Ala Tyr Ala Ile Val Pro Ser Pro Glu Trp Asp Ala Glu  
 380 385 390  
 Asn Gln Gly Ser Phe Cys Asn Gly Cys Lys Arg Leu Leu Thr Val  
 395 400 405  
 Ser Ser His Asn Leu Glu Ser Lys Ser Thr Lys Arg Asp Ile Leu  
 410 415 420  
 Val Ala Phe Lys Gly Gly Cys Ser Ile Leu Pro Leu Pro Tyr Met  
 425 430 435  
 Ile Gln Cys Lys His Phe Val Thr Gln Tyr Glu Pro Val Leu Ile  
 440 445 450  
 Glu Ser Leu Lys Asp Met Met Asp Pro Val Ala Val Cys Lys Lys  
 455 460 465  
 Val Gly Ala Cys His Gly Pro Arg Thr Pro Leu Leu Gly Thr Asp  
 470 475 480  
 Gln Cys Ala Leu Gly Pro Ser Phe Trp Cys Arg Ser Gln Glu Ala  
 485 490 495  
 Ala Lys Leu Cys Asn Ala Val Gln His Cys Gln Lys His Val Trp  
 500 505 510  
 Lys Glu Met His Leu His Ala Gly Glu His Ala  
 515 520

<210> 2  
 <211> 201  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7474024CD1

<400> 2

WO 01/79291

PCT/US01/11861

```

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

```

&lt;210&gt; 3

&lt;211&gt; 753

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 2480555CD1

&lt;400&gt; 3

```

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

```

WO 01/79291

PCT/US01/11861

Pro Leu Pro Val Thr Glu Gly Leu Val Glu Val Arg Leu Pro Asp  
 185 190 195  
 Gly Trp Ser Gln Val Cys Asp Lys Gly Trp Ser Ala His Asn Ser  
 200 205 210  
 His Val Val Cys Gly Met Leu Gly Phe Pro Ser Glu Lys Arg Val  
 215 220 225  
 Asn Ala Ala Phe Tyr Arg Leu Leu Ala Gln Arg Gln Gln His Ser  
 230 235 240  
 Phe Gly Leu His Gly Val Ala Cys Val Gly Thr Glu Ala His Leu  
 245 250 255  
 Ser Leu Cys Ser Leu Glu Phe Tyr Arg Ala Asn Asp Thr Ala Arg  
 260 265 270  
 Cys Pro Gly Gly Gly Pro Ala Val Val Ser Cys Val Pro Gly Pro  
 275 280 285  
 Val Tyr Ala Ala Ser Ser Gly Gln Lys Lys Gln Gln Gln Ser Lys  
 290 295 300  
 Pro Gln Gly Glu Ala Arg Val Arg Leu Lys Gly Gly Ala His Pro  
 305 310 315  
 Gly Glu Gly Arg Val Glu Val Leu Lys Ala Ser Thr Trp Gly Thr  
 320 325 330  
 Val Cys Asp Arg Lys Trp Asp Leu His Ala Ala Ser Val Val Cys  
 335 340 345  
 Arg Glu Leu Gly Phe Gly Ser Ala Arg Glu Ala Leu Ser Gly Ala  
 350 355 360  
 Arg Met Gly Gln Gly Met Gly Ala Ile His Leu Ser Glu Val Arg  
 365 370 375  
 Cys Ser Gly Gln Glu Leu Ser Leu Trp Lys Cys Pro His Lys Asn  
 380 385 390  
 Ile Thr Ala Glu Asp Cys Ser His Ser Gln Asp Ala Gly Val Arg  
 395 400 405  
 Cys Asn Leu Pro Tyr Thr Gly Ala Glu Thr Arg Ile Arg Leu Ser  
 410 415 420  
 Gly Gly Arg Ser Gln His Glu Gly Arg Val Glu Val Gln Ile Gly  
 425 430 435  
 Gly Pro Gly Pro Leu Arg Trp Gly Leu Ile Cys Gly Asp Asp Trp  
 440 445 450  
 Gly Thr Leu Glu Ala Met Val Ala Cys Arg Gln Leu Gly Leu Gly  
 455 460 465  
 Tyr Ala Asn His Gly Leu Gln Glu Thr Trp Tyr Trp Asp Ser Gly  
 470 475 480  
 Asn Ile Thr Glu Val Val Met Ser Gly Val Arg Cys Thr Gly Thr  
 485 490 495  
 Glu Leu Ser Leu Asp Gln Cys Ala His His Gly Thr His Ile Thr  
 500 505 510  
 Cys Lys Arg Thr Gly Thr Arg Phe Thr Ala Gly Val Ile Cys Ser  
 515 520 525  
 Glu Thr Ala Ser Asp Leu Leu Leu His Ser Ala Leu Val Gln Glu  
 530 535 540  
 Thr Ala Tyr Ile Glu Asp Arg Pro Leu His Met Leu Tyr Cys Ala  
 545 550 555  
 Ala Glu Glu Asn Cys Leu Ala Ser Ser Ala Arg Ser Ala Asn Trp  
 560 565 570  
 Pro Tyr Gly His Arg Arg Leu Leu Arg Phe Ser Ser Gln Ile His  
 575 580 585  
 Asn Leu Gly Arg Ala Asp Phe Arg Pro Lys Ala Gly Arg His Ser  
 590 595 600  
 Trp Val Trp His Glu Cys His Gly His Tyr His Ser Met Asp Ile  
 605 610 615  
 Phe Thr His Tyr Asp Ile Leu Thr Pro Asn Gly Thr Lys Val Ala  
 620 625 630  
 Glu Gly His Lys Ala Ser Phe Cys Leu Glu Asp Thr Glu Cys Gln  
 635 640 645  
 Glu Asp Val Ser Lys Arg Tyr Glu Cys Ala Asn Phe Gly Glu Gln

WO 01/79291

PCT/US01/11861

|                 |     |                                             |     |                     |     |
|-----------------|-----|---------------------------------------------|-----|---------------------|-----|
| Gly Ile Thr Val | 650 | Gly Cys Trp Asp Leu Tyr                     | 655 | Arg His Asp Ile Asp | 660 |
| Cys Gln Trp Ile | 665 | Ile Thr Asp Val Lys Pro Gly Asn Tyr Ile     | 670 |                     | 675 |
| Leu Gln Val Val | 680 | Ile Asn Pro Asn Phe Glu Val Ala Glu Ser Asp | 685 |                     | 690 |
| Phe Thr Asn Asn | 695 | Ala Met Lys Cys Asn Cys Lys Tyr Asp Gly His | 700 |                     | 705 |
| Arg Ile Trp Val | 710 | His Asn Cys His Ile Gly Asp Ala Phe Ser Glu | 715 |                     | 720 |
| Glu Ala Asn Arg | 725 | Arg Phe Glu Arg Tyr Pro Gly Gln Thr Ser Asn | 730 |                     | 735 |
| Gln Ile Ile     | 740 |                                             | 745 |                     | 750 |

&lt;210&gt; 4

&lt;211&gt; 511

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 3187086CD1

&lt;400&gt; 4

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

WO 01/79291

PCT/US01/11861

```

275      280      285
Ala Gly Lys Glu Asp Ala Arg Thr Arg Ile Ile Arg His Ala
290      295      300
Ala Leu Glu Phe Glu Asp Gly Met Tyr Ala Asn Leu Gly Ile Gly
305      310      315
Ile Pro Leu Leu Ala Ser Asn Phe Ile Ser Pro Ser Met Thr Val
320      325      330
His Leu His Ser Glu Asn Gly Ile Leu Gly Leu Gly Pro Phe Pro
335      340      345
Thr Glu Asp Glu Val Asp Ala Asp Leu Ile Asn Ala Gly Lys Gln
350      355      360
Thr Val Thr Val Leu Pro Gly Gly Cys Phe Phe Ala Ser Asp Asp
365      370      375
Ser Phe Ala Met Ile Arg Gly Gly His Ile Gln Leu Thr Met Leu
380      385      390
Gly Ala Met Gln Val Ser Lys Tyr Gly Asp Leu Ala Asn Trp Met
395      400      405
Ile Pro Gly Lys Lys Val Lys Gly Met Gly Gly Ala Met Asp Leu
410      415      420
Val Ser Ser Gln Lys Thr Arg Val Val Val Thr Met Gln His Cys
425      430      435
Thr Lys Asp Asn Thr Pro Lys Ile Met Glu Lys Cys Thr Met Pro
440      445      450
Leu Thr Gly Lys Arg Cys Val Asp Arg Ile Ile Thr Glu Lys Ala
455      460      465
Val Phe Asp Val His Arg Lys Lys Glu Leu Thr Leu Arg Glu Leu
470      475      480
Trp Glu Gly Leu Thr Val Asp Asn Ile Lys Lys Ser Thr Gly Cys
485      490      495
Ala Phe Ala Val Ser Pro Asn Leu Arg Pro Met Gln Gln Val Ala
500      505      510
Pro

```

```

<210> 5
<211> 99
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1274566CD1

```

```

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

```

```

<210> 6
<211> 389
<212> PRT
<213> Homo sapiens

```

WO 01/79291

PCT/US01/11861

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1349442CD1  
  
 <400> 6  
 Met Arg Gly Gly Lys Cys Asn Met Leu Ser Ser Leu Gly Cys Leu  
 1 5 10 15  
 Leu Leu Cys Gly Ser Ile Thr Leu Ala Leu Gly Asn Ala Gln Lys  
 20 25 30  
 Leu Pro Lys Gly Lys Arg Pro Asn Leu Lys Val His Ile Asn Thr  
 35 40 45  
 Thr Ser Asp Ser Ile Leu Leu Lys Phe Leu Arg Pro Ser Pro Asn  
 50 55 60  
 Val Lys Leu Glu Gly Leu Leu Leu Gly Tyr Gly Ser Asn Val Ser  
 65 70 75  
 Pro Asn Gln Tyr Phe Pro Leu Pro Ala Glu Gly Lys Phe Thr Glu  
 80 85 90  
 Ala Ile Val Asp Ala Glu Pro Lys Tyr Leu Ile Val Val Arg Pro  
 95 100 105  
 Ala Pro Pro Pro Ser Gln Lys Lys Ser Cys Ser Gly Lys Thr Arg  
 110 115 120  
 Ser Arg Lys Pro Leu Gln Leu Val Val Gly Thr Leu Thr Pro Ser  
 125 130 135  
 Ser Val Phe Leu Ser Trp Gly Phe Leu Ile Asn Pro His His Asp  
 140 145 150  
 Trp Thr Leu Pro Ser His Cys Pro Asn Asp Arg Phe Tyr Thr Ile  
 155 160 165  
 Arg Tyr Arg Glu Lys Asp Lys Glu Lys Lys Trp Ile Phe Gln Ile  
 170 175 180  
 Cys Pro Ala Thr Glu Thr Ile Val Glu Asn Leu Lys Pro Asn Thr  
 185 190 195  
 Val Tyr Glu Phe Gly Val Lys Asp Asn Val Glu Gly Gly Ile Trp  
 200 205 210  
 Ser Lys Ile Phe Asn His Lys Thr Val Val Gly Ser Lys Lys Val  
 215 220 225  
 Asn Gly Lys Ile Gln Ser Thr Tyr Asp Gln Asp His Thr Val Pro  
 230 235 240  
 Ala Tyr Val Pro Arg Lys Leu Ile Pro Ile Thr Ile Ile Lys Gln  
 245 250 255  
 Val Ile Gln Asn Val Thr His Lys Asp Ser Ala Lys Ser Pro Glu  
 260 265 270  
 Lys Ala Pro Leu Gly Gly Val Ile Leu Val His Leu Ile Ile Pro  
 275 280 285  
 Gly Leu Asn Glu Thr Thr Val Lys Leu Pro Ala Ser Leu Met Phe  
 290 295 300  
 Glu Ile Ser Asp Ala Leu Lys Thr Gln Leu Ala Lys Asn Glu Thr  
 305 310 315  
 Leu Ala Leu Pro Ala Glu Ser Lys Thr Pro Glu Val Glu Lys Ile  
 320 325 330  
 Ser Ala Arg Pro Thr Thr Val Thr, Pro Glu Thr Val Pro Arg Ser  
 335 340 345  
 Thr Lys Pro Thr Thr Ser Ser Ala Leu Asp Val Ser Glu Thr Thr  
 350 355 360  
 Leu Val Leu Ser Lys Arg Thr Pro Glu Thr Leu Gln Thr Ile Leu  
 365 370 375  
 Ile Pro Gln Phe Glu Leu Pro Leu Ser Thr Leu Gly Lys Lys  
 380 385

<210> 7  
 <211> 322  
 <212> PRT  
 <213> Homo sapiens

WO 01/79291

PCT/US01/11861

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1400156CD1  
  
 <400> 7  
 Met Ala Leu Pro Pro Gly Pro Ala Ala Leu Arg His Thr Leu Leu  
 1 5 10 15  
 Leu Leu Pro Ala Leu Leu Ser Ser Gly Trp Gly Glu Leu Glu Pro  
 20 25 30  
 Gln Ile Asp Gly Gln Thr Trp Ala Glu Arg Ala Leu Arg Glu Asn  
 35 40 45  
 Glu Arg His Ala Phe Thr Cys Arg Val Ala Gly Gly Pro Gly Thr  
 50 55 60  
 Pro Arg Leu Ala Trp Tyr Leu Asp Gly Gln Leu Gln Glu Ala Ser  
 65 70 75  
 Thr Ser Arg Leu Leu Ser Val Gly Gly Glu Ala Phe Ser Gly Gly  
 80 85 90  
 Thr Ser Thr Phe Thr Val Thr Ala His Arg Ala Gln His Glu Leu  
 95 100 105  
 Asn Cys Ser Leu Gln Asp Pro Arg Ser Gly Arg Ser Ala Asn Ala  
 110 115 120  
 Ser Val Ile Leu Asn Val Gln Phe Lys Pro Glu Ile Ala Gln Val  
 125 130 135  
 Gly Ala Lys Tyr Gln Glu Ala Gln Gly Pro Gly Leu Leu Val Val  
 140 145 150  
 Leu Phe Ala Leu Val Arg Ala Asn Pro Pro Ala Asn Val Thr Trp  
 155 160 165  
 Ile Asp Gln Asp Gly Pro Val Thr Val Asn Thr Ser Asp Phe Leu  
 170 175 180  
 Val Leu Asp Ala Gln Asn Tyr Pro Trp Leu Thr Asn His Thr Val  
 185 190 195  
 Gln Leu Gln Leu Arg Ser Leu Ala His Asn Leu Ser Val Val Ala  
 200 205 210  
 Thr Asn Asp Val Gly Val Thr Ser Ala Ser Leu Pro Ala Pro Gly  
 215 220 225  
 Pro Ser Arg His Pro Ser Leu Ile Ser Ser Asp Ser Asn Asn Leu  
 230 235 240  
 Lys Leu Asn Asn Val Arg Leu Pro Arg Glu Asn Met Ser Leu Pro  
 245 250 255  
 Ser Asn Leu Gln Leu Asn Asp Leu Thr Pro Asp Ser Arg Ala Val  
 260 265 270  
 Lys Pro Ala Asp Arg Gln Met Ala Gln Asn Ser Arg Pro Glu  
 275 280 285  
 Leu Leu Asp Pro Glu Pro Gly Gly Leu Leu Thr Ser Gln Gly Phe  
 290 295 300  
 Ile Arg Leu Pro Val Leu Gly Tyr Ile Tyr Arg Val Ser Ser Val  
 305 310 315  
 Ser Ser Asp Glu Ile Trp Leu  
 320  
  
 <210> 8  
 <211> 587  
 <212> PRT  
 <213> Homo sapiens  
  
 <220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1610347CD1  
  
 <400> 8  
 Met His Pro Leu Gln Cys Val Leu Gln Val Gln Arg Ser Leu Gly  
 1 5 10 15  
 Trp Gly Pro Leu Ala Ser Val Ser Trp Leu Ser Leu Arg Met Cys

WO 01/79291

PCT/US01/11861

|                 |                         |                     |
|-----------------|-------------------------|---------------------|
| 20              | 25                      | 30                  |
| Arg Ala His Ser | Ser Leu Ser Ser Thr Met | Cys Pro Ser Pro Glu |
| 35              | 40                      | 45                  |
| Arg Gln Glu Asp | Gly Ala Arg Lys Asp Phe | Ser Ser Arg Leu Ala |
| 50              | 55                      | 60                  |
| Ala Gly Pro Thr | Phe Gln His Phe Leu Lys | Ser Ala Ser Ala Pro |
| 65              | 70                      | 75                  |
| Gln Glu Lys Leu | Ser Ser Glu Val Glu Asp | Pro Pro Pro Tyr Leu |
| 80              | 85                      | 90                  |
| Met Met Asp Glu | Leu Leu Gly Arg Gln Arg | Lys Val Tyr Leu Glu |
| 95              | 100                     | 105                 |
| Thr Tyr Gly Cys | Gln Met Asn Val Asn Asp | Thr Glu Ile Ala Trp |
| 110             | 115                     | 120                 |
| Ser Ile Leu Gln | Lys Ser Gly Tyr Leu Arg | Thr Ser Asn Leu Gln |
| 125             | 130                     | 135                 |
| Glu Ala Asp Val | Ile Leu Leu Val Thr Cys | Ser Ile Arg Glu Lys |
| 140             | 145                     | 150                 |
| Ala Glu Gln Thr | Ile Trp Asn Arg Leu His | Gln Leu Lys Ala Leu |
| 155             | 160                     | 165                 |
| Lys Thr Arg Arg | Pro Arg Ser Arg Val Pro | Leu Arg Ile Gly Ile |
| 170             | 175                     | 180                 |
| Leu Gly Cys Met | Ala Glu Arg Leu Lys Glu | Glu Ile Leu Asn Arg |
| 185             | 190                     | 195                 |
| Glu Lys Met Val | Asp Ile Leu Ala Gly Pro | Asp Ala Tyr Arg Asp |
| 200             | 205                     | 210                 |
| Leu Pro Arg Leu | Leu Ala Val Ala Glu Ser | Gly Gln Gln Ala Ala |
| 215             | 220                     | 225                 |
| Asn Val Leu Leu | Ser Leu Asp Glu Thr Tyr | Ala Asp Val Met Pro |
| 230             | 235                     | 240                 |
| Val Gln Thr Ser | Ala Ser Ala Thr Ser Ala | Phe Val Ser Ile Met |
| 245             | 250                     | 255                 |
| Arg Gly Cys Asp | Asn Met Cys Ser Tyr Cys | Ile Val Pro Phe Thr |
| 260             | 265                     | 270                 |
| Arg Gly Arg Glu | Arg Ser Arg Pro Ile Ala | Ser Ile Leu Glu Glu |
| 275             | 280                     | 285                 |
| Val Lys Lys Leu | Ser Glu Gln Gly Leu Lys | Glu Val Thr Leu Leu |
| 290             | 295                     | 300                 |
| Gly Gln Asn Val | Asn Ser Phe Arg Asp Asn | Ser Glu Val Gln Phe |
| 305             | 310                     | 315                 |
| Asn Ser Ala Val | Pro Thr Asn Leu Ser Arg | Gly Phe Thr Thr Asn |
| 320             | 325                     | 330                 |
| Tyr Lys Thr Lys | Gln Gly Gly Leu Arg Phe | Ala His Leu Leu Asp |
| 335             | 340                     | 345                 |
| Gln Val Ser Arg | Val Asp Pro Glu Met Arg | Ile Arg Phe Thr Ser |
| 350             | 355                     | 360                 |
| Pro His Pro Lys | Asp Phe Pro Asp Glu Val | Leu Gln Leu Ile His |
| 365             | 370                     | 375                 |
| Glu Arg Asp Asn | Ile Cys Lys Gln Ile His | Leu Pro Ala Gln Ser |
| 380             | 385                     | 390                 |
| Gly Ser Ser Arg | Val Leu Glu Ala Met Arg | Arg Gly Tyr Ser Arg |
| 395             | 400                     | 405                 |
| Glu Ala Tyr Val | Glu Leu Val His His Ile | Arg Glu Ser Ile Pro |
| 410             | 415                     | 420                 |
| Gly Val Ser Leu | Ser Ser Asp Phe Ile Ala | Gly Phe Cys Gly Glu |
| 425             | 430                     | 435                 |
| Thr Glu Glu Asp | His Val Gln Thr Val Ser | Leu Leu Arg Glu Val |
| 440             | 445                     | 450                 |
| Gln Tyr Asn Met | Gly Phe Leu Phe Ala Tyr | Ser Met Arg Gln Lys |
| 455             | 460                     | 465                 |
| Thr Arg Ala Tyr | His Arg Leu Lys Asp Asp | Val Pro Glu Glu Val |
| 470             | 475                     | 480                 |
| Lys Leu Arg Arg | Leu Glu Glu Leu Ile Thr | Ile Phe Arg Glu Glu |
| 485             | 490                     | 495                 |

WO 01/79291

PCT/US01/11861

Ala Thr Lys Ala Asn Gln Thr Ser Val Gly Cys Thr Gln Leu Val  
 500 505 510  
 Leu Val Glu Gly Leu Ser Lys Arg Ser Ala Thr Asp Leu Cys Gly  
 515 520 525  
 Arg Asn Asp Gly Asn Leu Lys Val Ile Phe Pro Asp Ala Glu Met  
 530 535 540  
 Glu Asp Val Asn Asn Pro Gly Leu Arg Val Arg Ala Gln Pro Gly  
 545 550 555  
 Asp Tyr Val Leu Val Lys Ile Thr Ser Ala Ser Ser Gln Thr Leu  
 560 565 570  
 Arg Gly His Val Leu Cys Arg Thr Thr Leu Arg Asp Ser Ser Ala  
 575 580 585  
 Tyr Cys

<210> 9  
 <211> 173  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 187209CD1

<400> 9  
 Met Glu Glu Met Arg Pro Ala Gly His Gly Val Ser Asn Val Cys  
 1 5 10 15  
 Val Ala Phe Lys Val Ala Cys His Ser Cys Leu Pro Arg Leu Phe  
 20 25 30  
 Asn Ala Leu Ile Pro Ser Pro Asp Arg Asn Gly Ala Ala Leu Leu  
 35 40 45  
 Gly Gly Gln Ala Ser Ala Asp Ser Lys Ser Glu Ala Arg Arg Asn  
 50 55 60  
 Gln Cys Asp Ser Met Leu Leu Arg Asn Gln Gln Leu Cys Ser Thr  
 65 70 75  
 Cys Gln Glu Met Lys Met Val Gln Pro Arg Thr Met Lys Ile Pro  
 80 85 90  
 Asp Asp Pro Lys Ala Ser Phe Glu Asn Cys Met Ser Tyr Arg Met  
 95 100 105  
 Ser Leu His Gln Pro Lys Phe Gln Thr Thr Pro Glu Pro Phe His  
 110 115 120  
 Asp Asp Ile Pro Thr Glu Asn Ile His Tyr Arg Leu Pro Ile Leu  
 125 130 135  
 Gly Pro Arg Thr Ala Val Phe His Gly Leu Leu Thr Glu Ala Tyr  
 140 145 150  
 Lys Thr Leu Lys Glu Arg Gln Arg Ser Ser Leu Pro Arg Lys Glu  
 155 160 165  
 Pro Ile Gly Lys Thr Thr Arg Gln  
 170

<210> 10  
 <211> 325  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 2607963CD1

<400> 10  
 Met Gln Gly Arg Val Ala Gly Ser Cys Ala Pro Leu Gly Leu Leu  
 1 5 10 15  
 Leu Val Cys Leu His Leu Pro Gly Leu Phe Ala Arg Ser Ile Gly

WO 01/79291

PCT/US01/11861

```

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

```

&lt;210&gt; 11

&lt;211&gt; 733

&lt;212&gt; PRT

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 412044CD1

&lt;400&gt; 11

```

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

```

WO 01/79291

PCT/US01/11861

|                 |                                             |     |
|-----------------|---------------------------------------------|-----|
| 80              | 85                                          | 90  |
| Ile Phe Thr Thr | Pro Asn Phe Asp Glu Thr Ser Ser Ala Thr Thr |     |
| 95              | 100                                         | 105 |
| Ile Ser Thr Ser | Glu Asp Gly Ser Tyr Gly Thr Asp Val Thr Arg |     |
| 110             | 115                                         | 120 |
| Cys Ile Cys Gly | Phe Thr His Asp Asp Gly Tyr Met Ile Cys Cys |     |
| 125             | 130                                         | 135 |
| Asp Lys Cys Ser | Val Trp Gln His Ile Asp Cys Met Gly Ile Asp |     |
| 140             | 145                                         | 150 |
| Arg Gln His Ile | Pro Asp Thr Tyr Leu Cys Glu Arg Cys Gln Pro |     |
| 155             | 160                                         | 165 |
| Arg Asn Leu Asp | Lys Glu Arg Ala Val Leu Leu Gln Arg Arg Lys |     |
| 170             | 175                                         | 180 |
| Arg Glu Asn Met | Ser Asp Gly Asp Thr Ser Ala Thr Glu Ser Gly |     |
| 185             | 190                                         | 195 |
| Asp Glu Val Pro | Val Glu Leu Tyr Thr Ala Phe Gln His Thr Pro |     |
| 200             | 205                                         | 210 |
| Thr Ser Ile Thr | Leu Thr Ala Ser Arg Val Ser Lys Val Asn Asp |     |
| 215             | 220                                         | 225 |
| Lys Arg Arg Lys | Lys Ser Gly Glu Lys Glu Gln His Ile Ser Lys |     |
| 230             | 235                                         | 240 |
| Cys Lys Lys Ala | Phe Arg Glu Gly Ser Arg Lys Ser Ser Arg Val |     |
| 245             | 250                                         | 255 |
| Lys Gly Ser Ala | Pro Glu Ile Asp Pro Ser Ser Asp Gly Ser Asn |     |
| 260             | 265                                         | 270 |
| Phe Gly Trp Glu | Thr Lys Ile Lys Ala Trp Met Asp Arg Tyr Glu |     |
| 275             | 280                                         | 285 |
| Glu Ala Asn Asn | Asn Gln Tyr Ser Glu Gly Val Gln Arg Glu Ala |     |
| 290             | 295                                         | 300 |
| Gln Arg Ile Ala | Leu Arg Leu Gly Asn Gly Asn Asp Lys Lys Glu |     |
| 305             | 310                                         | 315 |
| Met Asn Lys Ser | Asp Leu Asn Thr Asn Asn Leu Leu Phe Lys Pro |     |
| 320             | 325                                         | 330 |
| Pro Val Glu Ser | His Ile Gln Lys Asn Lys Lys Ile Leu Lys Ser |     |
| 335             | 340                                         | 345 |
| Ala Lys Asp Leu | Pro Pro Asp Ala Leu Ile Ile Glu Tyr Arg Gly |     |
| 350             | 355                                         | 360 |
| Lys Phe Met Leu | Arg Glu Gln Phe Glu Ala Asn Gly Tyr Phe Phe |     |
| 365             | 370                                         | 375 |
| Lys Arg Pro Tyr | Pro Phe Val Leu Phe Tyr Ser Lys Phe His Gly |     |
| 380             | 385                                         | 390 |
| Leu Glu Met Cys | Val Asp Ala Arg Thr Phe Gly Asn Glu Ala Arg |     |
| 395             | 400                                         | 405 |
| Phe Ile Arg Arg | Ser Cys Thr Pro Asn Ala Glu Val Arg His Glu |     |
| 410             | 415                                         | 420 |
| Ile Gln Asp Gly | Thr Ile His Leu Tyr Ile Tyr Ser Ile His Ser |     |
| 425             | 430                                         | 435 |
| Ile Pro Lys Gly | Thr Glu Ile Thr Ile Ala Phe Asp Phe Asp Tyr |     |
| 440             | 445                                         | 450 |
| Gly Asn Cys Lys | Tyr Lys Val Asp Cys Ala Cys Leu Lys Glu Asn |     |
| 455             | 460                                         | 465 |
| Pro Glu Cys Pro | Val Leu Lys Arg Ser Ser Glu Ser Met Glu Asn |     |
| 470             | 475                                         | 480 |
| Ile Asn Ser Gly | Tyr Glu Thr Arg Arg Lys Lys Gly Lys Lys Asp |     |
| 485             | 490                                         | 495 |
| Lys Asp Ile Ser | Lys Glu Lys Asp Thr Gln Asn Gln Asn Ile Thr |     |
| 500             | 505                                         | 510 |
| Leu Asp Cys Glu | Gly Thr Thr Asn Lys Met Lys Ser Pro Glu Thr |     |
| 515             | 520                                         | 525 |
| Lys Gln Arg Lys | Leu Ser Pro Leu Arg Leu Ser Val Ser Asn Asn |     |
| 530             | 535                                         | 540 |
| Gln Glu Pro Asp | Phe Ile Asp Asp Ile Glu Glu Lys Thr Pro Ile |     |
| 545             | 550                                         | 555 |

WO 01/79291

PCT/US01/11861

```

Ser Asn Glu Val Glu Met Glu Ser Glu Glu Gln Ile Ala Glu Arg
560 565
Lys Arg Lys Met Thr Arg Glu Glu Arg Lys Met Glu Ala Ile Leu
575 580
Gln Ala Phe Ala Arg Leu Glu Lys Arg Glu Lys Arg Arg Glu Gln
590 595
Ala Leu Glu Arg Ile Ser Thr Ala Lys Thr Glu Val Lys Thr Glu
605 610
Cys Lys Asp Thr Gln Ile Val Ser Asp Ala Glu Val Ile Gln Glu
620 625
Gln Ala Lys Glu Glu Asn Ala Ser Lys Pro Thr Pro Ala Lys Val
635 640
Asn Arg Thr Lys Gln Arg Lys Ser Phe Ser Arg Ser Arg Thr His
650 655
Ile Gly Gln Gln Arg Arg Arg His Arg Thr Val Ser Met Cys Ser
665 670
Asp Ile Gln Pro Ser Ser Pro Asp Ile Glu Val Thr Ser Gln Gln
680 685
Asn Asp Ile Glu Asn Thr Val Leu Thr Ile Glu Pro Glu Thr Glu
695 700
Thr Ala Leu Ala Glu Ile Ile Thr Glu Thr Glu Val Pro Ala Leu
710 715
Asn Lys Cys Pro Thr Lys Tyr Pro Lys Thr Lys Lys Val
725 730

```

<210> 12  
 <211> 242  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 638118CD1

```

<400> 12
Met Pro Pro Arg Leu Pro Pro Met Pro Ala Val Leu Gly Lys Leu
1 5 10 15
Pro Arg Thr Leu Gly Glu Arg Pro Glu Asn Leu Arg Arg Lys Pro
20 25 30
Pro Gly Leu Leu Ala Thr Cys Ser Val Ser Leu Pro Ala Pro Leu
35 40 45
Pro Ser Gly Ile Arg Lys Arg Ala Gly Pro Cys Ala Pro Ser Pro
50 55 60
Leu Pro Arg Ala Ala Asn Asn Thr Pro Pro Trp Gly Ala Ser Phe
65 70 75
Leu Leu Trp Lys Leu Arg His Trp Thr Glu Gly Thr Gly Leu Arg
80 85 90
Gly Ala Asp Arg Gly Pro Val Leu Leu Gly Ala Leu Arg Thr Arg
95 100 105
Gly Arg Arg Gly His Gly Gln Glu Pro Gln Pro Arg Val Leu Ala
110 115 120
Phe Leu Leu Arg Arg Ser Pro Pro Lys Ser Thr Gln Arg Leu Glu
125 130 135
Gln Pro Ser Thr Gln Pro Glu Glu Gly Arg Ala Pro Pro Pro Ala
140 145 150
Leu Gly Gly Gly Val Trp Pro Phe Leu Pro Phe Pro Arg Pro Pro
155 160 165
Glu Ala Pro Thr Gln Phe Ser Val Thr Ser Ser Gly Arg Lys Ala
170 175 180
Ser Arg Cys Pro Pro Glu Leu Leu Trp Ala Gln Gly Trp Leu Arg
185 190 195
Asp His Leu Met Asp Val Leu Gly Ser Met Gly Ser Gln Gly Ser
200 205 210

```

WO 01/79291

PCT/US01/11861

Ile Pro Ser Cys Ser Pro Thr Pro Pro Gln Leu Pro Gly Gly Trp  
 215 220 225  
 Ala His Glu Gly Ser Gly Asp Thr Ser Ile Gly Lys Gly Pro Gly  
 230 235 240  
 Thr Leu

<210> 13  
 <211> 153  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 743323CD1

<400> 13  
 Met Ser Ala Val Phe Gly Arg Pro His Ala Cys Gln Pro His Ala  
 1 5 10 15  
 Val Leu Leu Arg Leu Phe Pro Ser His Pro Ser Gly Cys Leu Thr  
 20 25 30  
 Pro Leu Thr Ala Ser Leu Ser Cys His Leu Arg Ala Ala Ser Gly  
 35 40 45  
 Asn Arg Lys Thr Gly Leu Cys Pro Ser Ile Asn Pro Phe Ile His  
 50 55 60  
 Lys Phe Ser Ile Ser Met Ser Pro Gly Glu Leu Gln Gly Cys Ser  
 65 70 75  
 Gln Glu Pro Arg Ser Gln Gly Trp Ser Trp Leu Cys Cys Cys Thr  
 80 85 90  
 Arg Ala Ala Phe Pro Thr Phe Ser Arg Gly Thr Cys Ser Thr Ala  
 95 100 105  
 Arg Arg Thr Ser Thr Glu His Pro Glu Gly Ser Arg Pro Arg Pro  
 110 115 120  
 Gln Gly Thr Pro Arg Pro Leu Gln Arg Gly Pro Val Ser Gly Ser  
 125 130 135  
 Leu Gly Ala Val Val Leu Arg Gly His Ile Pro Ala Glu Trp Pro  
 140 145 150  
 Cys Ser Val

<210> 14  
 <211> 134  
 <212> PRT  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1691509CD1

<400> 14  
 Met Tyr Ser Ala Met Met Phe Leu Phe Gln Leu Ile Leu Gly Ile  
 1 5 10 15  
 Pro Glu Gln Ala Leu Ser Leu Leu His Met Ala Ile Glu Pro Ile  
 20 25 30  
 Leu Ala Asp Gly Ala Ile Leu Asp Lys Gly Arg Ala Met Phe Leu  
 35 40 45  
 Val Ala Lys Cys Gln Val Ala Ser Ala Ala Ser Tyr Asp Gln Pro  
 50 55 60  
 Lys Lys Ala Glu Ala Leu Glu Ala Ala Ile Glu Asn Leu Asn Glu  
 65 70 75  
 Ala Lys Asn Tyr Phe Ala Lys Val Asp Cys Lys Glu Arg Ile Arg  
 80 85 90  
 Asp Val Val Tyr Phe Gln Ala Arg Leu Tyr His Thr Leu Gly Lys

WO 01/79291

PCT/US01/11861

95 100 105  
 Thr Gln Glu Arg Asn Arg Cys Ala Met Leu Phe Arg Gln Leu His  
 110 115 120  
 Gln Glu Leu Pro Ser His Gly Val Pro Leu Ile Asn His Leu  
 125 130

<210> 15  
 <211> 1566  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7473577CB1

<400> 15  
 atgctgtgtg cctgtctctt cctgtccagc ctctgtgggg ccaccagggc cagccccacc 60  
 tcaggccccc aggagtgtgc aaagggtccc acggtgtggt gtcaggatct gcagacagct 120  
 gccagggtggc gggctgtggg gtactgtcaa ggggcccgtat ggaacaaacc caccgcgaag 180  
 tctctgcctc gcgacgtatg ccaggacata gcagccgcgc ctggcaatgg gctgaacctc 240  
 cagccacagg agtctgacat cctgtctttg gtgatgaaga cctgtgagtg gctccccagc 300  
 caggagtctt cagccggatg caagtggatg gtggatgccc acagttcggc catcctgagc 360  
 atgtctctgt gggccccgga cagtgtcccc gcacagggtg gcacagcgct cagcctctgt 420  
 agcccgctgc agaggcacct ggccaccctg aggccactct ccaagaggga cacccttgag 480  
 gctgtggctc cgttcattgg caatggggccc cttaccttcc acccccgcca ggcgcttgaa 540  
 ggaactctgt gccaaagactg tgtacggcag gtctcccgac tccaggaggg tgtccggctc 600  
 aacttgacct tggccgactt gaacatccag gagcagtggt agtccttggg gctggtgctg 660  
 gccgtctctt gcaagaacta cctcttccag tttttgttcc ctgtgacca agcactgagg 720  
 ctcttcccccc cgcaggagct ctgcagggaag gggggattct gtgaggagct aggggcacct 780  
 gcccggttga ctcaagtagt ggccatggac ggggtcccc cctggagctc ggggttgcca 840  
 aggaacacga gcgagatgca gatgaaggcc ggtatgacct gtgaggtgtg catgaaactg 900  
 gtgcagaagc tggaccactg gctcatgtcc aacagctctg agctcatgat caccatgccc 960  
 ctggagcgcg tgtgctcggt aatgcctgcc tctatcaca aggagtgcac catcttggtg 1020  
 gacacctaca gccctccttt ggtgcagctt gtggccaaaa tcaccccgag gaaggtgtgc 1080  
 aagttcatcc gtctgtgtgg caaccggagg cgggcccggg cagtccatga tgcctatgcc 1140  
 atcgtgcctg cccagagtg gtacggggag aaccagggca gcttctgcaa tgggtgcaag 1200  
 aggtctgtca cgggtctctc ccacaacctg gagagcaaga gcaccaagcg agacatcctg 1260  
 gtggccttca aggtgtgctg cagcatcctg ccgctgcctt atatgtacca gtgcaagcac 1320  
 ttctcacacc agtacgagcc cgtgtctatt gagagtctca aggacatgat ggaccccgct 1380  
 cgtgtgtgca agaaggtggg ggcctgtccac gggcccagga cccactgctc gggcacccag 1440  
 cagtgtgccc tgggcccagg ctctgtgtgc aggagccagg aggcggccaa gctgtgcaac 1500  
 gctgtgcaac actgccagaa gcatgtatgg aaagagatgc acctccacgc tggggaacac 1560  
 gcgtga 1566

<210> 16  
 <211> 939  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 7474024CB1

<400> 16  
 ggcgcgctcg cctccctcgc tccacgcggc cccggacgcy ggggcaggc ttgcgcgcgg 60  
 ttccccctccc ggtgggcgga ttccctggga agatgaagtg ggtgtgggag ctcttctgtg 120  
 tggcggcgctc ggcagcgggc cgcgcggagc gcgactgcgc agtgagcagc ttccgagtca 180  
 aggagaacctc cagacaaggct cgcttctctg ggaccttgta cgcctatggc aagaaggacc 240  
 ccgaggggctc cttctgtcag gacaacatcg tcgcggagtt ctccgtggac gagaccggcc 300  
 agatgagcgc cacagccaag ggcagagtcg gttctttgaa taactgggac gtgtgcgcag 360  
 acatgggtggg caccttcaca gacaccgagg accctgcca gttcaagatg aagtactggg 420  
 gcgtagcctc cttctccag aaaggaaatg atgaccactg gatcgtcgac acagactacy 480  
 acacgtatgc cgtgcagtac tctgcgcgcc tctggaacct cgatggcacc tgtgtgtaca 540

WO 01/79291

PCT/US01/11861

```

gtctactcctt cgtgttttcc cgggacccca acggcctgcc cccagaagcg cagaagattg 600
taaggcagcg gcaggaggag ctgtgcctgg ccaggcagta caggctgacg gtcacaaacg 650
gttactgaga tggcagatca gaaagaaacc ttttgttaga atatcaagaa tcttagtttca 720
ctctgagaact tctgattagc tctcagctct cagctctatt tatcttagga gtttaatttg 780
ccctctcttc cccatcttcc ctccagttccc ataaaaacct cattacacat aaagatacac 840
gtgggggtca gtgaattctg tgccttttcc tgaagaattt tggggcttaa gattccagac 900
tctgattcat taaactatag tcaccctgga aaaaaaaaa 939

```

```

<210> 17
<211> 2785
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Inyte ID No: 2480555CB1

```

```

<400> 17
ctgattctcca ggaccagcac tcttctccca gcccttaggg tcttgcctgg ccaaggcctt 60
ccctgccatg cgacctgtca gtgtctggca gtggagcccc tgggggctgc tgcgtgcct 120
gctgtgcagt tctgtcttgg ggtctccgtc ccttctccag gcccttgaga agaagcccg 180
ggccagaggg cttcggttcc ggtcggctgg cttcccagag aagccttacc agggcccgct 240
ggagatcacg cgagctggtg aatggggcac catctcgat gatgacttca cgtcgaggcg 300
tgcacacatc ccttgccggg agctgggctt cacagagggc acaggctgga cccacagtcg 360
caaatatggc cctggaacag gccgcactcg gctggacaac ttgagctgca gtgggaccga 420
gcagagtgtg actgaatgtg cctccggggg ctgggggaaac agtgactgta cgcacgatga 480
ggatgctggg gtcatctgca aagaccagcg cctccctggc ttctcggact ccaatgtcat 540
tgaggtagag catcacctgc aagtggaggga ggtgcgaatt cgaccccgcc tggggtgggg 600
cagacgaccc ctgcccgtga cggaggggct ggtggaagt aggcctctcg acgctggctg 660
gcaagtgtgc gacaaagctt ggaagcgcca caacagccac gtggtctggg ggaatgctgg 720
cttcccagcg gaaaagaggg tcaacgcggc cttctacagg ctgtagcccc aacggcagca 780
acactccttt ggtctgcatg ggtggcgtg cgtgggcacg gaggccccc tctccctctg 840
ttccctggag ttctatctg ccaatgacac cgccaggctg cctggggggg gccctgcagt 900
ggtgagctgt gtgccaggcc ctgtctacgc ggcatccagt ggcccaaga agcaacaaca 960
gtcgaagcct cagggggagg ccctgttccg tctaaagggt ggcgccccc cttgagaggg 1020
ccgggtagaa gtccctgaagg ccagcacatg gggcacagt ttgtaccgca agtgggacct 1080
gcattgcagc agctgtgtgt gtccggagct gggcttcggg agtgctcgag aagctctgag 1140
tggcgtctcg atggggcagg gcatgggtgc tatccacctg agtgaagttc gctgctctg 1200
acaggagctc tccctctgga agtgccecca caagaacatc acagctgagg attgttaca 1260
tagccaggat gccggggtcc ggtgcaacct accttacact ggggcagaga ccaggatccg 1320
actcagtggg ggccgcagcc aactgaggg gcgagtcgag gtgcaaatag ggggacctgg 1380
ggccctctcg tggggcttca tctgtgggga tgactgggg accctggagg ccatgtgtggc 1440
ctgtaggcaa ctgggtctgg gctacgccaa ccaacggctg caggagacct ggtactggga 1500
ctctgggaat ataacagagg tggtagtag tggagtgcgc tgcacaggga ctgagctgtc 1560
cctggatcag tgtgccatc atggcaccga catcacctgc aagaggacag ggaaccgctt 1620
caactctgga gtcatctgt ctgagactgc atcagatctg ttgtgctact cagcactggt 1680
gcaggagacc gccatcatcg aagaccggcc cctgcatatg ttgtactgtg ctgcgggaaga 1740
gaactgcctg gccagctcag cccgctcagc caactggccc tatggtcacc ggcgtctgct 1800
ccgattctcc tccagatccc acaacctggg acgagctgac ttacggccca aggtggggcg 1860
ccactcctgg gtgtggcacg agtgccatgg gcattaccac agcatggaca tcttcaacta 1920
ctatgatatc ctaccccaca atggcaccga ggtggtgag ggccacaaag ctagtttctg 1980
tctgaaagac actgagtgtc agggagatgt ctccaagcgg tatgagtgtg ccaactttgg 2040
agagcaaggc atcactgtgg gtgtctggga tctctaccgg catgacaltg actgtcagtg 2100
gattgacatc acgagtgtga agccaggaaa ctacatttc caggttgtca tcaacccaaa 2160
ctttgaagta gcagagagtg actttaccaa caatgcaatg aatgttaact gcaaatatga 2220
tggacataga atctgggtgc acaactgccca cattggtgat gccctcagtg aagaggccaa 2280
cagggggttt gaacgctacc ctggccagac cagcaaccag attatctaag tgccactgcc 2340
ctctgcaaac caccactggc ccctaatggc aggggtctga ggctgccatt acctcaggag 2400
cttaccaaaga aacctatgtc agcaaccgca ctcatcagac catgcaactt ggaatgtgaa 2460
ctgtcaagca gaagttttca cctctcttca gaggcagct gtgagtatct gtgaccaagc 2520
atgggaaatc ttgctccag gccccagcac gagcagaaca gaccagagcc caccacacca 2580
caaagagcag cacttgacta actgcccaca aaagatggca gcagctcatt tctttaata 2640
ggaggtcagg atggtcagct ccaagtatct ccctaagttt aggggggatac agctttacct 2700

```

WO 01/79291

PCT/US01/11861

ctagcctttt ggtgggggaa aagatccagc cctccacact cattttttac tataatatgt 2760  
 gaatagcaca agtattttata taaaa 2785

<210> 18  
 <211> 1733  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 3187086CB1

<400> 18  
 cgggcccgaact atggcggcgc tgcggctcct ggcgtcagtg ctggggcgcg ggggtccccg 60  
 cggcgggctca gggctcgcgc tgtcccaggg ctgcgcgcgc tgcctttgcca ccagtcgccg 120  
 gctccgtgccc aagttctctacg cggacccgggt ggagatgggt aaggacatct ctgacggggc 180  
 gaccgtcatg atcggggggct tcgggctctg cgggatlcccc gagaacctga tcgcgcgctt 240  
 gctcaggacc cgcgtgaaag acctgcaggt ggtcagcagc aacgtggggc tggaggacct 300  
 cggcctgggc ctctcgtctg cgcgcaggca ggtcgtctgc atcgtctgtt cctacgtggg 360  
 cgagaaacacc ctgtgcgaga gccagtacct ggcagagag ctggagctgg agctcacgcc 420  
 ccagggcacc ctggccgagc gcacccgcgc gtggggcgcc ggggtgcccc cttctacac 480  
 cccacaggcg tacgggaacc tggteccagg agggggcgcc cccatccgct acaccocga 540  
 cggccaccctg gcgctcatga gccagcccg agaggtgagg gaggttcaac ggcacacct 600  
 ccttttggag cgcgccatcc ggcgcagactt cgcctgggt aaagggtgga aggcgacgc 660  
 ggcaggaaac gtggtcttca ggagaagcgc cgcgaattc aacgtgcccc tgtgcaaac 720  
 tgcagacgtc tacggcgggt gagggtgggg ctccccca gaagacatcc acgttctcaa 780  
 catttatgta ggtcgcgtga taaaggggca gaaatacag aaacgaattg agcgtctaac 840  
 gatccggaaa gaggaagatg gagacgctgg aaaggaagag gacgccagga cgcgcatcat 900  
 cagacacga gctctggaat ttgaggacgg catgtacgcc aatctgggca taggcacccc 960  
 ctgcttgccc agcaactca tcagtccag catgactgtc catcttcaca gtgagaacgg 1020  
 gatccctggc ctgggcccgt tccccacga agatgaggt gatgcgacc tcataaatgc 1080  
 aggcagcag acggtcacgg tgcctcccg ggcgtgctt ttcccgagc acgactcctt 1140  
 gcacatgac cgagggggac acatccaaact aaccatgctt ggagccatgc aggtttccaa 1200  
 atacggcgac ctggcgaact gcatgatccc tggcaagaag gtgaaaggca tggcggtgct 1260  
 catggacttg gtgtccagtc agaagaccag agtggtggtc accatgcagc actgcacaaa 1320  
 ggacaacacc cccaagatca tggagaatg caccatgcc ctgacgggga agcgtgctt 1380  
 ggaaccgcatc atcaccgaga aggcctgtgt tgacgtgcac aggaagaaag agctgacgct 1440  
 gagggagctc tgggagggccc tgacgttgga caacatcaaa aagagcacgg ggtgtgcctt 1500  
 tgcgtgtgcc cgaacctca ggcacatgca gcaggtggca cctgacggg acctggatct 1560  
 gggcggggtg gtgcgctcct caggcgagat gccacgggt tccccaggg aatacatgtc 1620  
 cccagctctg ggagggggtt gctactggcc tctactttc ctccctaggt ggacagtgtc 1680  
 cctctagaga gctgcgactt taattaaaaa caacaggaaa acaaaaaaaa aaa 1733

<210> 19  
 <211> 1148  
 <212> DNA  
 <213> Homo sapiens

<220>  
 <221> misc\_feature  
 <223> Incyte ID No: 1274566CB1

<400> 19  
 caacatcccc actgaaatag aatgtcttat gtctttgaat gccctaaaaa gatttaaaga 60  
 aataataact gatccttgag cacatatacc tacaggata tagcaalcat tgaactaat 120  
 aattagctaa ttaaaaaatt gttttgcaaa tgtagaatg actatttcac tctgtttgtg 180  
 gtgcttctgt aaccttgtgt tctgtcccc atgtggacag tgtgccacat ctagtattctg 240  
 catagatttc aagagagaca taaggacaag tttttlatgt gtgaggatgc agcttagggc 300  
 agcaaacattg cacaccaatt ataaaccaat aaagtctctg tccctgctc tccctcagcg 360  
 tctccctcac cagccagtat cagcagatgg tctatctcat tctctatggg aaaacagaaa 420  
 ctgttcatct tatgcttggg aagcctctct ctcttaacct tgagatttcc cttcagagat 480  
 ggtcgtttcc tcttccctc ctgcctaata gaataagtta tctctctgc ctaatagaat 540  
 aagtgtgtct cttcgccata tagaataagt tgttctcttc tctttacct cttttgagac 600

WO 01/79291

PCT/US01/11861

```

ttagcctctt  taaagattcc  cttttgtttt  agttgtcttc  gatttttctt  tcaactggct  660
taattccctc  agagacataa  tttaaatcat  ttcttgata  ctgtctctcg  ctcagtttac  720
accttagctt  actccttact  gtcaatgaaa  atcttgggag  ggttgatata  gcttcttagc  780
tctcttgctt  catgattact  cctcagccct  cagcatttgc  atgccaaact  ctgaacttcc  840
atgtttctcg  catgttagta  aaactgttct  tgtttcaitt  taettaagtt  cgaagcactg  900
tattattgct  cccacagccc  tcagtccaaa  acttcaagtgc  attgtaattg  ttaaaatctt  960
catcacattg  tattttaatg  gtctgcgttc  atatgtttcc  tccgtagact  aagcttctag  1020
aaggcagaaa  aatggattta  tatagattgc  atatttcctt  tggagttaaa  tgtaggtcat  1080
gacacataaa  tatatggaaa  atacactatc  caataaatcc  acaataaata  ttgggtgaaa  1140
aaaaaaaaa  1148

```

```

<210> 20
<211> 1213
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1349442CB1

```

```

<400> 20
ggttgcgagg  caccaccag  catcatttcc  catgcgaggt  ggcaaatgca  acatgtcttc  60
cagtttgggg  tgttacttcc  tctgtggaag  tattacacta  gccctgggaa  atgcacagaa  120
attgccaaaa  ggtaaaaggc  caaacctcaa  agtccacatc  aataccacaa  gtgactccat  180
cctcttgaag  ttcttgctgc  caagtccaaa  tgtaaagctt  gaaggtcttc  tcctgggata  240
tggcagcaat  gtatcaccaa  accagtactt  cctcttccc  gctgaaggga  aattcacaga  300
agctatagtt  gatgcagagc  cgaatatctt  gatagttgtg  cgacctgctc  caccctcaag  360
tcaaaagaag  tcatgttcag  gtaaaactcg  ttctgcgaaa  cctctgcagc  tgggtggttg  420
caactctgac  ccgagctcag  ttcttctgtc  ctggggttcc  ctcatcaacc  cacacatga  480
ctggacattg  ccaagtact  gtcccaatga  cagattttat  acaattcgct  atcgagaaaa  540
ggataaagaa  aagaagtggg  tttttcaaat  ctgtccagcc  actgaaacaa  ttgtgaaaaa  600
cctaaagccc  aacacagttt  atgaatttgg  agtgaagac  aatgcggaag  gtggaaattg  660
gagtaagatt  ttcaatcaca  agactgttgt  tggaaagtaa  aaagtaaatg  ggaaaatcca  720
aagtaacctat  gaccaagacc  acacagtgcc  agcataatgt  ccaaggaac  taatcccaat  780
aacaatctac  aagcaagtga  ttcagaatgt  tactcacaag  gattcagcta  aatccccaga  840
aaaagctcca  ctgggaggag  tgatactagt  ccaccttatt  attccaggtc  ttaatgaaac  900
tactgtaaaa  ctctctgcat  ccctaattgt  tgagatttca  gatgcactca  agacacaaat  960
agctaagaat  gaaaccttgg  cattactctgc  cgaatctaaa  acaccagagg  ttgaaaaaat  1020
ctcagcacga  cccacaacag  tgactcctga  aacagttcca  agaagcacta  aaccoactac  1080
gtctagtgcg  ttagatgttt  cagaacaac  actggttctc  agcaaaagga  ccccggaac  1140
attgcnaact  attctaatac  ctccagttga  atlgccactg  agcactctag  gtaaaaaata  1200
ataaatactg  cag  1213

```

```

<210> 21
<211> 2298
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1400156CB1

```

```

<400> 21
ggtgctcggg  cggcgctcgg  ggagggagcc  agcagcctag  ggcttaggcc  cgggccaacca  60
tggcgctgcc  tccaggccca  gcgcgccctc  ggcacacact  gctgctctcg  ccagcccttc  120
tgagctcagg  ttggggggag  ttggagccac  aaatagatgg  tcagacctgg  gctgagcggg  180
caacttcggg  gaataaargc  caecgcttca  cctgccgggt  ggcaaggggg  cctggcaacc  240
ccagattggc  ctggtatctg  gatggacagc  tgcaggaggc  cagcacctca  agactgctga  300
gcgtggggag  ggaggccttc  tctggaggca  ccagcacctt  cactgtcact  gccatcgagg  360
cccagcatga  gctcaactgc  tctctgcagg  accccagaag  tggccgatca  gccaacgcct  420
ctgtcatcct  taatgtgcaa  ttcaagccag  agattgccca  agtcggcgcc  aagtaccagg  480
aagctcaggg  cccaggcctc  ctggttgtcc  tgtttgcctt  ggtgcgtgcc  aaccgcggcg  540
ccaatgtcac  ctggatcgac  caggatgggc  cagtgcactgt  caacacctct  gacttccctg  600

```

WO 01/79291

- PCT/US01/11861

```

tgctggatgc gcagaactac ccttggctca ccaaccacac ggtgcagctg cagctccgca 660
gcctgggaca caacctctcg gtgggtggcca ccaatgacgt ggggtgcacc agtgcgtcgc 720
ttccagcccc aggcacctcc cggcaccatc ctctgatatc aagtgcactcc aacaacctaa 780
aactcaacaa cgtggccctg ccacggggaga acatgtccct ccggtccaac ctccagctca 840
atgacctcac tcagattccc agagcagtgga aaccagcaga ccggcagatg gctcagaaca 900
acagctggcc agagcttctg gaccgggagc ccggcggcct cctcaccagc caaggtttca 960
tccgctcccc agtgcgtggc tatactctatc gagtgtccag cgtgagcagt gatgagctct 1020
ggctctgagc cagggcgagc acaggagtat tctcttgccc tctggacacc ctccattccc 1080
tccaaggcat cctctacctg gctaggtcac caacgtgaag aagttatgcc actgccactt 1140
ttgcttgccc tcttggctgg ggtgccctcc atgtcatgca cgtgatgcat ttcaactggc 1200
tgttaaccgc aggggcacag gtatcttttg caaggctacc agttggagct aagccctcca 1260
tgctgactca ggttgggccc tgcatgtgat gactgggccc ttccagaggg agctcttttg 1320
ccagggtgtg tcagatgtca tccagcatcc aagtgtggca tggcctgctg tataccccc 1380
cccagttact cacagcacct tgtacagtag gcatgggggc gtgctgtgtg gggggacagg 1440
gaggggcctg catggatttt cctccttccc atgctatgta gccttgttcc ctcaggtaaa 1500
atttaggacc ctgctagctg tgcagaaccc aattgccctt tgcacagaaa ccaacccctg 1560
accacgggtg accggccaag cacaaacgtc ctttttgctg cacacgtctc tgccttccac 1620
ttctctctct ctgtcccccac cctctcttgg gaattctagg ttacacgttg gacctctctc 1680
actacttccac tgggcaclag actttctctat tggcctgtgc catcgcccag tatlagcaca 1740
agttagggag gaagaggcag gcgatgagtc tagtagcacc caggacggct tbtagctatg 1800
catcattttc ctacggcgtt agcactttaa gcacatcccc tagggagggg ggtgagtgag 1860
gggcccagag cctcttttgt ggcttcccca cgtttggcct tctgggatc actgtgagt 1920
tcttgagctc tgggggttga tggtttttct ctacgatgt cctctccacc accgggacccc 1980
agccttgacc aaccatgggt tgcctcatca gcaggaaggt gccttctctg gaggatgttc 2040
gccacagcca cataattcaa cagtgtggaa gcttttaggg aacatggaga aagaaggaga 2100
ccacataccc caaagtgacc taagaacact ttaaaaagca acatgtaaat gattggaaat 2160
taatatagta cagaatatat ttttcccttg ttgagatctt cttttgtaat gtttttcatg 2220
tacttgctta gggcggtgct gagcacacag caagtttaat aaacttgact gaattcattt 2280
acaaaaaaaa aaaaaaaaaa

```

&lt;210&gt; 22

&lt;211&gt; 2079

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 1610347CB1

&lt;400&gt; 22

```

agtgcgttgt gtatgaacgc agcggcggac ctgtgagggg atccgacttg ccggcagaac 60
ttacgctgctg ggaaccccggt cactgttgtct gctgcggggg actgtgggct gtttagtgcc 120
atgcacccctt taacagtgctg cctccaagtg cagaggcttc tggggtgggg accattggcc 180
ttctgtgtctt ggctgtcgtc gaggatgtgc agggcacaca gcagctcttc tagtaacctg 240
tgtccacgtc cagagaggca ggaagatgga gctcggaagg atttcagctc caggctggct 300
gtggacccga cttttcaaca ttttttaaaa agtgcctcag cctctcagga gaagctgtct 360
tcagaagtgg aagaccaccc tccctatctc atgatggatg aactcttttg aaggcagaga 420
aaagtctacc tcgagacctg tggctgccag atgaatgtga atgacacaga gatagcctgg 480
tccatcttacc agaagagtgg ctacctgcgg accagtaacc tccaagaggc agatgtgatt 540
ctcctttgtca cgtgctctat cagggagaaag gctgagcaga ccactctgaa ccgtttacat 600
cagcttaaaag ccttgaagac nagggcgccc cgctcccggt ttctctctga gattggaaat 660
ctaggctgca tggctgagag gttgaaggag gagattctca acagagagaa aatggtagat 720
attttggctg gtctctgagc ctaccgggac cttcccccgc tgcctggctgt tgcagagtcg 780
ggccagccaag ctgccaaact gctgctctct ctggacgaga cctatgctga tgcctatgca 840
gtccagacaa ggcgcagtgcc cactgtctgct tttgtgtcaa tcatgctgag ctgtgacaa 900
atgtgtagct actgcattgt tcttttcaac cggggcaggg agaggagtcg gcttattgcc 960
tccattctag aggaagtga gaaactttct gaagctttct gaaagaagt gaaacttctt 1020
ggtcagaatg ttaatagttt tcgggacaat tcggaggtcc agttcaacag tgcagtgcct 1080
accaatctca gtcgtggctt taccaccaac tataaaacca agcaaggagg acttcgtttt 1140
gctcatcttc tggatcaggt ctccagagta gatcctgaaa tggaggtccc ttttacctct 1200
ccccaccca aggattttcc tgatgaggtt ctgcagctga ttcattgag agataacatc 1260
tgtaaacaga tccacctgcc agccacaggt ggaagcagcc gtgtgttgga ggccatcgcg 1320
aggggalatt caagagaagc tlatgtggag ttagtccacc atattagaga actatttcca 1380

```

WO 01/79291

PCT/US01/11861

```

gggtgtgagcc tcagcagcga tttcattgct ggctttttgt gtgagacgga ggaagatcac 1440
gtccagacag tctcttttgc cggggaagtt cagtacaaca tgggcttccct ctttgcctac 1500
agcatgagac agaagacacg ggcataatcat aggcctgaagg atgatgtccc ggaagaggta 1560
aaatcaaggc gtttggagga actcatcact atcttccgag agaagcaac aaagccaat 1620
cagacctctg tgggtgttac ccagtttggtg ctagtggaaag ggctcagtaa acgctctgcc 1680
actgacctgt gtggcaggaa tgatggaaac ctttaaggtga tcttccctga tcgagagatg 1740
gaggatgtca ataaccctgg gctcagggtc agagcccagc ctggggacta tgtgctggtg 1800
aagatcacctc cagccagttc tcagacactt aggggacatg ttctctgcag gaccactctg 1860
agggactctt ctgcataatt ctgacctgag aggatggcct cagagctgac ttgggcaatc 1920
ctcccaca ggaaggggag acattgcctg ccactgagga aacaggtcat gaaggtggag 1980
ataagctgca agggcggaag caactttatg tcagtggaaa acgtgtctct ttaaagctgc 2040
tatgtgaaca gcttttacag tcattaaatt tacctaaac 2079

```

```

<210> 23
<211> 846
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 187209CB1

```

```

<400> 23
tggctcgtaag gaaaccgagt gagaaggga tgcaacagaa gaaaaagacc aaagacctgg 60
gtttcagggc tgggaaagaa agcaagacag aatggaggaa atgaggcctg caggacatgg 120
cgtctcaaat gtttgcgtcg cctttaaagt agcctgtcac agctgccttc cagactcttt 180
caatgccctc atcccttctc cagatagaaa tggagcagct cttcttggag gccaggtctc 240
agctgatagc aagtctgagg ccaggaggaa ccagtgtgac tccatgctgc tcagaaacca 300
acagctgtgc tcccatgtc aagaatgaa atgtgtacaa ccagaacaa tgaatatccc 360
agatgatcca aaagcatcct ttgagaattg tatgagttat agaagagtc ttcatcaacc 420
caatttccag actacacctg agcctttcca tgatgacatc ccaacagaaa acattacta 480
cagactgcc atctctgggc ccaggacagc tgtcttccac ggattactga caggagccta 540
caaaactcta aaagagagac aacgtctctc cltgcccaga aaggaaccaa taggcaagac 600
aacggggcag tgagcggtag gagctcatca cctcccagac tcccagagag aaaaataacc 660
cgccaagcca atctttgaca ctggcacctt ctctccaca ttttctctct tctcccaaaa 720
gatgatttaa ttttgccttc ctaagattgc tggattctca gctcttacct ctatgtttct 780
tctcacgtat cctaaagaca aaattgttta atttacctga ttataaagat ctgtttatga 840
aaatgg 846

```

```

<210> 24
<211> 1148
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2607963CB1

```

```

<400> 24
caggatgcag ggcgcgtgg caggagagct cgctcctctg ggctgtctcc tggctgtctc 60
tcactccca ggcctctttg cccggagcat cgggtgtgtg gaggagaaag ttccccaaaa 120
cttcgggaac aacttgcttc agctcggaca accttccctc actggccctc ctaactctga 180
acatccgacg ccgctctctg accctaggct caatgacttg gcaaggttbc ctctgaagct 240
cagcgtgcct ccactgagtg gcttcccaac tgacggagtg tctgcagtg agaggtggcc 300
tccatcgtgg ggcgtgctg ccatggattc ctggccctct gaggatcctt gccagatgat 360
ggctgctggc gctgaggacc gcctggggga agcgtgcct gaagaactct ctacctctc 420
cagtgctggc gccctgctc cgggcagtg ccctttgctt gggagttct ctcccgatgc 480
cacaggccct tcactgagg cttaactcct ccaccaggac tcggagtcca gacgactgac 540
cgttctaat tcactggag ccgggggaaa aatcctttcc caacgcctc cctggctctc 600
catccacagg gttctgctg atcaccctg ggtaccctg aatcccagtg tgcctctggg 660
aggtggagcc cctgggactg gttggggaac gaggccatg ccacaccctg agggaatctg 720
gggtatcaat aatcaacccc caggtaccag ctggggaaat attaatcgtt atccaggagg 780
cagctgggga aatattaatc ggtatccagg aggcagctgg ggggaatatta atcggtatcc 840

```

WO 01/79291

PCT/US01/11861

```

aggaggcagc tgggggaata ttcactata cccaggtatc aataacccat ttctctctgg 900
agtctccgc cctctgggt cttcttgga catccagct ggttcccta atctcccaag 960
cctcaggttg cagtggggt agagcagat agagggaaac ccaacattgg gaggtagagt 1020
cttgcctccc ccccttgctg tgggggtca atccaggccc tgttaacatg ttccagcac 1080
tatccccact ttctagtgcc tccctgctc atctccaata aaataaaagc acctatggaa 1140
aaaaaaa 1148

<210> 25
<211> 3076
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 412044CB1

<400> 25
gtgacactga gcggcgagc ggggcccagc cggagaccgt gccggagttc gggagcggca 60
acagagtggg catagacact ccgagcagcc tcgccgtcgt ctctgcgttc ctgttgactg 120
cctggctggc cctccccccta ctccctcggtt cctgggtgaag aggcctgcgc ctgctgtttg 180
gggaggggggt ggtgaggacc ggtcctctgt tcgcagtggt ctgctgtcgg ggggctgcct 240
gttcgcggag gtcgggagag actccttggg ggtcgagcac ataacggggt tcgggtgtct 300
cgtgtgtgaa catcacaggg tttgtgagt cacttagatg ttgcaatga gcaactgtgc 360
tggcatgccc cagtgttttg gataccaatg cataggactc catagtaac gaattacca 420
gagcgcaacg tcactgagcat agtgatccca ttgggggttg atacagcaga gacgtcatac 480
ttggaatagg ctgcaggttc agaaccagaa tccgtagaag ctgcccctgt ggtagttgag 540
aaatccaaaca gttatcccca ccagttatat accagcagct cacatcattc acacagttac 600
atgggtttgc cctatgcgga ccataattat ggtgctcgtc ctctccgac acctccggt 660
tccctctctc catcagtcct tatagcaaaa aatgaaglag gcatactac cactccaat 720
ttgatgaaa ctccagtgcc tactacaatc agcaactctg aggtaggag ttatgtgact 780
gatgtaacca ggtgcataat tgggtttaca catgatgat gatacatgat ctgttgtgac 840
aaatgcagcg tttggcaaca tatgactgc atggggattg ataggcagca tatcctgat 900
acatatctat gtgaacgttg tcagcctagg aatttgata aagagagggc agtgctacta 960
caacgcggga aaagggaaaa tatgtcagat ggtgatacca gtgcaactga gattggtgat 1020
gaggttcttg tggaaattata tactgcattt cagcatactc caacatcaat tactttaact 1080
gcttcaagag ttcccaaatg taatgataaa agaagggaaa aaagcgggga gaaagacaaa 1140
cacatttcaa aatgtaaaaa ggcatttctg gaaggatcta ggaagtcac aagagtttaag 1200
ggttcagctc cagagattga tccttctctc gatggttcaa attttgatg ggagacaaag 1260
atcaaaagcat ggalggatcg atatgaagaa gcaataaaca accagtacag tgagggtggt 1320
cagagggagg cacaagaat agctctgaga ttaggcaatg gaaatgacaa aaaagagatg 1380
aataaaatccg attlgaatac caacactctg ctcttcaaac ctctgtaga gagcaata 1440
caaaagata agaaattct taatctgca aagatttgc ctctgtatgc acttatcatt 1500
gaatacagag ggaagtattat gctgagagaa cagtttgaag caaatgggta ttctttaa 1560
agaacatacc cttttgtgtt attctactc aaatttcatg ggctagaaat gtgtgttgat 1620
gcaaggactt ttgggaatga ggtcgtatc atcagcggtt ctgtacacc caatgcagag 1680
gtgagcgatg aaattcaaga tggaaaccata catctttata ttattctat acacagtatt 1740
ccaaagggaa ctgaaattac tattgccttt gattttgact atggaaattg taagtacaa 1800
gtgactgtg catgcctcaa agaaaaccca gactgcccgt ttctaaacg tagttctgaa 1860
tcctgggaaa atatcaatag tggttatgag accagacgga aaaaagaaaa aaaagacaaa 1920
gatatttcaa aagaaaaaga tacacaaaat cagaatatta ctttggtatg tgaaggaaac 1980
accacaaaaa tgaagagccc agaaactaaa caaagaaagc ttctccact gagactatca 2040
gtatcaata atcagggaacc agattttatt gatgatag agaaaaaac tcctattagt 2100
aatgaagtag aatggaatc agaggagcag attgcagaaa ggaagggaaa gatgacaa 2160
gaagaagaaa aatggaagc aattttgcaa gcttttgcaa gacttgaaa gagagaaa 2220
agaagagaa acgttttggg aaggatcagc acagccaaa ctgaagttaa aactgaatg 2280
aagatacac agattgtcag tgatgtgaa gttattcagg aacaagcaaa agaagaaaat 2340
gctagcaagc caaccctgc caaagtaaat agaactaaac agagaaaaag ttttctcgg 2400
agtaggactc acattggaca gcaagcgtcg agacacagaa ctgtcagcat gtgttcagat 2460
atccagccat ctctcctga tatagaagtt acttcacaa aaaaatgat tgaaaatact 2520
gtacttacaa tagaaccaga aactgaaact gcaactagca aaataattac tgaactgaa 2580
ttccagcac ttaataaatg tcctaccaag taccocaaaa caaagaaggt atgattctaa 2640
tgaatgtaag aactgttttt ctaacagttt cttatattaa ttatattgtt gttttaaaa 2700
ttggttttt aagacctcat aataataaga ggcagttttt atactgcag attttaaaac 2760

```

WO 01/79291

- PCT/US01/11861

```

taagaatgag aattccaaaa ctgtaaaatt aatataaatg ttgtcattac tgtgaagata 2820
aagttacatt cagtttatca ggaacttttat ggtattgcaa ttcattgattt ctttaataaa 2880
gtttgtctac ttatgttaca aaatatatac ttctctgaaa ctgggttttag atgtgtctac 2940
ctttatattt ttataaattt cattgtatag gtatattata agaattaaat gtgaagaaat 3000
tgatttccac agaattgtact atgaaaaatt gtaagaaaga gtatgttttag gtgtaatat 3060
taataaaaat ctctgc 3076

```

&lt;210&gt; 26

&lt;211&gt; 2102

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 638118CB1

&lt;400&gt; 26

```

gttctctcca ttatatatgt gtgtgcactg tctggtcagg agcagtgctc atgtgcgttt 60
ctgtggccgc agtgcaactg ggtgtgggtgg ggggtgcgta cgtgtgggct ttaggtttaa 120
gatgtctctg cgtgtgtgtt gtggggctgg aggccgagcc aggcacggca gctggccctc 180
tccaggcgca tggagggaatt ctccctttgc cgttccaaag atgcccgccg ctccctccgg 240
ggctcccggg ggggtcagcc aacctacaa agggcagcga gtctcaggg cgcccgagg 300
ccaccagctc gctctcgggc tctactcaga ctacgcgcaa cagtctccg ggacagactc 360
gcattcagcg cctcggggct cggcctcctc ccataggctc ctcttctct ctttcatcca 420
agtcttcaaa ctttgaaaaa aatatgagct ttgtccaaat aagacgggat agtttatgga 480
gtttttcag gcgcattgaa ctctggtcag cccctcagac gtgggctcgg ccatcagcag 540
gcccgggtgg ttggccgggc tggccctccc aggcctgctt tctctctggt cgccggcctg 600
cccgcccggg cctccaccgc cgcattcat gctgggtccg cgcccgagc cccactcctg 660
gtatgcctga gatctccag gggccaaagc ggtcctctt cccggccac tcttccagg 720
ccaggacgct cccgaggagc cagcccatcg tggctggggg tcccaggggc tggaggcctg 780
ctgggttagc ctctgtttcc cagacattga ctcgaggcgc ctccagctc ctcaaccccc 840
tgaactttcaa gtcttttcaa gtctctgtgc cctgaccaag ttatgaggaa aagaggcgct 900
agggcctcga ggtgggctcg caccctctcg aaccaccta gatgtctct cctccccagc 960
ctatgagttc agacttgggg ggtcccatct gtccctcctt gaggctcctc atctgttccc 1020
tgctctgatt ctgtcagcag atcctgaaca ctcttgggt gggaccatg cctccaagcc 1080
tgccccaat gccagctgtc ctggggaac tcccagaac tctcgggaa aggcccgaga 1140
accttcggag gaagcctcgg ggaactcctg ccacctgtc tgtttctct cggcccccac 1200
tgcttcagg tatcaggaa agggcaggcc cctgtgccc ctctccctt cccaggcgcg 1260
ccaacaatac tccaccatgg ggcgcttctt ttcttctct gaaactaaga cattggactg 1320
aaggaaactgg cctgaggggg gctgaccggg ggcctgtgct gctgggagct ctacgcacaa 1380
gaagcctcgg gggcacgga caggaaactc agcccaagg gctggcctt ctcttcgca 1440
ggagttcccc aaatccaca cagcagctcg agcagcctt cactcagcca gaggagggtc 1500
gggctcccc cctgtctct ggtgtgggg tctggcctt cctgcccgtt ccaaggcctc 1560
cggaagcccc aaccagttc tcagttacct cctctggcgg gaaggcctca cgttgtcttc 1620
cagagctgct ctgggccag ggcctgctcc gtgaccacct gatggacgtg ctgggcagca 1680
tgggctccca aggcagcctc cccagctgct ctcccacccc accccagctg cccggcggtc 1740
gggcccacga ggcctctggg gacacttcca taggaagggg gcctggcact ttgtgattgc 1800
cacgtgtttc ctgttaagcc gcttgcccc agtgcaaatc tctgtgtttt gctctctctc 1860
gaacaaaaat gtaaacggac gccggaagac aaatggatc gaataccttt ttgcccata 1920
aggggtttta caaagaacgt gtctcactag gaactaggac actacatcgg ccgagaccaa 1980
ctctgggata aaaggaataa ggaagaatct gttgtgaac agttacagga ttgtttcttt 2040
cctggggttt aaacttttgt taaattgtga aattatggag gagttttata gaaaaaaaaa 2100
aa 2102

```

&lt;210&gt; 27

&lt;211&gt; 807

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; Incyte ID No: 743323CB1

WO 01/79291

PCT/US01/11861

```

<400> 27
gtctgtgtgcc agggccaggg agggggggcac tggctgcttc tgtattttgg gggttggggc 60
cctggagctt cccatgcgga attgcccgtcc ctcctcctag gcgagtccca gggccacccc 120
atccccacag gaccggggcg ccagcttctg aaagcatggg scatctgcgg aagaactggg 180
ttgtttccca gctttgtgcc ctggggaggg gcgatacggc cctctcatgt cagcagtggt 240
tgggtgtcca catgcttgct agccccacgc tgtgtcctcg cgtctcttcc cgtctcatcc 300
atctggatgc ttgacacctc tgacagcatc ccttctctgt catcttaggg cagcttcagg 360
aaaccgaaaa acagggttgt gtcccttccat taaccccttt atccacaagt tcagttacag 420
catgagccct ggggagctcc aagcgtgcag ccaggagccc cgtagccagg gatggtcctg 480
gctgtgtctg tgcaccaggg ccgcttcccc cactttttcc agaggaaact gttctacggc 540
cagaagaaca agtaccgagc accccgaggg aagccggccc cggcctcagg ggacaccagg 600
acccttgcaa aggggtccag tgtccgggag cctggggcca gtgggttga ggggcccacat 660
tccagctgag tggccttgct ctgtgtgagc cccgtgcgag ggcctgctt gtactgtggc 720
cctggaacct tctgtagcta agaggggaatc ctggcccccct ccccgagaagc catttgtcaa 780
taaaccattt ctaagaaaaa aaaaaaaa 807

<210> 28
<211> 1049
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1691509CB1

<400> 28
gtttaaccat attgcccagg ctggtcttga actcatggcc tcaagtatc ttctgcttc 60
atgtctcccaa agtgttggga ttataggcat gagccactgt gccagctct agtgtaccat 120
tttaattctc tctttttttt ttaaacatac taacacccaa cggatctct tgtttcttt 180
gctttattct tttagattat ttcttggtat gttacctttc tcccccag ttcaacccct 240
cagaataccc aagctaaagg atattaatgc tggatctttt tgtaagaca agaacccttt 300
catagtgtga ataaccgccc cctgtcatat ttgtcatagt ctgaaatcag attgtccat 360
gttctgtttt ctaggaaagt ggttcctgtt ttaacgtagt acatgtcagg taaaaggtaa 420
gccagagata gccattcagg agagaactgc cagaaatgaa acgcttcctg gtgaagggca 480
tggggtttgg gtatgtacag tgccatgatg tttctttttc agctcattct tggaaatcca 540
gaacaggcct taagtcttct ccacatggcc atcgagccca tcttggtga cggggctatc 600
ctggacaaag gtctgtccat gttcttagtg gccaaagtgc agtggtcttc agcagcttcc 660
tacgatcagc cgaagaaagc agaagctctg gaggtgccta tcgagaacct caatgaagcc 720
aagaactact ttgcaaaagt tgactgcaaa gagcgcatca gggacgtcgt ttacttcagg 780
gccagactct accataccct ggggaagacc caggagagga accggtgtgc gatgtcttc 840
cggcagctgc atcaggagct gcctctctcat ggggtacct tgataacca tctctagaga 900
ggacatccct gctgggtgc tgtgcagagt ataaagattt ggaactgttc atgtccctcc 960
tctccctata aatgatgtat ttgtgacacc ctatcttctg aataaacagc attctgatta 1020
gtttgtctta aaaaaaaaaa aaaaaaaaaa 1049

```

## 【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  
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number  
WO 01/079291 A3(51) International Patent Classification: C12N 15/12,  
C07K 14/47, 16/18, C12Q 1/68, A61K 38/17, 39/395,  
G01N 33/53, 33/68

(21) International Application Number: PCT/US01/11861

(22) International Filing Date: 11 April 2001 (11.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/197,854 14 April 2000 (14.04.2000) US  
60/202,373 4 May 2000 (04.05.2000) US  
60/205,899 18 May 2000 (18.05.2000) US  
60/210,155 1 June 2000 (01.06.2000) US  
60/209,401 1 June 2000 (01.06.2000) US(63) Related by continuation (CON) or continuation-in-part  
(CIP) to earlier applications:US 60/197,854 (CIP)  
Filed on 14 April 2000 (14.04.2000)  
US 60/202,373 (CIP)  
Filed on 4 May 2000 (04.05.2000)  
US 60/205,899 (CIP)  
Filed on 18 May 2000 (18.05.2000)  
US 60/210,155 (CIP)  
Filed on 1 June 2000 (01.06.2000)  
US 60/209,401 (CIP)  
Filed on 1 June 2000 (01.06.2000)(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): GRIFFIN, Jennifer,  
A. [US/US]; 181 Irene Court #11, Belmont, CA 94002  
(US); YAO, Monique, G. [US/US]; 111 Frederick Court,  
Mountain View, CA 94043 (US); BRUNS, Christopher,  
M. [US/US]; 375 S. Rengstorff Avenue #126, Mountain  
View, CA 94040 (US); YUE, Henry [US/US]; 826 Lois  
Avenue, Sunnyvale, CA 94087 (US); DELEGEANE,  
Angelo, M. [US/US]; 594 Angus Drive, Milpitas, CA  
95035 (US); HAFALLA, April [US/US]; 2227 Calle de  
Primavera, Santa Clara, CA 95054 (US); PATTERSON,Chandra [US/US]; 490 Sherwood Way #1, Menlo Park,  
CA 94025 (US); POLICKY, Jennifer, L. [US/US]; 1511  
Jarvis Court, San Jose, CA 95118 (US); TRIBOULEY,  
Catherine, M. [US/US]; 1121 Tennessee Street #5, San  
Francisco, CA 94107 (US); BAUGHN, Mariah, R.  
[US/US]; 14244 Santiago Road, San Leandro, CA 94577  
(US); NGUYEN, Daniel, B. [US/US]; 1403 Ridgewood  
Drive, San Jose, CA 95118 (US); LAL, Preed [IN/US];  
P.O. Box 5142, Santa Clara, CA 95056 (US); TANG,  
V. Tom [US/US]; 4330 Ranwick Court, San Jose, CA  
95118 (US); HILLMAN, Jennifer, L. [US/US]; 230  
Monroe Drive #17, Mountain View, CA 94040 (US); LU,  
Dyung, Aina, M. [US/US]; 233 Coy Drive, San Jose, CA  
95123 (US); BATRA, Sajeev [US/US]; 555 El Camino  
Real #709, Sunnyvale, CA 94087 (US); AU-YOUNG,  
Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA  
94005 (US); REDDY, Roopa [IN/US]; 1233 W. McKinley  
Drive #3, Sunnyvale, CA 94086 (US); AZIMZAI, Valda  
[US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA  
94552 (US).(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,  
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,  
DE, DK, DM, DZ, EC, ES, FI, GB, GD, GH, GR, GM, HK,  
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,  
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,  
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,  
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.(84) Designated States (regional): ARIPO patent (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, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,  
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:  
1 May 2003For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

(54) Title: SECRETED PROTEINS

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



WO 01/079291 A3

## 【国際調査報告】

## INTERNATIONAL SEARCH REPORT

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                                                                                                                                       |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|
| International Application No.<br>PCT/US 01/11861                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                       |
| <b>A. CLASSIFICATION OF SUBJECT MATTER</b><br>IPC 7 C12N15/12 C07K14/47 C07K16/18 C12Q1/68 A61K38/17<br>A61K39/395 G01N33/53 G01N33/68                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                       |
| According to International Patent Classification (IPC) or to both national classification and IPC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                                                                                                       |
| <b>B. FIELDS SEARCHED</b><br>Minimum documentation searched (classification system followed by classification symbols)<br>IPC 7 C07K C12N C12Q A61K G01N                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                                                                                                                                       |
| 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)<br>EPO-Internal, BIOSIS, MEDLINE, WPI Data, PAJ                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                                                                                                       |
| <b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                                                                                                       |
| Category *                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | Citation of document, with indication, where appropriate, of the relevant passages<br>Relevant to claim No.                           |
| X                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | WO 00 05361 A (SMITHKLINE BEECHAM PLC)<br>3 February 2000 (2000-02-03)<br><br>SEQ ID NOS:1-3<br>page 9-17; claims 1-18<br>---<br>-/-- |
| <input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |                                                                                                                                       |
| * Special categories of cited documents :<br>"A" document defining the general state of the art which is not considered to be of particular relevance<br>"E" earlier document but published on or after the international filing date<br>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)<br>"O" document referring to an oral disclosure, use, exhibition or other means<br>"P" document published prior to the international filing date but later than the priority date claimed<br>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention<br>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone<br>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.<br>"Z" document member of the same patent family |                                                                                                                                       |
| Date of the actual completion of the international search<br>23 August 2002                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Date of mailing of the international search report<br>05.12.2002                                                                      |
| Name and mailing address of the ISA<br>European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 51 651 epo nl<br>Fax: (+31-70) 340-3016                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Authorized officer<br>Brouns, G                                                                                                       |

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

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 01/11861

| C/(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT |                                                                                                                                                                                                                                                                                                                                           |                            |
|------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------|
| Category *                                           | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                                                                                                                                        | Relevant to claim No.      |
| A                                                    | HIESBERGER T ET AL: "Cellular uptake of saposin (SAP) precursor and lysosomal delivery by the low density lipoprotein receptor-related protein (LRP)." THE EMBO JOURNAL. ENGLAND 17 AUG 1998, vol. 17, no. 16, 17 August 1998 (1998-08-17), pages 4617-4625, XP002210336 ISSN: 0261-4189 cited in the application abstract; figures 1,6,8 | 1-7,<br>10-12,<br>43,45,59 |
| A                                                    | BLOBEL G ET AL: "TRANSLOCATION OF PROTEINS ACROSS MEMBRANES: THE SIGNAL HYPOTHESIS AND BEYOND" SYMPOSIA OF THE SOCIETY FOR EXPERIMENTAL BIOLOGY, COMPANY OF BIOLOGISTS, CAMBRIDGE, GB, 1979, pages 9-36, XP001010405 ISSN: 0081-1386 table 1                                                                                              | 1                          |

Form PCT/ISA/210 (continuation of second sheet) July 1999

| INTERNATIONAL SEARCH REPORT                                                                                                                                                                                                                                                                                                   | International application No.<br>PCT/US 01/11861 |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------|
| <b>Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)</b>                                                                                                                                                                                                                |                                                  |
| This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:                                                                                                                                                                                      |                                                  |
| 1. <input checked="" type="checkbox"/> Claims Nos.:<br>because they relate to subject matter not required to be searched by this Authority, namely:<br>see FURTHER INFORMATION sheet PCT/ISA/210                                                                                                                              |                                                  |
| 2. <input checked="" type="checkbox"/> Claims Nos.: 20,21,23,24<br>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:<br>see FURTHER INFORMATION sheet PCT/ISA/210 |                                                  |
| 3. <input type="checkbox"/> Claims Nos.:<br>because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).                                                                                                                                                           |                                                  |
| <b>Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)</b>                                                                                                                                                                                                                        |                                                  |
| This International Searching Authority found multiple inventions in this international application, as follows:                                                                                                                                                                                                               |                                                  |
| see additional sheet                                                                                                                                                                                                                                                                                                          |                                                  |
| 1. <input type="checkbox"/> As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.                                                                                                                                                          |                                                  |
| 2. <input type="checkbox"/> As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.                                                                                                                                              |                                                  |
| 3. <input type="checkbox"/> 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. <input checked="" type="checkbox"/> 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.:<br>1-44 partially 45 and 59 complete                                  |                                                  |
| Remark on Protest                                                                                                                                                                                                                                                                                                             |                                                  |
| <input type="checkbox"/> The additional search fees were accompanied by the applicant's protest.                                                                                                                                                                                                                              |                                                  |
| <input type="checkbox"/> No protest accompanied the payment of additional search fees.                                                                                                                                                                                                                                        |                                                  |

International Application No. PCT/US 01/11861

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

## Continuation of Box I.1

Although claims 18, 21 and 24 are directed to a method of treatment of the human or animal body, the search has been carried out and based on the alleged effects of the composition.  
Although claims 32 and 34 are directed to a diagnostic method practised on the human or animal body, the search has been carried out and based on the alleged effects of the composition.

## Continuation of Box I.2

Claims Nos.: 20,21,23,24

Present claims 20, 21, 23 and 24 relate to compounds defined by reference to a desirable characteristic or property, namely that they are effective as an agonist or antagonist of a polypeptide of claim 1.

The claims cover all compounds having this characteristic or property, whereas the application provides no support within the meaning of Article 6 PCT or disclosure within the meaning of Article 5 PCT for such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search for the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compounds by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, no search has been carried out for the claims which relate to said compounds.

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 Application No. PCT/US 01/11861

## FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: in part 1-44, all as far as applicable,  
45 and 59 complete

## Invention 1:

Polypeptide comprising an amino acid sequence of SEQ ID NO:1, a naturally occurring sequence at least 90% identical to said sequence, a biologically active fragment or a immunogenic fragment having said sequence; an isolated polynucleotide encoding aforementioned polypeptides, possibly comprising the sequence of SEQ ID NO:15 or at least 60 contiguous nucleotides, a recombinant polynucleotide comprising a promoter operably linked to said polynucleotide and a cell or transgenic organism comprising said recombinant polynucleotide; a method of producing the aforementioned polypeptide; an isolated antibody binding said polypeptide and compositions thereof; a method of detecting a target polynucleotide in a sample; composition comprising said polypeptides; methods of treating a disease using said polypeptides and agonists or antagonists of said polypeptides; methods of screening for agonists, antagonists, binding compounds, modulating compounds, compounds that alter expression, of assessing toxicity, a diagnostic test, methods of diagnosis, method of preparing antibodies, method of detecting said polypeptides and a method of purifying said polypeptides

2. Claims: in part 1-44, 46-58, 60-72,  
all as far as applicable

## Inventions 2-14:

as invention 1, but limited to the subject-matter relating to SEQ ID NOS 2-14 and 16-28, wherein invention 2 is limited to SEQ ID NOS 2 and 16, invention 3 is limited to SEQ ID NOS 3 and 17, etc...., invention 14 is limited to SEQ ID NOS 14 and 28.

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/US 01/11861

| Patent document<br>cited in search report |   | Publication<br>date | Patent family<br>member(s) | Publication<br>date |
|-------------------------------------------|---|---------------------|----------------------------|---------------------|
| WO 0005361                                | A | 03-02-2000          | WO 0005361 A1              | 03-02-2000          |
|                                           |   |                     | EP 1098970 A1              | 16-05-2001          |
|                                           |   |                     | JP 2002524025 T            | 06-08-2002          |
| -----                                     |   |                     |                            |                     |

## フロントページの続き

| (51)Int.Cl. <sup>7</sup> | F I            | テーマコード (参考) |
|--------------------------|----------------|-------------|
| A 6 1 P 9/00             | A 6 1 P 9/00   | 4 C 0 8 4   |
| A 6 1 P 25/00            | A 6 1 P 25/00  | 4 C 0 8 5   |
| A 6 1 P 35/00            | A 6 1 P 35/00  | 4 H 0 4 5   |
| A 6 1 P 37/02            | A 6 1 P 37/02  |             |
| C 0 7 K 1/22             | C 0 7 K 1/22   |             |
| C 0 7 K 14/47            | C 0 7 K 14/47  |             |
| C 0 7 K 16/18            | C 0 7 K 16/18  |             |
| C 0 7 K 16/46            | C 0 7 K 16/46  |             |
| C 0 7 K 19/00            | C 0 7 K 19/00  |             |
| C 1 2 N 1/15             | C 1 2 N 1/15   |             |
| C 1 2 N 1/19             | C 1 2 N 1/19   |             |
| C 1 2 N 1/21             | C 1 2 N 1/21   |             |
| C 1 2 N 5/10             | C 1 2 P 21/02  | C           |
| C 1 2 P 21/02            | C 1 2 P 21/08  |             |
| C 1 2 P 21/08            | C 1 2 Q 1/02   |             |
| C 1 2 Q 1/02             | C 1 2 Q 1/68   | A           |
| C 1 2 Q 1/68             | G 0 1 N 33/15  | Z           |
| G 0 1 N 33/15            | G 0 1 N 33/50  | Z           |
| G 0 1 N 33/50            | G 0 1 N 33/53  | D           |
| G 0 1 N 33/53            | G 0 1 N 33/53  | M           |
| G 0 1 N 33/566           | G 0 1 N 33/566 |             |
|                          | C 1 2 N 5/00   | A           |
|                          | A 6 1 K 37/02  |             |

- (31)優先権主張番号 60/210,155  
 (32)優先日 平成12年6月1日(2000.6.1)  
 (33)優先権主張国 米国(US)  
 (31)優先権主張番号 60/209,401  
 (32)優先日 平成12年6月1日(2000.6.1)  
 (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,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,S G,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

- (72)発明者 ヤオ、モニック・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 ブランス、クリストファー・エム  
 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・# 1 2 6・サウスレンジスト  
 フアベニュー 5 7 5  
 (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 デレジーン、アンジェロ・エム  
 アメリカ合衆国カリフォルニア州9 5 0 3 5・ミルピタス・アングスドライブ 5 9 4  
 (72)発明者 ハファリア、エープリル

- 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 ・ コーレデブリマベール 2 2 2 7  
 (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  
 (72)発明者 ポリッキー、ジェニファー・エル  
 アメリカ合衆国カリフォルニア州 9 5 1 1 8 ・ サンノゼ ・ ジャービスコート 1 5 1 1  
 (72)発明者 トリボレー、キャサリン・エム  
 アメリカ合衆国カリフォルニア州 9 4 1 0 7 ・ サンフランシスコ ・ # 5 ・ テネシーストリート 1  
 1 2 1  
 (72)発明者 ボーグン、マライア・アール  
 アメリカ合衆国カリフォルニア州 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 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 ラル、プリーティ  
 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 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 ヒルマン、ジェニファー・エル  
 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 ・ # 1 7 ・ モンロードドライブ 2  
 3 0  
 (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 バトラ、サジীব  
 アメリカ合衆国カリフォルニア州 9 4 0 8 7 ・ サニーバイル ・ # 7 0 9 ・ エルカミーノリアル 5  
 5 5  
 (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 レディ、ルーパ  
 アメリカ合衆国カリフォルニア州 9 4 0 8 6 ・ サニーバイル ・ # 3 ・ ウェストマッキンレードライ  
 ブ 1 2 3 3  
 (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

F ターム(参考) 2G045 AA25 AA40 BB03 BB20 CB01 CB21 DA12 DA13 DA14 DA36  
 DA37 DA77 FB02 FB03  
 4B024 AA01 AA11 BA44 CA04 CA09 CA20 GA11 HA12 HA15  
 4B063 QA18 QA19 QQ02 QQ20 QQ42 QQ52 QQ79 QR55 QR59 QR62  
 QR69 QR80 QR82 QS12 QS24 QS25 QS28 QS34  
 4B064 AG01 AG27 CA10 CA19 CA20 CC24 DA01 DA13  
 4B065 AA93Y AB01 AC14 BA02 CA24 CA44 CA46  
 4C084 AA02 AA06 AA07 AA13 AA14 AA17 BA01 BA08 BA22 BA23  
 BA35 BA44 CA18 CA53 CA56 CA59 NA14 ZA012 ZA362 ZB072  
 ZB112 ZB262 ZC512  
 4C085 AA13 AA14 AA16 AA19 BB11 BB31 CC02 CC03 CC04 CC05  
 CC13 CC21 CC22 CC23 CC32 EE01 EE06 GG01 HH01 JJ01  
 KA04 KA05 KA27 KB82 LL20  
 4H045 AA10 AA11 AA20 AA30 BA10 CA40 DA76 DA86 EA21 EA22  
 EA23 EA27 EA28 EA50 EA51 FA72 FA74

|                |                                                                                                                                                                                                                                                                                                   |         |            |
|----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|------------|
| 專利名称(译)        | 分泌蛋白質                                                                                                                                                                                                                                                                                             |         |            |
| 公開(公告)号        | <a href="#">JP2004518402A</a>                                                                                                                                                                                                                                                                     | 公開(公告)日 | 2004-06-24 |
| 申請号            | JP2001576886                                                                                                                                                                                                                                                                                      | 申請日     | 2001-04-11 |
| [標]申請(專利权)人(译) | 洞察Genomics公司                                                                                                                                                                                                                                                                                      |         |            |
| 申請(專利权)人(译)    | 洞察基因组公司                                                                                                                                                                                                                                                                                           |         |            |
| [標]發明人         | グリフィンジェニファーエイ<br>ヤオモニークジー<br>ブランスクリストファーエム<br>ユエヘンリー<br>デレジーンアンジェロエム<br>ハファリアエープリル<br>パターソンチャンドラ<br>ポリッキージェニファーエル<br>トリボレーキャサリーンエム<br>ボーグンマライアアール<br>ニュエンダニエルビー<br>ラルプリーティ<br>タングワイトム<br>ヒルマンジェニファーエル<br>リュデュングアイナエム<br>バトラサジープ<br>オウヤングジャンス<br>レディルーパ<br>アジムザイヤルダ                                |         |            |
| 發明人            | グリフィン、ジェニファー・エイ<br>ヤオ、モニーク・ジー<br>ブランス、クリストファー・エム<br>ユエ、ヘンリー<br>デレジーン、アンジェロ・エム<br>ハファリア、エープリル<br>パターソン、チャンドラ<br>ポリッキー、ジェニファー・エル<br>トリボレー、キャサリーン・エム<br>ボーグン、マライア・アール<br>ニュエン、ダニエル・ビー<br>ラル、プリーティ<br>タング、ワイトム<br>ヒルマン、ジェニファー・エル<br>リュ、デュング・アイナ・エム<br>バトラ、サジープ<br>オウ・ヤング、ジャンス<br>レディ、ルーパ<br>アジムザイ、ヤルダ |         |            |
| IPC分类号         | G01N33/50 A61K38/00 A61K39/395 A61K45/00 A61K49/00 A61P9/00 A61P25/00 A61P35/00 A61P37/02 C07K1/22 C07K14/47 C07K16/18 C07K16/46 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/12 C12P21/02 C12P21/08 C12Q1/02 C12Q1/68 G01N33/15 G01N33/53 G01N33/566                           |         |            |

|           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CPC分类号    | A61P9/00 A61P25/00 A61P35/00 A61P37/02 C07K14/47                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| FI分类号     | C12N15/00.ZNA.A A61K39/395.D A61K39/395.N A61K45/00 A61K49/00.Z A61P9/00 A61P25/00 A61P35/00 A61P37/02 C07K1/22 C07K14/47 C07K16/18 C07K16/46 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/02 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A A61K37/02                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| F-TERM分类号 | 2G045/AA25 2G045/AA40 2G045/BB03 2G045/BB20 2G045/CB01 2G045/CB21 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/DA37 2G045/DA77 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA44 4B024/CA04 4B024/CA09 4B024/CA20 4B024/GA11 4B024/HA12 4B024/HA15 4B063/QA18 4B063/QA19 4B063/QQ02 4B063/QQ20 4B063/QQ42 4B063/QQ52 4B063/QQ79 4B063/QR55 4B063/QR59 4B063/QR62 4B063/QR69 4B063/QR80 4B063/QR82 4B063/QS12 4B063/QS24 4B063/QS25 4B063/QS28 4B063/QS34 4B064/AG01 4B064/AG27 4B064/CA10 4B064/CA19 4B064/CA20 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA93Y 4B065/AB01 4B065/AC14 4B065/BA02 4B065/CA24 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA06 4C084/AA07 4C084/AA13 4C084/AA14 4C084/AA17 4C084/BA01 4C084/BA08 4C084/BA22 4C084/BA23 4C084/BA35 4C084/BA44 4C084/CA18 4C084/CA53 4C084/CA56 4C084/CA59 4C084/NA14 4C084/ZA012 4C084/ZA362 4C084/ZB072 4C084/ZB112 4C084/ZB262 4C084/ZC512 4C085/AA13 4C085/AA14 4C085/AA16 4C085/AA19 4C085/BB11 4C085/BB31 4C085/CC02 4C085/CC03 4C085/CC04 4C085/CC05 4C085/CC13 4C085/CC21 4C085/CC22 4C085/CC23 4C085/CC32 4C085/EE01 4C085/EE06 4C085/GG01 4C085/HH01 4C085/JJ01 4C085/KA04 4C085/KA05 4C085/KA27 4C085/KB82 4C085/LL20 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/CA40 4H045/DA76 4H045/DA86 4H045/EA21 4H045/EA22 4H045/EA23 4H045/EA27 4H045/EA28 4H045/EA50 4H045/EA51 4H045/FA72 4H045/FA74 |
| 优先权       | 60/197854 2000-04-14 US<br>60/202373 2000-05-04 US<br>60/205899 2000-05-18 US<br>60/210155 2000-06-01 US<br>60/209401 2000-06-01 US                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| 外部链接      | <a href="#">Espacenet</a>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |

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。

| シオホ<br>加ハシD | シオホ<br>加ハシD | シオホ<br>加ハシD | シオホ<br>加ハシD | シオホ<br>加ハシD |
|-------------|-------------|-------------|-------------|-------------|
| 747357      | 747357      | 747357      | 747357      | 747357      |
| 747404      | 747404      | 747404      | 747404      | 747404      |
| 240555      | 240555      | 240555      | 240555      | 240555      |
| 318706      | 318706      | 318706      | 318706      | 318706      |
| 127156      | 127156      | 127156      | 127156      | 127156      |
| 134942      | 134942      | 134942      | 134942      | 134942      |
| 140156      | 140156      | 140156      | 140156      | 140156      |
| 161047      | 161047      | 161047      | 161047      | 161047      |
| 187209      | 187209      | 187209      | 187209      | 187209      |
| 260793      | 260793      | 260793      | 260793      | 260793      |
| 412044      | 412044      | 412044      | 412044      | 412044      |
| 638118      | 638118      | 638118      | 638118      | 638118      |
| 743323      | 743323      | 743323      | 743323      | 743323      |
| 1691509     | 1691509     | 1691509     | 1691509     | 1691509     |