

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2005-500008

(P2005-500008A)

(43) 公表日 平成17年1月6日(2005.1.6)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A
A 6 1 K 38/00	A 6 1 K 45/00	2 G O 4 5
A 6 1 K 45/00	A 6 1 P 1/04	4 B O 2 4
A 6 1 P 1/04	A 6 1 P 1/16	4 B O 2 9
A 6 1 P 1/16	A 6 1 P 1/18	4 B O 6 3
		4 B O 6 4
	審査請求 未請求 予備審査請求 有	(全 294 頁) 最終頁に続く

(21) 出願番号	特願2002-558506 (P2002-558506)	(71) 出願人	301005050 インサイト・ゲノミックス・インコーポレイテッド アメリカ合衆国カリフォルニア州94304・パロアルト・ポータードライブ 3160
(86) (22) 出願日	平成14年1月16日 (2002.1.16)	(74) 代理人	100089266 弁理士 大島 陽一
(85) 翻訳文提出日	平成15年7月16日 (2003.7.16)	(72) 発明者	リー、アーンステーション・エイ アメリカ合衆国カリフォルニア州94706・アルバニー・ケインズストリート 624
(86) 国際出願番号	PCT/US2002/001339		
(87) 国際公開番号	W02002/057454		
(87) 国際公開日	平成14年7月25日 (2002.7.25)		
(31) 優先権主張番号	60/262, 838		
(32) 優先日	平成13年1月19日 (2001.1.19)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/265, 927		
(32) 優先日	平成13年2月2日 (2001.2.2)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/271, 196		
(32) 優先日	平成13年2月23日 (2001.2.23)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 受容体および膜結合タンパク質

(57) 【要約】

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

【特許請求の範囲】

【請求項1】

以下の(a)乃至(e)からなる群から選択した単離されたポリペプチド。

(a) SEQ ID NO:1-15(配列番号1乃至15)からなる群から選択したアミノ酸配列を含むポリペプチド

(b) SEQ ID NO:1-14からなる群から選択したアミノ酸配列と少なくとも90%が同一であるような天然のアミノ酸配列を含むポリペプチド

(c) SEQ ID NO:15のアミノ酸配列に対して少なくとも98%が同一であるような天然アミノ酸配列を持つポリペプチド

(d) SEQ ID NO:1-15を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片 10

(e) SEQ ID NO:1-15を有する群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片

【請求項2】

SEQ ID NO:1-15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1に記載の単離されたポリペプチド。

【請求項3】

請求項1のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項4】

請求項2のポリペプチドをコードする単離されたポリヌクレオチド。 20

【請求項5】

SEQ ID NO:16-30(配列番号16乃至30)を有する群から選択したポリヌクレオチド配列を含む請求項4に記載の単離されたポリヌクレオチド。

【請求項6】

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

【請求項7】

請求項6に記載の組換えポリヌクレオチドを用いて形質転換した細胞。

【請求項8】

請求項6に記載の組換えポリヌクレオチドを含む遺伝形質転換体。 30

【請求項9】

請求項1のポリペプチドを生産する方法であって、

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

(b) そのように発現した前記ポリペプチドを回収する過程とからなる方法。

【請求項10】

前記ポリペプチドが、SEQ ID NO:1-15からなる群から選択したアミノ酸配列を持つことを特徴とする、請求項9に記載の方法。

【請求項11】

請求項1に記載のポリペプチドと特異的に結合する単離された抗体。 40

【請求項12】

以下の(a)乃至(e)からなる群から選択した単離されたポリヌクレオチド。

(a) SEQ ID NO:16-30からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド

(b) SEQ ID NO:16-30を有する群から選択したポリヌクレオチド配列と少なくとも90%が同一であるような天然のポリヌクレオチド配列を有するポリヌクレオチド

(c) (a)のポリヌクレオチドに相補的なポリヌクレオチド

(d) (b)のポリヌクレオチドに相補的なポリヌクレオチド

(e) (a)~(d)のRNA等価物 50

【請求項 13】

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

【請求項 14】

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

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

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

10

【請求項 15】

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

【請求項 16】

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

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

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

20

【請求項 17】

請求項 1 に記載のポリペプチドと、薬剤として許容できる賦形剤とを含む組成物。

【請求項 18】

前記ポリペプチドが、SEQ ID NO:1-15 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 17 の組成物。

【請求項 19】

機能性 REMAP の発現低下に関連する疾患又は病状を治療する方法であって、そのような治療を必要とする患者に対して請求項 17 に記載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 20】

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

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝すステップと、

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

【請求項 21】

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

40

【請求項 22】

機能性 REMAP の発現低下に関連する疾患又は病状を治療する方法であって、そのような治療を必要とする患者に対して請求項 21 に記載の組成物を投与する過程を含むことを特徴とする方法。

【請求項 23】

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

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝すステップと、

50

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

【請求項 24】

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

【請求項 25】

機能性 REMAP の過剰発現に関連する疾患又は病状の治療方法であって、そのような治療を必要とする患者に対して請求項 24 に記載の組成物を投与する過程を含むことを特徴とする方法。

【請求項 26】

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

(a) 請求項 1 のポリペプチドを適切な条件下で少なくとも 1 つの試験化合物に混合させる過程と、

(b) 請求項 1 のポリペプチドの試験化合物との結合を検出し、それによって請求項 1 のポリペプチドに特異結合する化合物を同定する過程を含む方法。

【請求項 27】

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

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

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

(c) 試験化合物の存在下での請求項 1 のポリペプチドの活性を、試験化合物の不存在下での請求項 1 のポリペプチドの活性と比較する過程を含み、試験化合物の存在下での請求項 1 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調節する化合物を標示するような方法。

【請求項 28】

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

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

(b) 前記標的ポリヌクレオチドの発現改変を検出する過程と、

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

【請求項 29】

試験化合物の毒性を算定する方法であって、

(a) 核酸を含む生物学的サンプルを前記試験化合物で処理する過程と、

(b) 処理した前記生物学的サンプルの核酸と、請求項 12 のポリヌクレオチドの少なくとも 20 の連続するヌクレオチドを持つプローブをハイブリダイズさせる過程であって、このハイブリダイゼーションが、前記プローブと前記生物学的サンプル中の標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われ、前記標的ポリヌクレオチドが、請求項 12 のポリヌクレオチドまたはその断片のポリヌクレオチド配列を持つポリヌクレオチドである、前記過程と、

(c) ハイブリダイゼーション複合体の収量を定量する過程と、

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

【請求項 30】

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

10

20

30

40

50

て、

(a) 前記生物学的サンプルと請求項 11 の抗体との混合を、前記抗体が前記ポリペプチドに結合し、抗体とポリペプチドとの複合体を形成するのに適した条件下で行う過程と、
(b) 前記複合体を検出する過程とを含み、前記複合体の存在が、前記生物学的サンプル中の前記ポリペプチドの存在と相関することを特徴とする方法。

【請求項 31】

前記抗体が、

- (a) キメラ抗体
- (b) 単鎖抗体
- (c) Fab断片
- (d) F(ab')₂断片
- (e) ヒト化抗体、のいずれかである抗体。

10

【請求項 32】

請求項 11 に記載の抗体と、許容できる賦形剤とを含む組成物。

【請求項 33】

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

【請求項 34】

前記抗体が標識されることを特徴とする請求項 32 に記載の組成物。

【請求項 35】

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

20

【請求項 36】

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

(a) 抗体反応を誘発する条件下で、SEQ ID NO:1-15 からなる群から選択したアミノ酸配列またはその免疫原性断片を含むポリペプチドを用いて動物を免疫化する過程と、
(b) 前記動物から抗体を単離する過程と、
(c) 前記単離された抗体を前記ポリペプチドでスクリーニングし、それによって、SEQ ID NO:1-15 からなる群から選択したアミノ酸配列を含むポリペプチドに特異結合するポリクローナル抗体を同定する過程とを含むような方法。

30

【請求項 37】

請求項 36 に記載の方法で産出したポリクローナル抗体。

【請求項 38】

請求項 37 に記載のポリクローナル抗体及び適切なキャリアーを含む組成物。

【請求項 39】

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

(a) 抗体反応を誘発する条件下で、SEQ ID NO:1-15 からなる群から選択したアミノ酸配列またはその免疫原性断片を含むポリペプチドを用いて動物を免疫化する過程と、
(b) 前記動物から抗体産出細胞を単離する過程と、
(c) 前記抗体産出細胞と不死化した細胞とを融合して、モノクローナル抗体を産出するハ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:1-15 からなる群から選択したアミノ酸配列を含むポリペプチドに特異結合するようなモノクローナル抗体を前記培養物から単離する過程とを含むことを特徴とする方法。

40

【請求項 40】

請求項 39 に記載の方法で産出したモノクローナル抗体。

【請求項 41】

請求項 40 に記載のモノクローナル抗体及び適切なキャリアーを含む組成物。

【請求項 42】

50

Fab発現ライブラリをスクリーニングすることにより産出されることを特徴とする請求項 11 に記載の抗体。

【請求項 43】

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

【請求項 44】

SEQ ID NO:1-15 からなる群から選択したアミノ酸配列を含むポリペプチドを検出する方法であって、

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

(b) 特異結合を検出する過程とを含み、該特異結合が、SEQ ID NO:1-15 からなる群から選択したアミノ酸配列を有するポリペプチドがサンプル中に存在することを標示することを特徴とする方法。

10

【請求項 45】

SEQ ID NO:1-15 からなる群から選択したアミノ酸配列を含むポリペプチドを精製する方法であって、

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

(b) 前記サンプルから前記抗体を分離し、SEQ ID NO:1-15 からなる群から選択したアミノ酸配列を有する精製ポリペプチドを得る過程、とを含むことを特徴とする方法。

20

【請求項 46】

マイクロアレイの少なくとも 1 つのエLEMENT が請求項 13 に記載のポリヌクレオチドであることを特徴とするマイクロアレイ。

【請求項 47】

ポリヌクレオチドを含むサンプルの発現プロフィールを作成する方法であって、

(a) サンプル中のポリヌクレオチドを標識化する過程と、

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

(c) サンプル中のポリヌクレオチドの発現を定量する過程とを含む方法。

【請求項 48】

固体基板上の固有の物理的位置に付着された種々のヌクレオチド分子を含むアレイであって、少なくとも 1 つの前記ヌクレオチド分子が、標的ポリヌクレオチドの少なくとも 30 の連続したヌクレオチドと特異的にハイブリダイズ可能なオリゴヌクレオチドまたはポリヌクレオチドを有し、前記の標的ポリヌクレオチドが請求項 12 に記載のポリヌクレオチドであることを特徴とするアレイ。

30

【請求項 49】

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

【請求項 50】

請求項 48 に記載のアレイで、前記の最初のオリゴヌクレオチドまたはポリヌクレオチドの配列が前記の標的ポリヌクレオチドの少なくとも 60 の連続したヌクレオチドに完全に相補的であることを特徴とするアレイ。

40

【請求項 51】

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

【請求項 52】

請求項 48 に記載のアレイで、マイクロアレイであることを特徴とするアレイ。

【請求項 53】

請求項 48 に記載のアレイで、前記のオリゴヌクレオチドまたはポリヌクレオチドの最初

50

の配列を含むヌクレオチド分子にハイブリダイズした前記の標的ポリヌクレオチドを有することを特徴とするアレイ。

【請求項 5 4】

請求項 4 8 に記載のアレイで、リンカーが少なくとも 1 つの前記のヌクレオチド分子と前記の固体基板を連結していることを特徴とするアレイ。

【請求項 5 5】

請求項 4 8 に記載のアレイで、基板上の固有の物理的位置の各々が複数のヌクレオチド分子を含み、任意の単一の固有の物理的位置でのその複数のヌクレオチド分子は同一の配列を有し、基板上の固有の物理的位置の各々は、基板上の別の固有の物理的位置でのヌクレオチド分子の配列とは異なる配列を有するヌクレオチド分子を含むことを特徴とするアレイ。

10

【請求項 5 6】

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

【請求項 5 7】

SEQ ID NO:2 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 5 8】

SEQ ID NO:3 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 5 9】

SEQ ID NO:4 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 6 0】

SEQ ID NO:5 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

20

【請求項 6 1】

SEQ ID NO:6 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 6 2】

SEQ ID NO:7 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 6 3】

SEQ ID NO:8 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 6 4】

SEQ ID NO:9 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 6 5】

30

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

【請求項 6 6】

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

【請求項 6 7】

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

【請求項 6 8】

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

【請求項 6 9】

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

【請求項 7 0】

40

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

【請求項 7 1】

SEQ ID NO:16 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 7 2】

SEQ ID NO:17 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 7 3】

SEQ ID NO:18 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 7 4】

SEQ ID NO:19 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 7 5】

50

SEQ ID NO:20のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 7 6】

SEQ ID NO:21のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 7 7】

SEQ ID NO:22のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 7 8】

SEQ ID NO:23のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 7 9】

SEQ ID NO:24のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 8 0】

SEQ ID NO:25のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 8 1】

SEQ ID NO:26のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 8 2】

SEQ ID NO:27のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 8 3】

SEQ ID NO:28のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 8 4】

SEQ ID NO:29のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 8 5】

SEQ ID NO:30のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、受容体および膜結合タンパク質の核酸配列とアミノ酸配列に関する。本発明はまた、これらの配列を利用した細胞増殖異常、自己免疫/炎症性の疾患、神経疾患、代謝異常、発達障害および内分泌疾患の診断・治療・予防に関する。本発明は、さらに受容体および膜結合タンパク質の核酸配列とアミノ酸配列の発現における外来性化合物の効果についての評価に関する。

【背景技術】

【0002】

シグナル伝達は、細胞が細胞外シグナルに応答する全般的プロセスである。原形質膜でのシグナル伝達は、ホルモン、神経伝達物質または成長因子などのシグナル分子が細胞膜受容体に結合することから開始される。このように活性化された受容体によって細胞内の生化学的カスケードが誘発され、このカスケードは転写因子などの細胞内標的分子の活性化で終る。シグナル伝達のこのプロセスは、細胞増殖、分化及び遺伝子転写を含む全てのタイプの細胞機能を制御する。

【0003】

生体膜はオルガネラと、小胞および細胞自体とを囲む。生体膜は、脂質の2重層シートから成る高度に選択的な透過性障壁であって、シートの構成素材は、ホスホグリセリド、脂肪酸、コレステロール、リン脂質、糖脂質、プロテオグリカン、およびタンパク質である。膜にはイオンポンプや、イオンチャネル、および特異的受容体類があり、受容体は外部刺激を受け、生化学的シグナルを膜の内側に伝達する。これらの膜にはまた、セカンドメッセンジャータンパク質類があって、これらのポンプやチャネルや受容体と相互作用し、これらのシグナルの伝達を増幅し調節する。

【0004】

形質膜タンパク質

形質膜タンパク質 (MP) は2群に分けられ、その分類基準は膜からのタンパク質の抽出方法である。膜の外にあるタンパク質すなわち膜表在性タンパク質の放出させるには、極端なイオン強度またはpHを用いるか、尿素などの、タンパク質相互作用の攪乱物質を用いる

10

20

30

40

50

。膜の内側にあるタンパク質すなわち膜内在性タンパク質を放出させる唯一の方法は、膜の脂質2重層を界面活性剤で溶解することである。

【0005】

既知の膜内在性タンパク質の大部分は膜貫通タンパク質(TM)であり、細胞外、膜貫通および細胞内ドメインを特徴とする。TMドメイン群は通常、15~25の疎水性アミノ酸からなり、これらのアミノ酸は1つの螺旋構造をとると予測される。TMタンパク質はバイトピック(bitopic)(タイプIおよびタイプII)と、ポリトピック(polytopic)(タイプIIIおよびタイプIV)(Singer, S.J. (1990) Annu. Rev. Cell Biol. 6:247-96)に分類され、バイトピックタンパク質は膜を1回貫通するが、ポリトピックタンパク質には、複数の膜貫通セグメントがある。TMタンパク質は種々の重要な細胞機能を果たしており、これには、シグナル伝達に關与する細胞表面受容体タンパク質としての作用が含まれる。これらの機能は、成長因子受容体および分化因子受容体、またショウジョウバエのpecanexやfrizzledタンパク質のような受容体相互作用タンパク質、LIV-1タンパク質、NF2タンパク質、およびGNS1/SUR4真核生物膜内在性タンパク質によって示される。TMタンパク質はまた、イオンまたは代謝産物のトランスポータ(例えばギャップ結合チャネル(コネキシン類)とイオンチャネル類)として作用し、また、細胞接着タンパク質(例えばレクチン類とインテグリン類とフィブロネクチン類)として作用する。TMタンパク質は、小胞オルガネラを形成する分子(例えばカベオリン(caveolin)類)、または、細胞認識分子(例えば分化クラスター(CD)抗原、糖タンパク質、およびムチン類)に存在する。

10

【0006】

多くの膜タンパク質(MP)には、タンパク質を特定の亜細胞内部に局在させるよう働く、アミノ酸配列モチーフがある。これらのモチーフの例に、PDZドメイン群や、KDEL、RGD、NGR、およびGSL配列モチーフ群や、フォンウィルブランド因子A(vWFA)ドメイン群や、EGF様ドメイン群がある。RGD、NGR、およびGSLモチーフ含有ペプチド類は、腫瘍血管系の標的癌を治療するための薬物送達剤として用いられている(Arap, W. 他(1998) Science, 279:377-380)。さらに、膜タンパク質にはまた、糖質認識ドメイン群(CRD)のような細胞外分子や細胞内分子と相互作用するように働くアミノ酸配列モチーフ群がある。

20

【0007】

アミノ酸残基側鎖の化学修飾により、MP類が他の分子(例えば膜リン脂質)と相互作用する方法が変わる。このような化学修飾の例に、グリコサミノグリカン類や、少糖類、リン脂質類、アセチル部分およびパルミトイル部分、ADPリボース、リン酸、および硫酸基との、共有結合の形成がある。

30

【0008】

膜タンパク質をコードするRNAは選択的スプライス部位を持つことがあり、これらの部位により、同一遺伝子がコードするタンパク質群が、異なるメッセンジャーRNAとアミノ酸配列とを持ち得る。スプライス変異膜タンパク質は、他のリガンドやタンパク質のアイソフォームと相互作用し得る。

【0009】

受容体

受容体の用語は他の分子を特異的に認識するタンパク質を意味する。そのカテゴリーは広く、様々の機能を持つタンパク質が含まれる。大部分の受容体は、細胞外リガンドと結合して、成長、分化、飲食細胞運動および免疫応答の領域において細胞内応答を生じる細胞表面タンパク質である。他の受容体は、小胞体からのタンパク質の選択的な輸送を促進し、細胞内の特定の部位に酵素を局在化させる。この用語はまた、リガンドの受容体として既知の、あるいは未知の化学的組成物と作用するタンパク質や、他の細胞構成成分と相互作用するタンパク質にも適用され得る。例えば、ステロイドホルモン受容体はDNAの転写物に結合し、調節する。

40

【0010】

細胞表面受容体は一般に原形質膜内在性タンパク質である。これらの受容体はカテコールアミンのようなホルモン、ペプチドホルモン、成長因子、分化因子、甲状腺刺激ホルモン

50

放出ホルモン、ガラニン、ソマトスタチン、およびタキキニンおよび循環系で運ばれるシグナル伝達分子を認識する。

免疫系細胞上の細胞表面受容体は、抗原、抗体、および主要組織適合性複合体 (MHC) 結合ペプチドを認識する。他の細胞表面受容体はリガンド類に結合し、リガンドはその細胞によって取り込まれる。この受容体介在性エンドサイトーシスは、低密度リポタンパク質 (LDL)、トランスフェリン、グルコース末端糖タンパク質、マンノース末端糖タンパク質、ガラクトース末端糖タンパク質、免疫グロブリン、ホスホビテロゲニン (phosphovitellogenin) 類、フィブリン、プロテイナーゼ-阻害剤複合体、プラスミノゲン活性化因子、およびトロンボスポンジンの取り込み時に機能する (Lodish, H. 他 (1995) *Molecular Cell Biology*, Scientific American Books, New York NY, p. 723; Mikhailenko, I. 他 (1997) *J. Biol. Chem.* 272:6784-6791)。

10

【0011】

受容体プロテインキナーゼ

多くの成長因子受容体 (例えば、上皮性成長因子、血小板由来成長因子、線維芽細胞成長因子、および成長調節因子 トロンピン等) には固有のプロテインキナーゼ活性がある。成長因子が受容体に結合すると、受容体のセリン、トレオニン、またはチロシン残基の自己リン酸化が誘発される。これらのリン酸化部位は他の細胞質シグナル伝達タンパク質結合の認識部位となる。これらのタンパク質は、細胞表面での最初の受容体の活性化を最終的に特異的細胞内標的分子の活性化に結びつけるシグナル伝達経路に参与する。チロシン残基自己リン酸化の場合、これらのシグナル伝達タンパク質には Src 相同性 (SH) ドメインと呼ばれる共通ドメインが含まれている。SH2 ドメインおよび SH3 ドメインはホスホリパーゼ C- β 、PI-3-K p85 調節サブユニット、Ras-GTPアーゼ活性化タンパク質および pp60^{c-src} に存在する。 (Lowenstein, E.J. 他 (1992) *Cell* 70:431-442)。サイトカインファミリーの受容体は異なった共通結合ドメインを共有しており、成長ホルモン (GH)、インターロイキン、エリスロポエチンおよびプロラクチンが含まれる。

20

【0012】

他の受容体およびセカンドメッセンジャー結合タンパク質は内在性セリン/トレオニンプロテインキナーゼ活性を有する。これらにはアクチビン/TGF- β /BMP-スーパーファミリー受容体、カルシウムおよびジアシルグリセロール活性化/リン脂質依存性プロテインキナーゼ (PK-C) および RNA 依存性プロテインキナーゼ (PK-P) が含まれる。さらに、線虫 Twitchin を含む他のセリン/トレオニンプロテインキナーゼはフィブロネクチン様、免疫グロブリン C2 様ドメインを有する。

30

【0013】

G タンパク質共役受容体

今までに同定されている遺伝子の最も大きいファミリーの 1 つによってコードされる G タンパク質共役受容体 (GPCR) は、原形質膜での細胞外シグナルの伝達において中心的役割を果たす。GPCR が治療の標的として成功していることは歴史的に証明されている。

【0014】

GPCR は、7 つの疎水性膜貫通ドメインを有することを特徴とする膜内在性タンパク質であり、それらのドメインがまとまって逆平行アルファ (α) 螺旋の束を形成するようになる。GPCR のサイズは、400 以下から 1000 以上のアミノ酸に及ぶ (Strosberg, A.D. (1991) *Eur. J. Biochem.* 196:1-10, Coughlin, S.R. (1994) *Curr. Opin. Cell Biol.* 6:191-197)。GPCR のアミノ末端は、細胞外にあり、長さが可変で、多くの場合グリコシル化される。カルボキシル末端は、細胞質内にあり、通常はリン酸化される。細胞外ループは、細胞内ループと交互に現れ、膜貫通ドメインに連結している。第 2 及び第 3 細胞外ループを結びつけるシステインジスルフィド架橋は、アゴニスト及びアンタゴニストと相互作用し得る。GPCR の最も保存されたドメインは、膜貫通ドメイン及び最初の 2 つの細胞質ループである。膜貫通ドメインによって、受容体の構造的及び機能的特徴がある程度決まる。大抵の場合、 α ヘルックスの束がリガンド結合ポケットを形成する。細胞外 N 末端セグメント、または 3 つの細胞外ループの内の 1 つまたは複数のループが、リガンド結合にも

40

50

関与し得る。リガンド結合によって、受容体の細胞内部分に構造的変化が誘発されて受容体が活性化される。次に、この活性化された受容体の大きな第3の細胞内ループがヘテロ3量体であるグアニンヌクレオチド結合Gタンパク質複合体と相互作用し、このGタンパク質複合体が更に細胞内のシグナル伝達作用を媒介する。この細胞内のシグナル伝達作用には、サイクリックAMP (cAMP)、ホスホリパーゼC、およびイノシトール三リン酸などのセカンドメッセンジャーの活性化、並びに活性化されたGPCRとイオンチャネルタンパク質との相互作用が含まれる (Watson, S.およびS. Arkininstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, 2-6ページ、Bolander, F.F. (1994) Molecular Endocrinology, Academic Press, San Diego CA, 162-176ページ、Baldwin, J.M. (1994) Curr. Opin. Cell Biol. 6:180190等参照)。

10

【0015】

GPCRには、感覚性シグナルメディエータ (例えば光及び嗅覚刺激性分子) の受容体、アデノシン、アミノ酪酸 (GABA)、肝細胞成長因子、メラノコルチン、ニューロペプチドY、オピオイドペプチド、オプシン、ソマトスタチン、タキキニン、血管作用性腸管ポリペプチドファミリー及びバソプレシン、生体アミン (例えばドーパミン、エピネフリン及びノルエピネフリン、ヒスタミン、グルタミン酸 (代謝調節作用)、アセチルコリン (ムスカリン様作用) 及びセロトニン)、ケモカイン、炎症の脂質メディエータ (例えばプロスタグランジン及びプロスタノイド、血小板活性化因子及びロイコトリエン) 及びペプチドホルモン (例えばボンベシン、ブラジキニン、カルシトニン、C5aアナフィラトキシン、エンドセリン、卵巣刺激ホルモン (FSH)、ゴナドトロピン放出ホルモン (GnRH)、ニューロキニン及び甲状腺刺激ホルモン放出ホルモン (TRH) 及びオキシトシン) がある。まだ同定されていない刺激の受容体として作用するGPCRは、オーファン受容体として知られている。

20

【0016】

GPCRの最も大きいファミリーはロドプシン様GPCRからなり、ロドプシン様GPCRはホルモン、神経伝達物質及び光を含む多様な細胞外シグナルを伝達する。ロドプシンは、動物の網膜に見られる感光性GPCRである。脊椎動物では、ロドプシン分子は光受容体 (杆体) 細胞に見られる膜性スタックに包埋されている。各ロドプシン分子は、cGMPレベルの低下を誘発し、これにより原形質膜ナトリウムチャネルが閉鎖することにより1光子の光に反応する。この方法で、可視シグナルが神経インパルスに変換される。その他のロドプシン様GPCRは、神経伝達物質への応答に直接関与する。このようなGPCRには、アドレナリンに対する受容体 (アドレナリン受容体)、アセチルコリンに対する受容体 (ムスカリン様受容体)、アデノシンに対する受容体、ガラニンに対する受容体及びグルタミン酸に対する受容体 (N-メチル-D-アスパラギン酸 / NMDA受容体) がある (Watson, S.およびS. Arkininstall (1994) The G-Protein Linked Receptor Facts Book, Academic Press, San Diego CA, 7-9, 19-22, 32-35, 130-131, 214-216, 221-222ページ、Habert-Ortoli, E. 他 (1994) Proc. Natl. Acad. Sci. USA 91:9780-9783に概説されている)。

30

【0017】

GPCRの最大のサブファミリーである嗅覚受容体もまた、ロドプシン様GPCRファミリーのメンバーである。この受容体は、嗅覚シグナルの伝達により機能する。異なる臭気を区別するためには、複数の異なる嗅覚受容体が必要である。各嗅覚知覚ニューロンは1種類の嗅覚受容体しか発現せず、異なった受容体を発現するニューロンが、鼻の経路中の異なる位置を占める。例えば、ラット脳ライブラリから単離したRA1c受容体は、脳のまさに特有の領域及び画定された嗅覚上皮の特定の区域だけに発現が限定されていることを示した (Raining, K. 他 (1998) 受容体 Channels 6:141-151)。

40

【0018】

嗅粘膜もまた、さらに臭気物質結合タンパク質群を所有すると考えられている。臭気物質結合タンパク質群は別の種類の臭気物質を認識し、結合する。例えば、嗅粘膜に高度に発現されるが、他の組織では検出されないmRNAに相当するラットのcDNAクローンが単離されている。これらのクローンによってコードされるタンパク質はリポ多糖類またはポリ塩

50

化ビフェニル類に結合するタンパク質に相同性であり、異なったタンパク質が粘膜組織の特定の部位に発現されると考えられる。これらのタンパク質は、臭気物質受容体による特異的認識の前後で臭気物質と相互作用する、おそらく、選択的シグナルフィルタとして作用すると考えられる (Dear, T.N. 他 (1991) *EMBO J.* 10:2813-2819; Vogt, R.G. et al. (1991) *J. Neurobiol.* 22:74-84)。

【 0 0 1 9 】

セクレチン様GPCRサブファミリーのメンバーは、セクレチン、カルシトニン、グルカゴン、成長ホルモン放出ホルモン、副甲状腺ホルモン、及び血管作用性小腸ペプチドなどのペプチドホルモンをリガンドとして有する。例えば、セクレチン受容体は、膵臓及び小腸で酵素及びイオンの分泌を刺激するペプチドホルモンであるセクレチンに対応する (前出の Watson, 278-283ページ)。セクレチン受容体は、長さが約 4 5 0 アミノ酸であり、胃腸細胞の原形質膜に存在する。セクレチンがその受容体に結合することにより、cAMPの産出が誘発される。

10

【 0 0 2 0 】

炎症及び免疫応答に結びつけられるセクレチン様GPCRの例には、EGFのモジュールを含んだムチン様ホルモン受容体 (Emrl) 及びCD97受容体タンパク質がある。CD97は、主に白血球で発現し、活性化されたB及びT細胞上で顕著に上方制御される (McKnight, A.J.及びS. Gordon (1998) *J. Leukoc. Biol.* 63:271-280)。これらのGPCRは、最近特徴付けられたEGF-TM7受容体サブファミリーのメンバーである。これら膜7回貫通ホルモン受容体は、*in vivo*でヘテロ二量体として存在し、3から7の潜在的カルシウム結合EGF様モチーフを含む。このEGFモチーフは長さが約 4 0 アミノ酸残基で、6つの保存されたシステイン残基を持ち、シグネチャ配列のN末端の近くにはカルシウム結合部位がある。アスパラギン酸またはアスパラギン残基の翻訳後のヒドロキシル化はいくつかのタンパク質のEGF様ドメインと関連している (Prosite PDOC00010 アスパラギン酸およびヒドロキシル化部位)。

20

【 0 0 2 1 】

カルシウム結合EGF様ドメインシグネチャ配列を有する多数のタンパク質が成長と分化に関与している。例としては、軟骨と骨の形成を誘導する骨形態形成タンパク質 1 ; ショウジョウバエの上皮発生タンパク質である *c r u m b s* ; 神経の成長と分化に関与する Notchおよび多数の相同体、およびトランスフォーミング成長因子ベータ 1 結合タンパク質等が挙げられる (Expasy PROSITE document PDOC00913; Soler, C. および Carpenter, G., Nicola, N.A. (1994) *The Cytokine Facts Book*, Oxford University Press, Oxford, UK, 193-197ページ)。EGF様ドメインはさまざまなタンパク質のタンパク質間相互作用を媒介する。たとえば、ECM糖タンパク質フィブリン 1 のEGF様ドメインは自己会合とフィブロネクチンへの結合の両方を媒介することが示されている (Tran, H. 他 (1997) *J. Biol. Chem.* 272:22600-22606)。ECMタンパク質のEGF様ドメインにおける点突然変異はマルファン症候群および偽軟骨芽形成 (pseudochondroplasia) 等のヒトの疾患の原因として同定されている (Maurer, P. 他 (1996) *Curr. Opin. Cell Biol.* 8:609-617)。

30

【 0 0 2 2 】

GPCR突然変異は、機能または恒常的活性化の喪失を招き得るものであり、多数のヒト疾患に関与してきた (前出の Coughlin)。例えば、色素性網膜炎はロドプシン遺伝子の突然変異から発生し得る。更に、甲状腺刺激ホルモン受容体の体細胞活性化突然変異は、機能亢進甲状腺アデノーマの原因となることが報告されており、恒常的活性化に感受性が高い或るGPCRが癌原遺伝子のように振舞い得ることを示唆している (Parma, J. 他 (1993) *Nature* 365:649-651)。以下のリガンド即ち黄体形成ホルモン (思春期早発症)、バソプレシン V_2 (X連鎖の腎原発性糖尿病)、グルカゴン (糖尿病及び高血圧)、カルシウム (副甲状腺機能亢進症、低カルシウム尿症 (hypocalcuria)、高カルシウム血症)、副甲状腺ホルモン (短肢小人症)、 β_3 -アドレナリン受容体 (肥満症、非インスリン依存型糖尿病)、成長ホルモン放出ホルモン (小人症) 及び副腎皮質刺激ホルモン (糖質コルチコイド欠乏症) に対するGPCR受容体には、ヒトの疾患に関連する突然変異も含まれる (Wilson, S.

40

50

他 (1998) Br. J. Pharmacol. 125:1387-.1392、Stadel, J.M. ら (1997) Trends Pharmacol. Sci. 18:430-437)。GPCRは、抑うつ症、分裂病、不眠症、高血圧、不安、ストレス、腎不全その他幾つかの心血管障害にも関与している (Horn, F. および G. Vriend (1998) J. Mol. Med. 76:464-468)。

【0023】

更に、GPCRの活性化及び阻害に直接作用する数百の新薬が過去20年のうちに認知されてきた。これらの薬剤の治療標的は、癌、骨粗鬆症、子宮内膜症のみならず心血管障害、胃腸障害、中枢神経系疾患を含めた広範囲の疾病及び疾患に及ぶ (前出のWilson、同Stadel)。例えば、ドーパミンアゴニストのLドーパはパーキンソン病の治療に用いられ、ドーパミンアンタゴニストは精神分裂病及びハンチントン病の初期段階の治療に用いられる。アドレナリン受容体のアゴニスト及びアンタゴニストは、喘息、高血圧その他の心血管障害及び不安の治療に用いられてきた。ムスカリン様アゴニストは、緑内障及び頻脈の治療に用いられ、セロトニン5HT1Dアンタゴニストは片頭痛に対して用いられ、ヒスタミンH1アンタゴニストはアレルギー性及びアナフィラキシー性反応、枯草熱、痒み及び動揺病に対して用いられる (前出のHorn)。

10

【0024】

核ホルモン受容体

核受容体又は細胞内受容体とも呼ばれる核ホルモン受容体群は、1つのタンパク質スーパーファミリーを構成する。このスーパーファミリーのメンバーは、受容体及び転写調節因子の両方である。核ホルモン受容体は、その受容体機能及び転写調節機能の両方によって、発生、恒常性、細胞増殖、及び細胞分化を含む広範な生体プロセスに影響を与える (Mangelsdorf, D. J.他 (1995) Cell 83 : 835-840 ; Wen, D. X. 及び D. P. McDonnell (1995) Curr. Opin. Biotechnol. 6 : 582-589 ; Perlmann, T. 及び R. M. Evans (1997) Cell 90 : 391-397 ; Tenbaum, S. 及び A. Baniahmad (1997) Int. J. Biochem. Cell Biol. 29 : 1325-1341 ; Moras, D. 及び H. Gronemeyer (1998) Curr. Opin. Cell Biol. 10 : 384-391 ; Willy, P. J. 及び D. J. Mangelsdorf (1998) Hormones and Signaling (編集 : B. W. O'Malley) vol. 1, Academic Press, San Diego CA, 307-358ページ ; Weatherman, R. V.他 (1999) Annu. Rev. Biochem. 68 : 559-581.)。

20

【0025】

受容体としての核ホルモン受容体

一般に、受容体とは、他の分子を特異的に認識するタンパク質を指す。核ホルモン受容体は受容体として、それらの同種リガンドを特異的に認識してそれらに結合する。核ホルモン受容体は細胞内に局在するが、多くの受容体は、細胞外リガンドと結合する細胞外の細胞表面タンパク質である。このような細胞外受容体は、成長、分化、エンドサイトーシス、及び免疫反応に影響を与える細胞応答を引き起こす。他の受容体は、小胞体からのタンパク質の選択的な輸送を促進し、細胞の特定の領域に酵素を局在化させる。核ホルモン受容体による転写調節、細胞外受容体による細胞内へのシグナル伝達、及び他の受容体によるタンパク質の輸送及び局在化は全て、受容体と様々な細胞成分との間の特異的な相互作用に依存する。多くの場合、受容体が結合する同種リガンドが同定されていない。このような受容体はオーファン受容体と呼ばれる。リガンドと結合せずに転写調節機能を果たす核ホルモン受容体もオーファン受容体である。

30

40

【0026】

転写調節因子としての核ホルモン受容体

多細胞生物は、構造及び機能が著しく異なる多様な細胞型から構成されている。細胞型は遺伝子発現パターンの特徴によって決定され、また異なった細胞型は発生中、重複するが固有の一組の遺伝子群を発現する。遺伝子発現の空間的及び時間的な調節は細胞増殖、細胞分化、アポトーシス、及び生物の発達に必要なその他のプロセスにとって極めて重要である。核ホルモン受容体は転写調節因子として、これらの基本的な生体プロセスの調節において重要な役割を果たしている。他の転写調節因子は、細胞間伝達の仲介および様々な細胞型の作用の調整を行なう細胞外シグナルに反応して遺伝子発現に影響を与える。

50

【0027】

一般に、核ホルモン受容体等の転写調節因子は、配列特異的な要領でプロモーター、エンハンサー、及び遺伝子上流の調節領域と結合して、遺伝子転写の開始、活性化、抑制、または停止を行う。しかしながら、或る種の転写調節因子は、遺伝子コード領域内或いはその下流の調節エレメントと結合する。転写調節タンパク質が、単体で或いは他のアクセサリーファクターと複合体を形成してDNAの特定の領域に結合し得る (Lewin, B. (1990) *Genes IV*, Oxford University Press, New York NY, and Cell Press, Cambridge MA, 554-570ページ)。

【0028】

核ホルモン受容体機能のメカニズム

核ホルモン受容体は、リガンドと結合していない状態では、hsp90等の熱ショックタンパク質及びhsp56等のイムノフィリンを含むシャペロンの多タンパク質複合体と結合して存在する。これらのシャペロンは、遊離リガンドの結合が受け入れられるような不活性化状態にリガンドのない受容体を維持し、リガンドのない受容体が核に転位するのを防止する。同種リガンドによって活性化されると、核ホルモン受容体はホモ二量体又はヘテロ二量体を形成することができ、そのホモ二量体又はヘテロ二量体は核に転位して、特定のDNA配列に結合し、転写調節機能を発揮する。核ホルモン受容体の調節の役割を効果的に果たすために、活性化した核ホルモン受容体はヒストンデアセチラーゼ含有コリプレッサー複合体から解離し、ヒストンアセチルトランスフェラーゼ含有コアクティベーター複合体と結合する (Xu, L.他 (1999) *Curr. Opin. Genet. Dev.* 9 : 140- 147)。活性化した受容体がコアクティベータタンパク質と結合してクロマチンが再構築されることにより、転写的に開いた活性化状態になり、活性化した核受容体の転写調節エレメントに到達できるようになる (Lemon, B. D. 及び L. P. Freedman (1999) *Curr. Opin. Genet. Dev.* 9 : 499-504)。

【0029】

核ホルモン受容体の構造

核ホルモン受容体は、ホルモンシグナルを転写反応に変換してシグナルトランスデューサーとして作用する。一般に、核ホルモン受容体は、可変アミノ末端ドメイン、高度に保存されたDNA結合ドメイン、及び保存されたC末端リガンド結合ドメインからなる。ステロイド結合核ホルモン受容体では、アミノ末端ドメインがAF-1と呼ばれるトランス活性化エレメントを匿っている。或る種の核ホルモン受容体はまた、AF-2と呼ばれるリガンド結合ドメインにトランス活性化エレメントを含む。核ホルモン受容体のDNA結合ドメイン及びリガンド結合ドメインは二量体化エレメントを含み得る。また、DNA結合ドメインは核局在化シグナルを含み得る (Weatherman, R. V.他 (1999) *Annu. Rev. Biochem.* 68 : 559-581)。

【0030】

核ホルモン受容体のDNA結合ドメインは、特定のDNA配列の認識を媒介する2つのZnフィンガーモチーフからなる。Znフィンガーモチーフは、 Zn^{+2} を配位結合する周期的に離間したヒスチジン残基及びシステイン残基を含む。この配列パターンの例には、C2H2型、C4型、及びC3HC4型(「RING」フィンガー)のZnフィンガー、及びPHDドメインが含まれる (Lewin, 前出 ; Aasland, R. 他 (1995) *Trends Biochem. Sci.* 20 : 56-59)。Znフィンガーモチーフは、構造及び隣接状態が亜鉛イオンによって維持されている1つの逆平行シート及び1つのヘリックスを含む。DNAとの接触は、ヘリックスの前のアルギニン残基によって、並びにヘリックスの第2の残基、第3の残基、及び第6の残基によってなされる。Znフィンガーモチーフは、タンパク質の各ZnフィンガーのヘリックスがDNAの二重螺旋の主な溝に接触するように、タンパク質内でタンデムに反復している。この蛋白質とDNAの間における接触が反復されることにより、強力かつ特異的なDNAタンパク質相互作用が生じる。相互作用の強さと特異性はタンパク質内におけるZnフィンガーモチーフの数によって調節される。Znフィンガーは、初めはDNAと直接相互作用する領域としてDNA結合タンパク質において見出されたが、その後DNAと結合しないタンパク質からも見出された

10

20

30

40

50

(例えばLodish, H.他 (1995) *Molecular Cell Biology*, Scientific American Books, New York NY, 447-451ページ参照)。

【0031】

核ホルモン受容体のリガンド結合ドメインは、リガンド、コアクティベータタンパク質、及びコリプレッサータンパク質との結合に参与する。このドメインは三層のヘリックスからなり、中間の層は多数の疎水性測鎖を含む2つのヘリックスからなる (Moras, D. 及び H. Gronemeyer (1998) *Curr. Opin. Cell Biol.* 10 : 384-391)。従って、これらの2つの中間層のヘリックスが、リガンド結合部位である疎水性ポケットを形成している。この疎水性リガンド結合部位のリガンド結合は、受容体タンパク質の内部に完全に埋もれ、溶液に露出していない。このことから、リガンドの結合にはリガンド結合ドメインにおける大きな構造の変化を伴うことが示唆される。リガンド結合ドメインのヘリックスの一つは、核受容体の二量体にサブユニット間の接触の多くを提供する。このヘリックスはリガンド結合ポケットにおいて結合している時はそのリガンドに接触することから、リガンド結合が受容体の二量体の形成に影響を与え得ることが示唆される (Weatherman, R. V.他 (1999) *Annu. Rev. Biochem.* 68 : 559-581)。

10

【0032】

核ホルモン受容体のクラス

核ホルモン受容体は、ステロイド受容体、RXR-ヘテロ二量体受容体、及びオーファン核ホルモン受容体の3つのクラスに大きく分類することができる。ステロイド受容体はステロイドホルモンに結合し、このクラスにはアンドロゲン受容体、ミネラルコルチコイド受容体、エストロゲン受容体、糖質コルチコイド受容体、及びプロゲステロン受容体が含まれる。RXR-ヘテロ二量体受容体は非ステロイドリガンドと結合し、このクラスには、甲状腺ホルモン受容体、レチノイン酸受容体、ビタミンD受容体、エクジソン受容体、及びペルオキシソーム増殖因子活性化受容体が含まれる。オーファン核ホルモン受容体には、ステロイド産生因子1、神経成長因子誘導性受容体、及びX染色体連鎖オーファン受容体DAX-1が含まれる。

20

【0033】

ステロイドホルモン受容体は、特定のステロイドホルモンと結合すると活性化する。リガンドの結合によって起こる構造の変化により、熱ショックタンパク質から受容体が解離し、ホルモン応答エレメント (HRE) と呼ばれる特定のパンドロームDNA配列を認識する受容体ホモ二量体が形成される。ステロイドホルモン受容体ホモ二量体はHREに結合すると標的遺伝子の転写を調節し得る。

30

例えばプロゲステロン受容体 (PR) は、女性の生殖周期を変化させる重要な要素であるコレステロール由来のプロゲステロン即ち4-プレグネン-3,20-ジオンによって活性化されるステロイドホルモン受容体である。これらの変動は、月経及び妊娠を含む解剖学的及び形態学的変化と相関する。プロゲステロンの作用はPRによって仲介される。細胞質では、PRは、幾つかの他のタンパク質及びPRヘテロ複合体として知られる因子と結合している。このヘテロ複合体には、hsp70、hsp90、hsp27、p59 (hsp56)、p48、及びp23等のイムノフィリン及び熱ショックタンパク質が含まれる (Johnson, J. L.他 (1994) *Mol. Cell. Biol.* 14 : 1956-1963)。プロゲステロンの結合で活性化されたPRが核に転移し、正準DNA転写エレメントに結合して分化及び細胞周期に関係するプロゲステロン調節性遺伝子を調節する (Moutsatsou, P 及び C. E. Sekeris (1997) *Ann. N. Y. Acad. Sci.* 816 : 99-115)。妊娠を終了させるのに用いられ得るPRアンタゴニストRU486は、ステロイドホルモン受容体を標的とした市販の治療薬の一例である。

40

【0034】

RXRヘテロ二量体核受容体は、そのグループのメンバーがレチノイドX受容体 (RXR) とヘテロ二量体を形成するとその標的DNA配列と結合することから、ステロイドホルモン受容体と区別される (Mangelsdorf, D. J. 及び R. M. Evans (1995) *Cell* 83 : 841-850)。RXRの3種類のアイソフォームが同定された (Minucci, S. 及び K. Ozato (1996) *Curr. Opin. Genet. Dev.* 6 : 567-574)。レチノイン酸受容体 (RAP) はRXRヘテロ二量体核受

50

容体の例である。レチノイン酸 (RA) は、主に魚の肝油、肝臓、卵黄、バター、及びクリームに存在する脂溶性ビタミンであるビタミンA (レチノール) の生物学的に活性な代謝産物である。9-シス-RAはRAR/RXRに結合するが、オール-トランス-RAはRXRのみと結合する。RAR/RXRヘテロ二量体は、レチノイン酸応答エレメント (RARE) として知られる特定のDNA配列と高親和的に結合するため、RA依存性転写の制御因子として作用する。

【0035】

ペルオキシソーム増殖因子活性化受容体 (PPAR) は、脂肪酸及びエイコサノイドによって誘導される治療的に重要なRXR-ヘテロ二量体核受容体である。それぞれが特定の発現パターンを有するPPARの3種類の既知のイソ型が存在し、これらはグルコース及び脂質の全身の恒常性に関する遺伝子の調節に関する (Kliwer, S. A. 及び T. M. Willson (1998) *Curr. Opin. Genet. Dev.* 8 : 576-581 ; Michalik, L. 及び W. Wahli (1999) *Curr. Opin. Biotechnol.* 10 : 564-570)。従って、PPARは糖尿病、異脂肪血症、及び肥満症等の疾患の治療標的である (Smith, S. A. (1996) *Pharmacol. Rev. Commun.* 8 : 57-64 ; Willson, T. M. 及び W. Wahli (1997) *Curr. Opin. Chem. Biol.* 1 : 235-241 ; Barroso, I.他 (1999) *Nature* 402 : 880-883)。

10

【0036】

オーファン核受容体は、活性化するリガンドを有していない核受容体、或いはリガンド結合してもその影響がなく転写調節活性を発揮し得る核受容体である。例えば、線虫において、X染色体によってコードされる核受容体相同体SEX-1は、性決定遺伝子xol-1の転写を調節する (Carmi, I.他 (1998) *Nature* 396 : 168-173)。SEX-1はリガンド結合に依存するのではなく、X染色体カウント機構による性分化調節において、用量応答的に転写調節因子として影響を与える。

20

【0037】

或る種の核ホルモン受容体は、通常は核ホルモン受容体ファミリーと結合する通常のDNA結合ドメインを含まない。DAX-1は、通常のDNA結合ドメインを含まない核ホルモン受容体の一つであって、DAX-1における突然変異は先天性X染色体連鎖副腎形成不全を引き起こすことが分かっている (Zanaria, E. F.他 (1994) *Nature* 372 : 635-641)。DAX-1はオーファン核受容体であって、ステロイド産生因子1 (SF-1) と直接相互作用する (Ito, M.他 (1997) *Mol. Cell. Biol.* 17 : 1476-1483)。DAX-1はまた、性特異的遺伝子発現におけるSF-1の作用を調節する能力がある (Nachtigal, M. W.他 (1998) *Cell* 445-454)。SF-1は、副腎及び生殖腺の幾つかのステロイド産生酵素遺伝子の転写調節因子として作用するオーファン核受容体であり、(Lala, D. S.他 (1992) *Mol. Endocrinol.* 6 : 1249-1258 ; Lynch, J. P. 他 (1993) *Mol. Endocrinol.* 7 : 776-786 ; Clemens, J. W. 他 (1994) *Endocrinology* 134 : 1499-1508)、下垂体性腺刺激細胞で発現される幾つかの遺伝子も調節し得る (Barnhart, K. M. 及び P. L. Mellon (1994) *Mol. Endocrinol.* 8 : 878-885 ; Ingraham, H. A. 他 (1994) *Genes Dev.* 8 : 2302-2312 ; Halvorson, L. M. 他 (1996) *J. Biol. Chem.* 271 : 6645-6650 ; Keri, R. A. 及び J. H. Nilson (1996) *J. Biol. Chem.* 271 : 10782-10785)。

30

【0038】

SF-1はまた小ヘテロ二量体 (SHP ; short heterodimer partner) パートナーの強力なトランス活性化因子として作用する (Lee, Y. K. 他 (1999) *J. Biol. Chem.* 274 : 20869-20873)。SHPもまた、通常のDNA結合ドメインを含まない核ホルモン受容体である (Seol, W. 他 (1996) *Science* 272 : 1336-1339 ; Lee, H.-K. 他 (1998) *J. Biol. Chem.* 273 : 14398-14402)。SHPは核ホルモン受容体ファミリーの多くのメンバーと相互作用する。その中には、レチノイド受容体、エストロゲン受容体、甲状腺ホルモン受容体、及びオーファン受容体CARが含まれる。SHPは、エストロゲン受容体への結合においてコアアクティベーターと競合することで、エストロゲン受容体仲介性転写活性化のインヒビターとして作用する (Johansson, L. 他 (1999) *J. Biol. Chem.* 274 : 345-353)。SHPはまた、オーファン受容体肝細胞核因子4及びレチノイドX受容体によってトランス活性化を阻害する (Lee, Y. K. 他 (2000) *Mol. Cell. Biol.* 20 : 187-195)。

40

50

【0039】

不完全な転写調節による結果

ヒトにおける腫瘍疾患の多くは不適切な遺伝子発現によって起こり得る。悪性の細胞成長は、腫瘍促進遺伝子の過剰な発現又は腫瘍抑制遺伝子の不十分な発現によって起こり得る (Cleary, M. L. (1992) *Cancer Surv.* 15 : 89-104)。染色体転座によって、或る遺伝子のコード配列を別の関連のない遺伝子の調節領域と結合させるキメラ座 (chimeric loci) が生成され得る。このような配列は不適切な遺伝子転写を引き起こし、悪性腫瘍を引き起こす可能性もある。

【0040】

更に、免疫系は、細胞防御機構の漸進的な選択、増幅、及び動員を調整するイベントのカスケードを活性化することにより感染や外傷に应答する。遺伝子の活性化及び抑制の複雑かつバランスの取れたプログラムが、このプロセスに関与している。しかしながら、遺伝子発現の不適切或いは不十分な調節によって起こる免疫系の過度の活性化は、組織や器官に相当な損傷を与え得る。この損傷は、関節炎、アレルギー、心臓発症、発作、及び感染に関連する免疫応答についての文献に詳しく記載されている (例えば Isselbacher 他 (1996) *Harrison's Principles of Internal Medicine*, 13/e, McGraw Hill, Inc. and Teton Data Systems Software参照)。

【0041】

更に、多細胞生物の成長は発生の適切な段階における細胞分化の誘導及び調整に基づいている。このプロセスの中心は、体中の細胞及び組織に固有性を与える示差的な遺伝子発現である。発生中に遺伝子発現の調節に失敗すると発生障害が起こり得る。

【0042】

リガンド依存性受容体イオンチャネル

リガンド依存性受容体イオンチャネルは2つのカテゴリーに分かれる。第1のカテゴリーである細胞外リガンド依存性受容体イオンチャネル (ELG) は、神経伝達物質結合を迅速なシナプス神経伝達のような電気シグナルに急速に変換する。ELG機能は翻訳後修飾によって調節される。第2のカテゴリーである細胞内リガンド依存性受容体イオンチャネル (ILG) は多くの細胞内セカンドメッセンジャーによって活性化され、チャネル開放応答に作用する際、翻訳後修飾を必要としない。

【0043】

ELGは、活動電位の発生の閾値にまで興奮性細胞を脱分極する。非興奮性細胞においては、ELGはアゴニストの存在時に限られたカルシウムイオン流入を可能にする。ELGにはアセチルコリン、L-グルタミン酸、グリシン、ATP、セロトニン、GABAおよびヒスタミンのような神経伝達物質によって直接制御されるチャネルが含まれる。ELG遺伝子は構造的かつ機能的に強い類似性を有するタンパク質をコードする。ILGは異なった、関連しない遺伝子ファミリーによってコードされ、また、cAMP、cGMP、カルシウムイオン、ATPおよびアラキドン酸の代謝産物に対する受容体群を含む。

【0044】

マクロファージスカベンジャー受容体

マクロファージスカベンジャー受容体類は広いリガンド特異性を持っており、低密度リポタンパク質 (LDL) や外来抗原との結合に関与し得る。スカベンジャー受容体タイプIおよびIIは三量体膜タンパク質であり、各サブユニットには小さいN末端細胞内ドメインと、1つの膜貫通ドメインと、1つの大きな細胞外ドメインと、C末端システインリッチドメインとがある。細胞外ドメインには1つの短いスパードメインと、1つの螺旋コイルドコイルドメインと、1つの3重螺旋コラーゲン性ドメインとがある。これらの受容体は或る範囲のリガンド群と結合することが示され、リガンドの例には、化学修飾されたりポタンパク質およびアルブミンと、ポリリボヌクレオチドと、ポリサッカライドと、リン脂質と、アスベストとがある (Matsumoto, A. 他 (1990) *Proc. Natl. Acad. Sci.* 87:9133-9137; および Elomaa, O. 他 (1995) *Cell* 80:603-609)。スカベンジャー受容体が主要な役割を果たすと思われるのはアテローム発生においてであり、これは、動脈壁内の、

10

20

30

40

50

修飾されたLDLの取り込みを媒介して働き、また宿主防御においては、細菌性エンドトキシンと、細菌と、原虫とに結合して働く。

【0045】

T細胞受容体

T細胞は免疫系において作動体および調節因子としての二つの役割を果たしており、感染細胞で細胞死を誘発し、他の免疫細胞の増殖を刺激するシグナルの伝達と抗原認識を連関する。T細胞の集団は広範囲の異なった抗原を認識できるが、個々のT細胞は一つの抗原のみを、抗原提示細胞表面上の主要組織適合分子(MHC)との複合型ペプチドとしてT細胞受容体(TCR)に提示された場合のみに認識できる。大部分のT細胞上のTCRは免疫グロブリン様膜内在性糖タンパク質からなる。この免疫グロブリン様膜内在性糖タンパク質は、似た分子量の二つのポリペプチドサブユニットを含む。両方のTCRサブユニットは、可変領域と定常領域の両方を有する細胞外ドメイン、膜を1回貫通する膜貫通ドメイン、および短い細胞内ドメインを有する(Saito, H. 他 (1984) *Nature* 309:757-762)。TCRサブユニットの遺伝子は、異なった遺伝子区分の体細胞性再構成によって作成される。適切なMHC状況における抗原とTCRとの相互作用によって、免疫系の細胞要素の増殖、成熟および機能を誘発するシグナルカスケードが開始される(Weiss, A. (1991) *Annu. Rev. Genet.* 25: 487-510)。リンパ腫、白血病、自己免疫疾患および免疫不全疾患においてTCR遺伝子の再編成とTCR発現の変化が注目されている(Aisenberg, A.C. 他 (1985) *N. Engl. J. Med.* 313:529-533; Weiss, 前出)。 10

セレクチン

セレクチン(またはLEC-CAM)は主に炎症および白血球接着に関与する特殊なレクチンサブファミリーである(Lasky, L.A. (1991) *J. Cell. Biochem.* 45:139-146に概説されている)。セレクチンは循環系から急性の炎症部位に白血球を補充する過程を仲介し、サイトカインシグナル伝達に応答して血管内皮細胞の表面で発現する。セレクチンは白血球上の特定のリガンドに結合し、白血球が内皮表面に接着して表面に沿って移動することを可能にする。セレクチンがそのリガンドに結合すると、アクチン細胞骨格の分極性再構成をもたらし、白血球内のシグナル伝達を刺激する(Brenner, B. 他 (1997) *Biochem. Biophys. Res. Commun.* 231:802-807; Hidari, K. I. 他 (1997) *J. Biol. Chem.* 272:28750-28756)。セレクチンファミリーのメンバーは3つの特徴的なモチーフを持っている: すなわち、レクチンまたは糖質認識ドメイン、上皮成長因子様ドメイン、および異なった数の短いコンセンサスリピート(scrまたは「スシ」リピート)の3つである。補体調節タンパク質(CCP)モジュールあるいは、短いコンセンサスリピート(SCR)とも呼ばれるスシドメインは、多様な補体および接着タンパク質において見られる(Norman, D.G. 他 (1991) *J. Mol. Biol.* 219:717-725)。 20

【0046】

ネトリン受容体

ネトリンは拡散性誘引物質および忌避物質として働いて発達中の神経系内で遊走している細胞や軸索をその標的に誘導する分子のファミリーである。このネトリンの受容体には線虫タンパク質UNC-5および最近脊椎動物中に同定された相同体が含まれる(Leonardo, E.D. 他 (1997) *Nature* 386:833-838)。これらの受容体は免疫グロブリンスーパーファミリーのメンバーであり、またZU5ドメインと呼ばれる特徴的ドメインも含んでいる。ネトリン受容体ファミリーのマウスメンバーにおける突然変異であるRcm(吻側小脳形成異常)では、明らかな神経細胞異常遊走の結果である小脳および中脳欠損を生じている。 40

【0047】

VPS10ドメイン含有受容体

VPS10含有受容体ファミリーのメンバーはすべて、酵母の液胞ソーティングのタンパク質10(VPS10)受容体に相同性を持つドメインを含んでいる。このファミリーには、マウス胚性および出生後早期の神経系発達時に発現するSorCS、モザイク受容体SorLA、およびニューロテンシン受容体sortilinが含まれる(Hermey, G. 他 (1999) *Biochem. Biophys. Res. Commun.* 266:347-351; Hermey, G. 他 (2001) *Neuroreport* 12:29-32)。このファミリーの最近 50

同定されたメンバーであるSorCS2は発達中および成熟したマウスの中樞神経系に高度に発現される。その発現の主要部位は底板であり、ドーパミン性脳核および視床背側核を含む脳領域においても一過性に高度に検出される (Rezgaoui, M. (2001) *Mech. Dev.* 100:335-338)。

【0048】

膜結合タンパク質

テトラスパンファミリータンパク質

膜4回貫通型スーパーファミリー(TM4SF)、別名テトラスパンファミリーは多重遺伝子族であり、タイプIII膜内在性タンパク質をコードする (Wright, M.D. および Tomlinson, M.G. (1994) *Immunol. Today* 15:588)。このTM4SFは、細胞膜を4回貫通する膜タンパク質からなる。TM4SFのメンバーの例には、血小板および内皮細胞の膜タンパク質と、黒色腫関連抗原、白血球表面糖タンパク質、結腸癌腫抗原、腫瘍関連抗原、および住血吸虫表面タンパク質がある (Jankowski, S.A. (1994) *Oncogene* 9:1205-1211)。TM4SFの各メンバーは、約25~30%のアミノ酸配列同一性を互いに共有する。

10

多数のTM4SFメンバーが関与すると示唆されている作用は、シグナル伝達、細胞接着の制御、細胞の成長と増殖との調節 (例えば発生および発癌)、並びに細胞運動性 (例えば腫瘍細胞転移) である。TM4SFタンパク質の発現は多様な腫瘍に関連し、発現のレベルは、細胞が成長する際や活性化する際に改変され得る。

【0049】

腫瘍抗原

腫瘍抗原は表面分子であり、腫瘍細胞では正常細胞と異なる発現をする。腫瘍抗原は、腫瘍細胞を免疫的に正常細胞と区別し、ヒトの癌に関する診断と治療の標的を提供する (Takagi, S. 他 (1995) *Int. J. Cancer* 61: 706-715; Liu, E. 他 (1992) *Oncogene* 7: 1027-1032)。

20

【0050】

イオンチャネル

イオンチャネルは、事実上、身体のあらゆる細胞の原形質膜内に見られる。例えば、塩素イオンチャネルは多様な細胞機能を媒介し、その機能の例には、膜電位の調節と、上皮膜を通過するイオンの吸収と分泌とがある。塩素イオンチャネルが、ゴルジ体とエンドサイトーシス小胞の細胞内の膜にある場合、オルガネラのpHをも調節する (例えば Greger, R. (1988) *Annu. Rev. Physiol.* 50:111-122を参照)。塩素イオンチャネル類の電気生理学的および薬理学的特性、例えばイオンコンダクタンス、電流-膜電位関係、およびモジュレーター (修飾因子) 類への感受性は、筋肉、神経細胞、線維芽細胞、上皮細胞、および白血球に、別々の塩素イオンチャネルがあることを示唆する。多くのチャネルは、1個若しくは数個のプロテインキナーゼによるリン酸化部位を有する。プロテインキナーゼには、プロテインキナーゼA、プロテインキナーゼC、チロシンキナーゼ及びカゼインキナーゼIIがあり、これらはすべて細胞のイオンチャネル活性を調節する。細胞内におけるタンパク質の不適切なリン酸化は、細胞周期の進行および細胞分化における変化に関連する。細胞周期における変化は、アポトーシスの誘発や癌の誘発に関連することが示されている。細胞分化における変化は、生殖系、免疫系、および骨格筋の、疾患や障害に関連することが示されている。

30

40

【0051】

小脳顆粒神経細胞は非不活化カリウム電流を持ち、これは神経伝達物質からの受容体刺激による発火頻度を調節し、また、静止膜電位を制御する。非不活化電流を示すカリウムチャネル類の例に、ether a go-go (EAG) チャネルがある。KCR1と名付けられた膜タンパク質は、ラットEAGと特異的にそのC末端領域で結合し、小脳非不活化カリウム電流を調節する。KCR1は、12の膜貫通ドメイン群と、細胞内にアミノ酸末端およびカルボキシル末端を含むと予測される。これらの膜貫通領域の構造的特徴はトランスポータースーパーファミリーの領域に類似しているように見えるが、KCR1と既知のトランスポーターとの間に相同性は見られないことから、KCR1は新規クラスのトランスポーターに属することが示唆される。KC

50

R1は、非不活化カリウムチャネルの調節成分と考えられる (Hoshi, N. 他 (1998) J. Bio l. Chem. 273:23080-23085)。

【 0 0 5 2 】

A B C 輸送体

ATP結合カセット (ABC) 輸送体、別名「輸送ATPアーゼ」は膜タンパク質のスーパーファミリーであり、原核生物と真核生物とにおける輸送機能とチャネル機能とを媒介する (Higgins, C.F. (1992) Annu. Rev. Cell Biol. 8:67-113)。ABCタンパク質類は、類似した全体構造と、顕著な配列相同性とを共有する。すべてのABCタンパク質には約200アミノ酸残基の或る保存されたドメインがあり、これには1つ以上のヌクレオチド結合ドメインが含まれている。ABC輸送体遺伝子群における突然変異は、高ビリルビン血症II/Dubin-Johnson症候群、劣性スタルガルト病、X染色体連鎖副腎白質ジストロフィー、多剤耐性、小児脂肪便症、および嚢胞性線維症など、様々な疾患に関与する。

【 0 0 5 3 】

セマホリンとニューロピリン

セマフォリン (semaphorin) 類は軸索誘導分子の1大グループで少なくとも30種のメンバーがあり、脊椎動物、非脊椎動物、更には数種のウィルスでも発見されている。すべてのセマフォリンは長さが約500アミノ酸のセマ (sema) ドメインを有する。セマフォリン受容体であるニューロピリン (neuropilin) は、*in vitro*で神経突起の生成を促進することが示されている。ニューロピリンの細胞外領域はCUB、ジスコイジン (discoidin)、およびMAMドメインの3種のドメインからなる。ニューロピリンのCUBモチーフおよびMAMモチーフはタンパク質間相互作用において幾つかの役割を果たすことが示唆されており、更にセマドメインおよびC末端ドメインを介してセマフォリン類の結合に関与していると考えられる (Raper, J.A. (2000) Curr. Opin. Neurobiol. 10:88-94の概説を参照)。

【 0 0 5 4 】

細胞間伝達に関連する膜タンパク質

細胞間伝達は、多細胞生物の発生と生存とに必須である。細胞は、タンパク質シグナル伝達分子の分泌と取り込みとによって互いに通信する。細胞内へのタンパク質の取り込みを達成する手段がエンドサイトーシスであり、ここでは、シグナル伝達分子と原形質膜表面との相互作用が、しばしば特異的受容体への結合を通じて行われる結果、形質膜由来小胞を形成し、これらの小胞が分子を取り囲み、細胞質内に運ぶ。細胞からのタンパク質の分泌を達成する手段がエキソサイトーシスであり、ここでは、細胞内の分子が、トランスゴルジ網に由来する、膜に囲まれた輸送小胞内に詰められる。これらの小胞は原形質膜と融合し、小胞の内容物を周囲の細胞外空間に放出する。エンドサイトーシスとエキソサイトーシスとの結果、原形質膜成分の除去と追加とが起こり、また、これらの成分のリサイクルは、原形質膜と、膜に囲まれた内部区画との双方の、完全性、同一性、および機能性の維持に必須である。

【 0 0 5 5 】

Nogo は中枢神経系ミエリンの成分として同定され、成熟した脊椎動物における軸索再生を防ぐ。Nogo-66 受容体の切断および軸索表面からのグリコホスファチジルイノシトール結合タンパク質は神経細胞をNogo-66に対して非感受性にして中枢神経系障害からの回復を促進する (Fournier, A.E. 他 (2001) Nature 409:341-346)。

【 0 0 5 6 】

スリットタンパク質は神経系の腹側正中線の細胞によって発現される細胞外基質タンパク質である。スリットタンパク質は反発的ガイダンス受容体のRoundabout (Robo)のためのリガンドであり、したがって、反発的軸索誘導の役割を果たす (Brose, K. 他 (1999) Cell 96:795-806)。

【 0 0 5 7 】

リソソームは、自己貪食時の細胞内物質分解の部位であり、また、エンドサイトーシス後の細胞外分子の分解の場でもある。リソソーム酵素は、トランスゴルジ網から出芽する小胞内に詰められる。これらの小胞はエンドソームと融合して成熟リソソームを形成し、成

10

20

30

40

50

熟リソソーム内では、エンドサイトーシスによって取り込まれた物質の加水分解性消化が起こる。リソソームはオートファゴソームと融合し、独特な区画を形成し得る。この区画の中では、オルガネラや他の細胞内成分の分解が起こる。

【0058】

輸送小胞、例えばエンドソームによるタンパク質選別は、多様な生理的過程にとって重要な結果をもたらす。このような過程の例に、細胞表面成長、異なる細胞内オルガネラ群の生合成、エンドサイトーシス、および、ホルモンと神経伝達物質との制御された分泌がある (Rothman, J.E. および Wieland, F.T. (1996) *Science* 272:227-234)。特に、神経変性障害その他の神経病理には、エンドソームのタンパク質選別時、または、エンドソーム生合成の際の生化学的欠陥が関連する (Mayer R.J. 他 (1996) *Adv. Exp. Med. Biol.* 389:261-269)。

【0059】

ペルオキシソーム類は、分泌経路とは独立したオルガネラである。これらは、細胞内の、過酸化物を生成する多くの酸化反応の場である。ペルオキシソームは、真核細胞のオルガネラの中でもサイズや数が独特であり、また、酵素内容は、生物と細胞タイプと代謝の需要によって異なる (Waterham, H.R. および Cregg, J.M. (1997) *BioEssays* 19:57-66)。ペルオキシソームタンパク質に遺伝子欠損があるとペルオキシソーム欠損症を生じ、この欠損が関連付けられる多数のヒト病理の例に、ツェルヴェーガー症候群、近位肢型点状軟骨異形成 (rhizomelic chondrodysplasia punctata)、X染色体連鎖副腎白質ジストロフィー、アシルCoAオキシダーゼ欠損症、二頭酵素欠損症、古典的Refsum病、DHAPアルキルトランスフェラーゼ欠損症、および無カタラーゼ血症がある (Moser, H.W. および Moser, A.B. (1996) *Ann. NY Acad. Sci.* 804:427-441)。また、Gartner, J. 他 (1991; *Pediatr. Res.* 29:141-146) は、ツェルヴェーガー症候群の患者における低密度ペルオキシソーム様細胞内分画に関連する或る22 kDaの膜内在性タンパク質を発見した。

【0060】

正常な胚発生と、生殖細胞成熟の制御とを変調する多数の分泌タンパク質は、それぞれの膜結合受容体と相互作用する。胚発生時の細胞運命は、アクチビン (activin) / TGF-スーパーファミリーのメンバー、カドヘリン、IGF-2、および他のモルフォゲンによって決定される。また、生殖細胞と生殖組織との増殖、成熟、および再分化は、例えばIGF-2、インヒピン、アクチビン、およびフォリスタチン (follistatins) によって調節される (Petraglia, F. (1997) *Placenta* 18:3-8; Mather, J.P. 他 (1997) *Proc. Soc. Exp. Biol. Med.* 215:209-222)。トランスフォーミング成長因子 (TGFベータ) シグナル伝達は、II型TGFベータ受容体 (TbetaR-II) とリン酸化I型TGFベータ受容体 (TbetaR-I) の連続して作用する二つの受容体セリン/トレオニンキナーゼによって媒介される。TbetaR-I-結合タンパク質-1 (TRECAP-1) はI型トランスフォーミング成長因子受容体の静止形態と活性形態を識別し、TGFベータシグナル伝達に関連している (Charng, M.J 他 (1998) *J. Biol. Chem.* 273:93659368)。

【0061】

レチノイン酸受容体 (RAR) はレチノイン酸によって誘発される成熟を媒介し、骨髄の発達に關与する。顆粒球分化中にレチノイン酸によって誘発される遺伝子には、造血特異的遺伝子のE3が含まれているが、E3は骨髄造血時の活性RARの直接の標的となる。

【0062】

μ -オピオイド受容体 (MOR) は、モルフィン、コデイン、メサドンおよびフェンタニル並びにヘロインなどの鎮痛薬の作用を媒介する。MORはGタンパク質活性化カリウムチャンネルに機能的に結合する (Mestek A. 他 (1995) *J. Neurosci.* 15:2396-2406)。多様なMORサブタイプが存在する。ソマトスタチン2、ドパミンD2、プロスタグランジンEP3、およびセロトニン受容体サブタイプ 5-ヒドロキシトリプタミン4 および5ヒドロキシトリプタミン7を含む多数のGタンパク質共役受容体と同様にMOR-1で選択的スプライシングが見られる (Pan, Y.X. 他 (1999) *Mol. Pharm.* 56:396403)

膜表在性タンパク質および固着膜タンパク質

いくつかの膜タンパク質は膜貫通しておらず、膜アンカー(固着)または膜内在性タンパク質との相互作用を介して原形質膜に固着している。膜アンカーは翻訳後のあるタンパク質に共役的に結合しており、プレニル、ミリスチルおよびグリコシルホスファチジルイノシトール基のような部分が含まれる。膜表在性タンパク質および固着タンパク質の膜局在化は受容体媒介シグナル伝達などのプロセスにおける機能にとって重要である。例えば、Rasのプレニル化は原形質膜への局在化に必要であり、またシグナル伝達における正常作用および発ガン性作用に必要とされる。

【0063】

細胞外メッセンジャー

細胞間伝達は多細胞生物の成長と生存に必須である。特に内分泌系、神経系、免疫系の機能に必須である。更に細胞間伝達は組織作製と器官形成などの発生プロセスにとって極めて大切である。その場合細胞増殖、細胞分化および形態形成が空間的また時間的に正確な協働作用で調整されなければならない。細胞は、ホルモン、成長因子、神経ペプチド、およびサイトカインなどの多様なタイプのシグナル伝達分子の分泌および取り込みによって互いに伝達する。

10

【0064】

ホルモン

ホルモンはシグナル伝達分子であり、胚形成から成人まで基本的な生理的プロセスを協働して調整する。これらのプロセスには代謝、呼吸、生殖、排出、胎児組織分化、器官形成、成長、発達、恒常性およびストレス反応が含まれる。ホルモン分泌と神経系は密接に組み込まれており、互いに依存している。ホルモンは内分泌腺により、主に視床下部と脳下垂体、甲状腺と副甲状腺、膵臓、副腎、卵巣、精巣で分泌される。

20

【0065】

循環へのホルモンの分泌はきちんと制御されている。ホルモンはしばしば日中に、拍動性及び周期性パターンで分泌される。ホルモン分泌は血液生化学の摂動、他の上流で作用するホルモン、神経インパルス、負のフィードバックループにより調整される。血中ホルモン濃度は常に監視され、最適の安定したレベルを維持するように調整される。いったん分泌されると、ホルモンは特異的受容体を発現する標的細胞だけに作用する。

【0066】

多くの内分泌系の疾患は、ホルモンの分泌不全あるいは分泌過多により引き起こされる。分泌不全はオリジンのホルモンの腺が損傷したりあるいは正常に機能しない場合に多くの場合発生する。分泌過多は、ホルモン分泌細胞に由来する腫瘍の増殖に多くの場合起因する。不適切なホルモンレベルは、調節フィードバックループあるいはホルモン前駆体のプロセッシングにおける障害により引き起こされる場合がある。内分泌系機能障害は、標的細胞がホルモンに応答できない場合にも発生し得る。

30

【0067】

ホルモンは生化学的に、ポリペプチド、ステロイド、エイコサノイドまたアミンとして分類される。インスリンや成長ホルモンなど多様なホルモンを含むポリペプチドホルモンは、大きさや機能が様々である。多くの場合細胞内で成熟、活性化状態へ処理される不活性の前駆体として合成される。アミンホルモンにはエピネフリン、ドーパミンがあり、神経内分泌シグナリングで機能するアミノ酸誘導体である。ステロイドホルモンにはコレステロールに由来するホルモンエストロゲン、テストステロンがあり、性的成長と生殖において機能する。エイコサノイドホルモンにはプロスタグランジン、プロスタシクリンがあり、多様なプロセスで機能する脂肪酸誘導体である。多くのポリペプチドホルモンとある種のアミンホルモンは、循環で可溶性であるが、循環では分泌後数秒以内にタンパク質分解に対して感受性が高い。ステロイドホルモンとエイコサノイドホルモンは不溶性であり、循環ではキャリアタンパク質で輸送する必要がある。今からの記述では、主にポリペプチドホルモンに焦点を合わせる。

40

【0068】

視床下部と脳下垂体腺によって分泌されたホルモンは、神経信号に反応して他の内分泌腺

50

からのホルモン分泌と協調的に調節することにより内分泌機能において重要な役割を演じる。視床下部ホルモンには、甲状腺刺激ホルモン放出ホルモン、性腺刺激ホルモン放出ホルモン、ソマトスタチン、成長ホルモン放出因子、副腎皮質刺激ホルモン放出ホルモン、サブスタンスP、ドーパミン、プロラクチン放出ホルモンがある。これらのホルモンは、脳下垂体の前葉からのホルモン分泌を直接調節する。下垂体の前葉ホルモンによって分泌されたホルモンには、副腎皮質刺激ホルモン(ACTH)、メラニン細胞刺激ホルモン、成長ホルモンおよびプロラクチンなどのソマトトロピンホルモン、甲状腺刺激ホルモンおよび黄体形成ホルモン(LH)また卵胞刺激ホルモン(FSH)などの糖蛋白質ホルモン、 α -リポトロピン、 β -エンドルフィンがある。これらのホルモンは、甲状腺、膵臓、副腎からホルモン分泌を調節する。また排卵と精子形成を刺激して生殖器官に直接作用する。下垂体後葉は、抗利尿ホルモン(ADH, バソプレシン)とオキシトシンを合成、分泌する。 10

【0069】

視床下部と下垂体の疾患は多くの場合、原発脳腫瘍及び腺腫、妊娠性梗塞、下垂体切除、動脈瘤、血管奇形、血栓症、感染症、免疫異常、頭部外傷による合併症などの病変から起こる。このような疾患は他の内分泌腺の機能に重大な影響を及ぼす。下垂体低下に関連した疾患には、性機能低下及びシーハン症候群、尿崩症、カルマン病、カルマン症候群、ハンドシュラークリスチャン病、レトラシヴェ病、サルコイドーシス、エンブティセラ症候群、小人症が含まれる。下垂体機能亢進症に関連した疾患には、良性線種によって発生しやすい不适当抗利尿ホルモン(ADH)分泌症候群(SIADH)及び先端巨大症、巨人症が含まれる。 20

【0070】

甲状腺、副甲状腺によって分泌されるホルモンは、それぞれ主に代謝速度と血清カルシウムレベルの調製を制御する。甲状腺ホルモンには、カルシトニン、ソマトスタチン、甲状腺ホルモンが含まれる。副甲状腺は、副甲状腺ホルモンを分泌する。甲状腺機能低下症に関連した疾患には、甲状腺腫及び粘液水腫、細菌感染性急性甲状腺炎、ウイルス感染性亜急性甲状腺炎、自己免疫性甲状腺炎(橋本病)、クレチン病を含む。甲状腺機能亢進症に関連した疾患には、甲状腺中毒症及びその様々な型、グレース病、前脛骨粘液水腫、中毒性多結節性甲状腺腫、甲状腺癌、プランマー病を含む。副甲状腺機能亢進症に関連する疾患には、骨再吸収と副甲状腺肥大を招くConn病(chronic hypercalcemia)がある。 30

【0071】

膵臓によって分泌されるホルモンは、糖質、脂肪、タンパク質代謝の割合を調節して血糖値を調節する。膵臓ホルモンにはインスリン、グルカゴン、アミリン、 α -アミノ酪酸、ガストリン、ソマトスタチンおよび膵臓性ポリペプチドが含まれる。膵臓機能不全に関連する主な疾患は、インスリン活性の不足によって生じる真性糖尿病である。真性糖尿病は一般にI型(インスリン依存型、若年性糖尿病)、またはII型(非インスリン依存型成人性糖尿病)に分類される。インシュリン代償療法による両者の治療は公知である。真性糖尿は通常、低血糖症(インシュリンショック)、昏睡、糖尿病性ケトアシドーシス、乳酸アシドーシスなどの急性合併症を引き起こす。また目、腎臓、皮膚、骨、関節、循環系、神経系の疾患や感染への抵抗低下を招く慢性合併症も引き起こす。 40

【0072】

ホルモン機能に関連する解剖、生理機能、疾患については、McCance, K.L. 及び Huether, S.E. (1994) Pathophysiology: The Biological Basis for Disease in Adults and Children, Mosby-Year Book, Inc., St. Louis, MO; Greenspan, F.S. 及び Baxter, J.D. (1994) Basic and Clinical Endocrinology, Appleton and Lange, East Norwalk, CT.を参照されたい。

【0073】

成長因子

成長因子は、細胞間伝達を仲介する分泌性蛋白質である。循環系を通過して長距離を移動するホルモンとは異なり、ほとんどの成長因子は主に局所的なメディエータであり、隣接する細胞に作用する。ほとんどの成長因子は、成長因子を分泌経路に導く疎水性N末端シグ 50

ナルペプチド配列を含む。またほとんどの成長因子は分泌経路内で翻訳後修飾される。このような修飾には、蛋白質分解、糖鎖形成、リン酸化反応、分子内ジスルフィド結合形成が含まれることもある。一旦分泌されると、成長因子は隣接する標的細胞の表面上で特定の受容体に結合する。そして結合された受容体は細胞内シグナル伝達経路を誘発する。これらのシグナル伝達経路は、標的細胞で特異的細胞応答を引き起こす。これらの応答には遺伝子発現の調節と細胞分裂、細胞分化、細胞運動の刺激または阻害が含まれることもある。

【0074】

成長因子は、少なくとも2つのクラスに大きく、重複して分類される。一番大きなクラスには、作用が広範囲にわたる大きなポリペプチド成長因子が含まれる。これらの因子には上皮成長因子(EGF)、線維芽細胞成長因子(FGF)、トランスフォーミング成長因子-(TGF-)、インスリン様成長因子(IGF)、神経成長因子(NGF)、および血小板由来成長因子(PDGF)が含まれ、各々多数の関連因子のファミリーを示す。大きいポリペプチド成長因子はNGFを除いて、多様な細胞タイプ上で分裂促進因子として作用して創傷治癒、骨形成および再形成、細胞外基質合成並びに、上皮、表皮、結合組織の増殖を刺激する。TGF-、EGFおよびFGFファミリーのメンバーもまた、胚組織の分化において誘導的シグナルとして作用する。NGFは特に、神経細胞の成長および分化を促進する神経栄養性因子として機能する。

10

【0075】

別の成長因子のクラスには造血性成長因子が含まれており、標的特異性において狭い。造血性成長因子は、Bリンパ球、Tリンパ球、赤血球、血小板、好酸球、好塩基球、好中球、マクロファージ、及びそれらの幹細胞前駆体のような血液細胞の増殖及び分化を刺激する。造血性成長因子にはコロニー形成活性化因子(G-CSF、M-CSF、GM-CSF、CSF1-3)、赤血球生成促進因子、サイトカインが含まれる。サイトカインは免疫システムの細胞によって分泌される特殊な造血性因子であり、以下に詳述する。

20

【0076】

成長因子は、*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大な役割を果たす。大きなポリペプチド成長因子の過剰な発現は、培地において細胞の増殖と形質転換を促進する。*in vivo*の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貢献しうる。造血性成長因子の不適切な活性によって、貧血、白血病、及びリンパ腫が結果として起こる。更に成長因子は潜在的にプロトオンコジーンを発癌生成物である腫瘍性タンパク質に構造的、機能的に関連する。特定のFGFとPDGFファミリーメンバー-自体が腫瘍性タンパク質に相同であるが、EGF、NGF及びFGFファミリーの特定のメンバーの受容体はプロトオンコジーンによってコードされる。成長因子もプロトオンコジーン及び腫瘍抑制因子遺伝子の双方の転写調節に影響を与える。(Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor, MI; McKay, I.及び Leigh, I., 編集 (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY; Habenicht, A., 編集 (1990) Growth Factors, Differentiation Factors, and Cytokines, Springer-Verlag, New York, NY.)。

30

【0077】

更に、一部の大きなポリペプチド成長因子は、発達中の胚の始原胚葉の誘導において重要な役割を果たす。この誘導により究極的には胎児の中胚葉、外胚葉、内胚葉の形成がなされる。そして完全な成人の体のためのフレームワークを提供する。この誘導プロセスの崩壊は、胚発育にとって破局的となるであろう。

40

【0078】

小さなペプチド因子 - 神経ペプチド及び血管介在物質

神経ペプチド及び血管介在物質(NP/VM)は、通常20以下のアミノ酸である小さなペプチド因子のファミリーを構成する。これらの因子は通常はニューロンの興奮と血管収縮/血管拡張、筋肉収縮、脳や他の内分泌組織からのホルモン分泌の阻害において機能する。このファミリーに含まれるのは、神経ペプチド及び神経ペプチドホルモンとして、ボンベシン

50

、神経ペプチドY、ニューロテンシン、ニューロメディンN、メラノコルチン、オピオイド、ガラニン、ソマトスタチン、タキキニン、ウロテンシンII、及び平滑筋刺激ペプチド、バソプレッシンペプチド、血管作用性小腸ペプチドに含まれる関連ペプチド、また循環系 - borneシグナル伝達分子として、アンジオテンシン、補体、カルシトニン、エンドセリン、ホルミルメチオニルペプチド、グルカゴン、コレシストキニン、ガストリン及び上記で取り上げた多くのペプチドホルモンである。NP/VMは直接信号を伝達することが可能であり、別の神経伝達物質及びホルモンの活性若しくは遊離を調節し、またシグナル伝達カスケードにおいて触媒酵素として働く。NP/VMの効果は、ごく短時間のものから長期間持続するものまで幅広い (Martin, C.R. 他 (1985) Endocrine Physiology, Oxford University Press, New York, NY, 57-62ページの概説を参照)。

10

【0079】

サイトカイン

サイトカインは、免疫システム及び炎症反応を調節するシグナル分子のファミリーを構成する。サイトカインは、損傷あるいは感染に反応して通常白血球あるいは白血球細胞によって分泌される。サイトカインは、Bリンパ球、Tリンパ球、単球、マクロファージ、顆粒球などの免疫システムの細胞に主に作用する成長及び分化因子として働く。他のシグナル伝達分子のようにサイトカインは特異的原形質膜受容体に結合する。そして遺伝子発現パターンを変異する細胞内シグナル伝達経路を誘発する。炎症および免疫系疾患の治療におけるサイトカインの使用に関しては、かなりの可能性がある。

【0080】

サイトカインの構造と機能は、in vitroで広範囲にわたって特徴付けられてきた。多くのサイトカインは30Kダルトン以下の小さなポリペプチドである。50を超えるサイトカインがヒト及びげっし類から同定されている。サイトカインのサブファミリーの例としては、インターフェロン(IFN-、- および -)、インターロイキン(IL1-IL13)、腫瘍壊死因子(TNF- と -)およびケモカインがある。多くのサイトカインは組換えDNA技術を用いて産生されており、また個々のサイトカインの活性はin vitroで測定されている。これらの活性には白血球増殖、分化および運動の調節である。

20

【0081】

in vitroの個々のサイトカインの活性は、in vivoのサイトカインの活性の完全な範囲を反映していないかもしれない。サイトカインはin vivoでは個々に発現しないが、その代わりに、生物体が刺激を受けると種々の他のサイトカインと結合して発現する。ともにこれらのサイトカインは、その特異な刺激に適切な方法で集団的に免疫応答を調節する。よってサイトカインの生理学的活性は、刺激自体によって決定される。また相乗作用と拮抗関係の双方によって示されることもあると同時に、発現するサイトカイン中の複合体双方向ネットワークによっても決定される。

30

【0082】

ケモカインは、30を超えるメンバーでサイトカインサブファミリー - を構成する。(Wells, T. N. C. 及び Peitsch, M. C. (1997) J. Leukoc. Biol. 61:545-550 を参照)。ケモカインは、炎症の部位に単球とマクロファージを動員する走化タンパク質として始めに同定された。最近の証拠によると、造血及びHIV-1感染においてケモカインは重要な役割を演じる。ケモカインは、分子量約6-15Kダルトンの範囲の小さなタンパク質である。更にケモカインは、重要なシステイン残基の数と位置に基づいてC、CC、CXC、あるいはCX₃Cに分類される。例えばCCケモカインは2つの連続したシステインから成る保存されたモチーフをそれぞれ含む。それに続いてそれぞれ24-及び16-残基間隔の下流で起こる2つの付加的なシステインを含む。(ExpASY PROSITE database, documents PS00472 及び PDOC00434)。これら4つのシステイン残基の存在と間隔は高度に保存される。しかし介在している残基は著しく異なる。しかしシステインダブレットのおよそ15残基下流に位置するチロシンは保存されており、走化活性にとって重要であるように思われる。CCケモカインをコードする多くのヒト遺伝子は、染色体17上に集まっている。しかし他の所にマッピングされるCCケモカイン遺伝子の例が、少数ある。ケモカインには、他にもリンフォタク

40

50

チン (lymphotactin) (C ケモカイン)、マクロファージ走化性及び活性因子 (MCAF/MCP-1、CC ケモカイン)、血小板因子 4 及び IL-8 (CXC ケモカイン) そしてフラクタルカイン (fractalkine) また neurotractin (CX₃C ケモカイン) がある (Luster, A. D. (1998) N. Engl. J. Med. 338:436-445の概説を参照)。

【0083】

クロモグラニンおよびセクレトグラニンは、内分泌および神経内分泌細胞の分泌顆粒に存在する酸性タンパク質である (Huttner, W.B. 他 (1991) Trends Biochem. Sci. 16:2730) (Simon, J.-P. 他 (1989) Biochem. J. 262:1-13)。グラニンは生物活性ペプチドの前駆体か、またはペプチドホルモンやニューロペプチドの詰め込み時のヘルパータンパク質かも知れないが、これらの正確な役割は不明である。

10

【0084】

新規の受容体および膜結合タンパク質、またそれらをコードするポリヌクレオチドの発見により、新規の組成物を提供することで当分野の要望に応えることができる。この新規の組成物は、細胞増殖異常、自己免疫/炎症性の疾患、神経疾患、代謝異常、発達障害および内分泌疾患の診断・治療・予防において有用であり、また、受容体および膜結合タンパク質の核酸配列及びアミノ酸配列の発現における外来性化合物の影響についての評価にも有用である。

【発明の開示】

【発明の効果】

【0085】

本発明は、総称して「REMAP」、個別にはそれぞれ「CSAP-1」、「REMAP-2」、「REMAP-3」、「REMAP-4」、「REMAP-5」、「REMAP-6」、「REMAP-7」、「REMAP-8」、「REMAP-9」、「REMAP-10」、「REMAP-11」、「REMAP-12」、「REMAP-13」、「REMAP-14」および「REMAP-15」と呼ぶ受容体および膜結合タンパク質である精製されたポリペプチドを提供する。或る実施態様において本発明は、(a) SEQ ID NO:1-15を有する群から選択したアミノ酸配列からなるポリペプチド、(b) SEQ ID NO:1-15を有する群から選択したアミノ酸配列と少なくとも90%が同一性のある天然のアミノ酸配列を有するポリペプチド、(c) SEQ ID NO:1-15を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO:1-15を有する群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を含む群から選択した単離されたポリペプチドを提供する。一実施態様では、SEQ ID NO:1-15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。

20

30

【0086】

また、本発明は(a) SEQ ID NO:1-15からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO:1-15からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を持つ或る天然アミノ酸配列を有するポリペプチド、(c) SEQ ID NO:1-15からなる群から選択した或るアミノ酸配列を持つポリペプチドの生物学的活性断片、または(d) SEQ ID NO:1-15からなる群から選択した或るアミノ酸配列を持つポリヌクレオチドの免疫原性断片、からなる群から選択した或るポリペプチドをコードする、単離されたポリヌクレオチドを提供する。一実施態様では、該ポリヌクレオチドは、SEQ ID NO:1-15からなる群から選択した或るポリペプチドをコードする。別の実施態様では、ポリヌクレオチドはSEQ ID NO:16-30からなる群から選択される。

40

【0087】

本発明は更に、(a) SEQ ID NO:1-15からなる群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO:1-15からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO:1-15からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO:1-15からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群から選択したポリペプチドをコードするようなポリヌクレオチドと機能的に連結したプロモーター配列を有する組換えポリヌクレオチドを提供する。一実施態様では、本発

50

明は組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様では、本発明は組換えポリヌクレオチドを含む遺伝形質転換体を提供する。

【0088】

更に本発明は、(a) SEQ ID NO:1-15からなる群から選択されたアミノ酸配列からなるポリペプチドと、(b) SEQ ID NO:1-15からなる群から選択されたアミノ酸配列と90%以上の同一性を有する天然のアミノ酸配列からなるポリペプチドと、(c) SEQ ID NO:1-15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:1-15とからなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドを製造する方法を提供する。製造方法は、(a)組換えポリヌクレオチドを用いて形質転換した細胞をポリペプチドの発現に適した条件下で培養する過程と、(b)そのように発現したポリペプチドを回収する過程とを有し、組換えポリヌクレオチドはポリペプチドをコードするポリヌクレオチドに機能的に連結したプロモーター配列を有する。

10

【0089】

本発明は更に、(a) SEQ ID NO:1-15からなる群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO:1-15からなる群から選択したアミノ酸配列と少なくとも90%同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO:1-15からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO:1-15からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片から構成される群から選択されたポリペプチドに特異結合するような単離された抗体を提供する。

20

【0090】

本発明は更に、(a) SEQ ID NO:16-30からなる群から選択したポリヌクレオチド配列を有するポリヌクレオチド、(b) SEQ ID NO:16-30からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c) (a)に相補的なポリヌクレオチド配列、(d) (b)に相補的なポリヌクレオチド配列、および(e) (a)~(d)のRNA等価物からなる群から選択された単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドは少なくとも60の連続したヌクレオチドを有する。

【0091】

本発明は更に、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは(a) SEQ ID NO:16-30からなる群から選択したポリヌクレオチド配列を有するポリヌクレオチド、(b) SEQ ID NO:16-30からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を有するポリヌクレオチド、(c) (a)に相補的なポリヌクレオチド、(d) (b)に相補的なポリヌクレオチド、および(e) (a)~(d)のRNA等価物からなる群から選択されたポリヌクレオチドの配列を有する。検出方法は、(a)サンプル中の上記標的ポリヌクレオチドに相補的な或る配列からなる少なくとも20の連続したヌクレオチド群からなる或るプローブを用いて該サンプルをハイブリダイズする過程と、(b)該ハイブリダイゼーション複合体の有無を検出し、複合体が存在すればオブションでその量を検出する過程からなる。該プローブと該標的ポリヌクレオチドあるいはその断片との間でハイブリダイゼーション複合体が形成されるような条件下で、プローブは、該標的ポリヌクレオチドに対し特異的にハイブリダイズする。一実施態様では、プローブは少なくとも60の連続したヌクレオチドを含む。

30

40

【0092】

本発明はまた、サンプル中の標的ポリヌクレオチドを検出する一方法を提供する。ここで、標的ポリヌクレオチドは、(a) SEQ ID NO:16-30からなる群から選択したポリヌクレオチド配列を持つポリヌクレオチド、(b) SEQ ID NO:16-30からなる群から選択したポリヌクレオチド配列との少なくとも90%の同一性を有する天然ポリヌクレオチド配列を持つポリヌクレオチド、(c) (a)のポリヌクレオチドに相補的なポリヌクレオチド、

50

(d)(b)のポリヌクレオチドに相補的なポリヌクレオチド、および(e)(a)~(d)のRNA等価物、からなる群から選択した、或るポリヌクレオチドの配列を持つ。検出方法は、(a)ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、(b)前記の増幅した標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含む。

【0093】

本発明は更に、有効量のポリペプチドと薬剤として許容できる賦形剤とを含む組成物を提供する。有効量のポリペプチドは、(a)SEQ ID NO:1-15からなる群から選択したアミノ酸配列を含むポリペプチド、(b)SEQ ID NO:1-15からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c)SEQ ID NO:1-15からなる群から選択したアミノ酸配列を含むポリペプチドの生物学的活性断片、および(d)SEQ ID NO:1-15からなる群から選択したアミノ酸配列の免疫原性断片からなる群から選択される。一実施例では、SEQ ID NO:1-15からなる一群から選択されたアミノ酸配列を含む組成物を提供する。更に、本発明は、患者にこの組成物を投与することを含む、機能的REMAP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。

10

【0094】

本発明はまた、(a)SEQ ID NO:1-15からなる群から選択したアミノ酸配列を含むポリペプチド、(b)SEQ ID NO:1-15からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c)SEQ ID NO:1-15からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d)SEQ ID NO:1-15からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群から選択されたポリペプチドのアゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a)ポリペプチドを含むサンプルを化合物に曝す過程と、(b)サンプル中のア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REMAP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。

20

30

【0095】

本発明は更に、(a)SEQ ID NO:1-15からなる群から選択したアミノ酸配列を含むポリペプチド、(b)SEQ ID NO:1-15からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c)SEQ ID NO:1-15からなる群から選択したアミノ酸配列の生物学的活性断片、および(d)SEQ ID NO:1-15からなる群から選択したアミノ酸配列のポリペプチドの免疫抗原性断片からなる群から選択されたポリペプチドのアンタゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a)ポリペプチドを含むサンプルを化合物に曝す過程と、(b)サンプル中のアンタゴニスト活性を検出する過程とを含む。一実施態様で本発明は、この方法によって同定したアンタゴニスト化合物と薬剤として許容できる賦形剤とを含む組成物を提供する。更なる別法で本発明は、機能的REMAPの過剰発現を伴う疾患や症状の治療を要する患者への、この組成物の投与方法を提供する。

40

【0096】

本発明は更に、(a)SEQ ID NO:1-15からなる群から選択したアミノ酸配列を含むポリペプチド、(b)SEQ ID NO:1-15からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c)SEQ ID NO:1-15からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d)SEQ ID NO:1-15からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を含む群から選択されたポリペプチドに特異結合する化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a)ポリペプチドを適切な条件下で少なくとも1

50

つの試験化合物と混合させる過程と、(b)試験化合物とのポリペプチドの結合を検出し、それによってポリペプチドに特異結合する化合物を同定する過程とを含む。

【0097】

本発明は更に、(a)SEQ ID NO:1-15からなる群から選択したアミノ酸配列を含むポリペプチド、(b)SEQ ID NO:1-15からなる群から選択したアミノ酸配列と少なくとも90%の相同性を有する天然のアミノ酸配列を含むポリペプチド、(c)SEQ ID NO:1-15からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d)SEQ ID NO:1-15からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を含む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。スクリーニング方法は、(a)ポリペプチドの活性が許容された条件下で、ポリペプチドを少なくとも1つの試験化合物と混合させる過程と、(b)ポリペプチドの活性を試験化合物の存在下で算定する過程と、(c)試験化合物の存在下でのポリペプチドの活性を試験化合物の不存在下でのポリペプチドの活性と比較する過程とを含み、試験化合物の存在下でのポリペプチドの活性の変化は、ポリペプチドの活性を調節する化合物であることを意味する。

10

【0098】

更に本発明は、SEQ ID NO:16-30からなる群から選択されたポリヌクレオチド配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、(a)この標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、(b)この標的ポリヌクレオチドの発現の変化を検出するステップと、(c)化合物の存在下での異なる量の標的ポリヌクレオチドの発現を化合物の不存在下での発現と比較するステップを含む、該スクリーニング方法を提供する。

20

【0099】

本発明は更に、(a)核酸を含む生物学的サンプルを試験化合物で処理する過程と、(b)(i)SEQ ID NO:16-30からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(ii)SEQ ID NO:16-30からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(iii)(i)に相補的な配列を有するポリヌクレオチド、(iv)(ii)のポリヌクレオチドに相補的なポリヌクレオチド、(v)(i)~(iv)のRNA等価物からなる群から選択したポリヌクレオチドの少なくとも20の連続したヌクレオチドを含むプローブを用いて、処理した生物学的サンプルの核酸をハイブリダイズする過程を含む、試験化合物の毒性の算定方法を提供する。ハイブリダイゼーションは、上記プローブと生物学的サンプル中の標的ポリヌクレオチドの間に特定のハイブリダイゼーション複合体が形成されるような条件下で発生し、上記標的ポリヌクレオチドは、(i)SEQ ID NO:16-30からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(ii)SEQ ID NO:16-30からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(iii)(i)のポリヌクレオチドに相補的な配列を有するポリヌクレオチド、(iv)(ii)のポリヌクレオチドに相補的なポリヌクレオチド、および(v)(i)~(iv)のRNA等価物からなる群から選択する。或いは、標的ポリヌクレオチドは、上記(i)~(v)からなる群から選択したポリヌクレオチド配列の断片と、(c)ハイブリダイゼーション複合体の量を定量する過程と、(d)処理した生物学的サンプルのハイブリダイゼーション複合体の量を、非処理の生物学的サンプルのハイブリダイゼーション複合体の量と比較する過程を含み、処理した生物学的サンプルのハイブリダイゼーション複合体の量の差は、試験化合物の毒性を示す。

30

40

【0100】

(本発明の記載について)

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

50

とも併せて理解されたい。

【0101】

請求の範囲及び明細書中で用いている単数形の「或る」及び「その(この)」の表記は、文脈から明らかにそうでないとされる場合を除いて複数のもを指す場合もあることに注意しなければならない。従って、例えば「或る宿主細胞」と記されている場合にはそのような宿主細胞が複数あることもあり、「或る抗体」と記されている場合には単数または複数の抗体、及び、当業者に公知の抗体の等価物等についても言及しているのである。

【0102】

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

【0103】

(定義)

用語「REMAP」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REMAPのアミノ酸配列を指す。

【0104】

用語「アゴニスト」は、REMAPの生物学的活性を強めたり、模倣する分子を指す。このアゴニストは、REMAPに直接相互作用するか、或いはREMAPが関与する生物学的経路の成分と作用して、REMAP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。

【0105】

用語「対立遺伝子変異配列」は、REMAPをコードする遺伝子の別の形を指す。対立遺伝子変異体は、核酸配列における少なくとも1つの突然変異から作製し得る。また、変異mRNAまたはポリペプチドを作製し得る。その構造または機能は、変異することもしないこともある。遺伝子は、天然の対立遺伝子変異体を全く有しないか、1個若しくは数個の天然の対立遺伝子変異体を有し得る。一般に対立遺伝子変異体を生じさせる通常の突然変異性変化は、ヌクレオチドの自然欠失、付加または置換に帰するものである。これら各変化は、単独或いは他の変化と共に、所定の配列内で1回若しくは数回生じ得る。

【0106】

REMAPをコードする「変異(altered)」核酸配列は、種々のヌクレオチドを欠失、挿入または置換する核酸配列を有し、REMAPと同一またはREMAPの機能的特徴を少なくとも1つ有するポリペプチドを産出する。この定義には、REMAPをコードするポリヌクレオチド配列の正常な染色体の遺伝子座ではない位置での対立遺伝子変異配列との不適當あるいは予期しないハイブリダイゼーション、並びにREMAP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あるいは検出困難な多型性を含む。コードされたタンパク質も「変異」し得るものであり、サイレント変化を生ぜしめて結果的に機能的に等価なREMAPとなるようなアミノ酸残基の欠失、挿入または置換を含み得る。生物学的或いは免疫学的にREMAPの活性が保持される範囲で、残基の極性、電荷、溶解性、疎水性、親水性、及び/または両親媒性についての類似性に基いて意図的なアミノ酸置換が成され得る。例えば、負に帯電したアミノ酸にはアスパラギン酸及びグルタミン酸があり、正に帯電したアミノ酸にはリジン及びアルギニンがある。親水性値が近似している非荷電極性側鎖を有するアミノ酸には、アスパラギンとグルタミン、セリンとトレオニンがある。親水性値が近似している非荷電側鎖を有するアミノ酸には、ロイシン、イソロイシンおよびバリン、グリシンとアラニン、フェニルアラニンとチロシンがある。

10

20

30

40

50

【0107】

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

【0108】

「増幅」は、核酸配列の追加複製に関連する。増幅は通常、当業者によく知られたポリメラーゼ連鎖反応 (PCR) 技術を用いて行う。

【0109】

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

10

【0110】

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

20

【0111】

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

30

【0112】

用語「アプタマー」は、特定の分子標的に結合する核酸またはオリゴヌクレオチドを指す。アプタマーは *in vitro* の進化過程 (例えば米国特許番号第5,270,163号に記載されたSEL EX (Systematic Evolution of Ligands by EXponential Enrichmentの略、試験管内選択法)) から由来するもので、そのような過程は大きな組み合わせライブラリから標的特異的なアプタマー配列を選択する。アプタマー組成は、2本鎖または1本鎖であってもよく、デオキシリボヌクレオチド、リボヌクレオチド、ヌクレオチド誘導體、または他のヌクレオチド様分子を含み得る。アプタマーのヌクレオチド構成要素は修飾された糖基 (例えば、リボヌクレオチドの2'-OH基が2'-Fまたは2'-NH₂で置換されている) を有することが可能で、これらの糖基は、例えば、ヌクレアーゼに対する耐性あるいは血中でのより長い生存期間などの望ましい性質を改善する。循環系からアプタマーが除去される速度を遅くするために、アプタマーを高分子量キャリア等の分子に抱合させることができる。アプタマー類は、例えば或る架橋剤の光活性化によって、それらの同種リガンド群と特異的に架橋させ得る (Brody, E.N. および L. Gold (2000) J. Biotechnol. 74:5-13等を参照。)

40

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

【0113】

50

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

【0114】

「アンチセンス」の語は、特定の核酸配列の「センス」(コーディング)鎖と塩基対を形成することが可能な任意の組成物を指す。アンチセンス組成物の例には、DNAや、RNAや、ペプチド核酸(PNA)や、修飾されたバックボーン連鎖たとえばホスホロチオ酸、メチルホスホン酸またはベンジルホスホン酸などを有するオリゴヌクレオチドや、修飾された糖基たとえば2'-メトキシエチル糖または2'-メトキシエトキシ糖などを有するオリゴヌクレオチドや、あるいは修飾された塩基たとえば5-メチルシトシン、2-デオキシウラシルまたは7-デアザ-2'-デオキシグアノシンなどを有するオリゴヌクレオチドがあり得る。アンチセンス分子は、化学合成または転写を含む任意の方法で製造することができる。相補的アンチセンス分子は、ひとたび細胞に導入されたら、細胞が形成した天然の核酸配列と塩基対を形成し、転写または翻訳を阻止する二重鎖を形成する。「負」または「マイナス」という表現は、或る参考DNA分子のアンチセンス鎖を指すことがあり、「正」または「プラス」という表現は、或る参考DNA分子のセンス鎖を意味し得る。

10

【0115】

「生物学的に活性」の語は、天然分子の構造的機能、調節機能または生化学的機能を有するタンパク質を指す。同様に、用語「免疫学的に活性」または「免疫原性」は、天然或いは組換え体のREMAP、合成のREMAP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。

20

【0116】

「相補(的)」または「相補性」の語は、塩基対形成によってアニーリングする2つの一本鎖核酸の間の関係を指す。例えば、配列「5'A-G-T3'」は、相補配列「3'T-C-A5'」に結合する。

【0117】

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

30

【0118】

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

40

【0119】

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

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	10
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	20
Val	Ile, Leu, Thr	

【0120】

保存アミノ酸置換では通常、(a)置換領域におけるポリペプチドのバックボーン構造、例えば シートや ヘリックス構造、(b)置換部位における分子の電荷または疎水性、及び/または(c)側鎖の大部分が保持される。

【0121】

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

【0122】

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

【0123】

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

【0124】

「示差発現」は少なくとも2つの異なったサンプルを比較することによって決められる、増加(上方調節)、あるいは減少(下方調節)、または欠損遺伝子またはタンパク発現の欠損を指す。このような比較は例えば、治療後サンプルと未治療のサンプルまたは病態のサンプルと正常サンプルの間で行われ得る。

【0125】

「エキソンシャフリング」は、異なるコード領域(エキソン)の組換えを意味する。1つのエキソンがコードされたタンパク質の1つの構造的または機能的ドメインを代表し得るため、安定したサブストラクチャーを再分類することによって、新しいタンパク質が組立てられることが可能であり、新しいタンパク質機能の進化を促進できる。

【0126】

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

10

【0127】

SEQ ID NO:16-30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:16-30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:16-30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:16-30を関連ポリヌクレオチド配列から区別する類似の方法に有用である。SEQ ID NO:16-30の断片の正確な長さ及び断片に対応するSEQ ID NO:16-30の領域は、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

20

【0128】

SEQ ID NO:1-15のある断片は、SEQ ID NO:16-30のある断片によってコードされる。SEQ ID NO:1-15の断片には、SEQ ID NO:1-15を特異的に同定する固有のアミノ酸配列領域が含まれている。例えば、SEQ ID NO:1-15の断片は、SEQ ID NO:1-15を特異認識する抗体を産出するための免疫抗原性ペプチドとして有用である。SEQ ID NO:1-15の断片及び断片に対応するSEQ ID NO:1-15の領域の正確な長さは、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

【0129】

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

30

【0130】

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

【0131】

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

40

【0132】

ポリヌクレオチド配列間の一致率は、MEGALIGN version 3.12e配列アラインメントプログラムに組み込まれているようなCLUSTAL Vアルゴリズムのデフォルトのパラメータを用いて決定できる。このプログラムは、LASERGENE ソフトウェアパッケージ(一組の分子生物学的分析プログラム)(DNASTAR, Madison WI)の一部である。このCLUSTAL Vは、Higgins, D. G. 及び P.M. Sharp (1989) CABIOS 5:151-153、Higgins, D.G. 他 (1992) CABIOS 8:189-191に記載されている。ポリヌクレオチド配列をペアワイズでアラインメントする際のデフォルトパラメータは、Ktuple=2、gap penalty=5、window=4、「diagonals saved」=4と設定する。「重み付けされた」残基重み付け表が、デフォルトとして選択された。CLUSTAL Vは、アラインメントされたポリヌクレオチド配列対間の「類似率」として一致率を報

50

告する。

【0133】

あるいは、米国国立バイオテクノロジー情報センター（NCBI）のBasic Local Alignment Search Tool（BLAST）が一般的に用いられ、且つ、無料で入手可能な配列比較アルゴリズム一式を提供している（Altschul, S.F. 他（1990）J. Mol. Biol. 215:403-410）。このアルゴリズムは、幾つかの情報源から入手可能であり、メリーランド州ベセスダにあるNCBIおよびインターネット（<http://www.ncbi.nlm.nih.gov/BLAST/>）からも入手可能である。BLASTソフトウェア一式には様々な配列分析プログラムが含まれており、既知のポリヌクレオチド配列を種々のデータベースから得た別のポリヌクレオチド配列とアラインメントする「blastn」もその1つである。その他にも、2つのヌクレオチド配列をペアワイズで直接比較するために用いる「BLAST 2 Sequences」と称されるツールも利用可能である。「BLAST 2 Sequences」は、<http://www.ncbi.nlm.nih.gov/gorf/bl2.html>にアクセスして対話形式で利用することが可能である。「BLAST 2 Sequences」ツールは、blastn 及び blastp（以下に記載）の両方に用いることができる。BLASTプログラムは、一般的には、ギャップ及びデフォルト設定に設定された他のパラメータと共に用いる。例えば、2つのヌクレオチド配列を比較するために、デフォルトパラメータとして設定された「BLAST 2 Sequences」ツールVersion 2.0.12（2000年4月21日）を用いてblastnを実行してもよい。デフォルトパラメータの設定例を以下に示す。

10

【0134】

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

20

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

30

【0135】

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

【0136】

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

40

【0137】

ポリペプチド配列間の一致率は、MEGALIGN version 3.12e配列アラインメントプログラムに組み込まれているようなCLUSTAL Vアルゴリズムのデフォルトのパラメータを用いて決定できる（既に説明したのでそれを参照されたい）。CLUSTAL Vを用いて、ポリペプチド配

50

列をペアワイズでアラインメントする際のデフォルトパラメータは、Ktuple=1、gap penalty=3、window=5、「diagonals saved」=5と設定する。デフォルトの残基重み付け表としてPAM250マトリクスを選択する。ポリヌクレオチドアラインメントと同様に、CLUSTAL Vは、アラインメントされたポリペプチド配列対間の「類似率」として一致率を報告する。

【0138】

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

【0139】

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

一致率は、完全に画定された（例えば特定の配列番号で画定された）ポリペプチド配列の長さと比較して測定し得る。一致率は、配列或いは、より短い長さ、例えばより大きな画定されたポリペプチド配列から得られた断片（例えば少なくとも15、20、30、40、50、70または150の連続した残基の断片）の長さと比較して一致率を測定してもよい。ここに挙げた長さは単なる例示的なものに過ぎず、表、図及び配列リストを含めた本明細書に記載された配列に裏付けられた任意の配列長さの断片を用いて、或る長さであってその長さに対して一致率を測定し得る長さを説明し得ることを理解されたい。

【0140】

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

【0141】

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。

【0142】

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

【0143】

一般に、ハイブリダイゼーションのストリンジェンシーは或る程度、洗浄ステップを実行する温度を基準にして表すことができる。このような洗浄温度は通常、所定のイオン強度及びpHにおける特定の配列の融点（T_m）より約5～20℃低くなるように選択する。このT_mは、所定のイオン強度及びpHの下で、完全に一致するプローブに標的配列の50%がハイブリダイズする温度である。T_mを計算する式及び核酸のハイブリダイゼーション条件は

10

20

30

40

50

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

【0144】

本発明の、ポリヌクレオチドとポリヌクレオチドのハイブリダイゼーションに対する高ストリンジェンシー条件には、約0.2×SSC及び約1%のSDS存在下で約68℃において1時間の洗浄条件が含まれる。或いは、65℃、60℃、55℃または42℃の温度で行ってもよい。SSC濃度は、約0.1%のSDS存在下で、約0.1~2×SSCの範囲で変化し得る。通常は、ブロッキング剤を用いて非特異ハイブリダイゼーションを阻止する。このようなブロッキング剤には、例えば、約100~200 µg/mlの変性サケ精子DNAがある。特定条件下で、例えばRNAとDNAのハイブリダイゼーションに有機溶剤、例えば約35~50% v/vの濃度のホルムアミドを用いることもできる。洗浄条件の有用なバリエーションは、当業者には自明であろう。ハイブリダイゼーションは、特に高ストリンジェント条件下では、ヌクレオチド間の進化的な類似性を示唆し得る。このような類似性は、ヌクレオチド及びヌクレオチドにコードされるポリペプチドに対する類似の役割を強く示唆している。

10

【0145】

用語「ハイブリダイゼーション複合体」は、相補的な塩基間の水素結合によって、形成された2つの核酸配列の複合体を指す。ハイブリダイゼーション複合体は、溶解状態で形成し得る(C₀tまたはR₀t解析等)。或いは、一方の核酸配列が溶解状態で存在し、もう一方の核酸配列が固体支持体(例えば紙、膜、フィルタ、チップ、ピンまたはガラススライド、或いは他の適切な基板であって細胞若しくはその核酸が固定される基板)に固定されているような2つの核酸配列間に形成され得る。

20

【0146】

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。

【0147】

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

30

【0148】

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REMAPのポリペプチド断片またはオリゴペプチド断片を指す。用語「免疫原性断片」はまた、本明細書で開示するまたは当分野で周知のあらゆる抗体生産方法に有用なREMAPのポリペプチド断片またはオリゴペプチド断片を含む。

【0149】

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

【0150】

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。

40

【0151】

用語「調節」は、REMAPの活性の変化を指す。調節することによって例えば、REMAPのタンパク質活性、結合特性その他の生物学的、機能的または免疫学的特性が増大または低下し得る。

【0152】

「核酸」及び「核酸配列」の語は、ヌクレオチド、オリゴヌクレオチド、ポリヌクレオチドまたはこれらの断片を指す。「核酸」及び「核酸配列」の語は、ゲノム起源または合成起源のDNAまたはRNAであって一本鎖または二本鎖であるか或いはセンス鎖またはアンチセ

50

ンス鎖を表し得るようなDNAまたはRNAや、ペプチド核酸（PNA）や、任意のDNA様またはRNA様物質を指すこともある。

【0153】

「機能的に連結した」は、第1の核酸配列と第2の核酸配列が機能的な関係にある状態を指す。例えば、プロモーターがコード配列の転写または発現に影響を及ぼす場合には、そのプロモーターはそのコード配列に機能的に連結している。同一のリーディングフレーム内で2つのタンパク質コード領域を接続する必要がある場合、一般に、機能的に連結したDNA配列は非常に近接するか、或いは連続的に隣接し得る。

【0154】

「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のバックボーンに結合した、少なくとも約5ヌクレオチドの長さのオリゴヌクレオチドを含む、アンチセンス分子または抗遺伝子剤を指す。末端のリシンは、この組成物に溶解性を与える。PNAは、相補的一本鎖DNAまたはRNAに優先的に結合して転写の伸長を停止するものであり、ポリエチレングリコール化して細胞におけるPNAの寿命を延長し得る。

【0155】

REMAP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が含まれ得る。これらのプロセスは、合成或いは生化学的に生じ得る。生化学的修飾は、REMAPの酵素環境に依存し、細胞の種類によって異なり得る。

【0156】

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

【0157】

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

【0158】

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

【0159】

プライマーとして用いるオリゴヌクレオチドの選択は、そのような目的のために本技術分野でよく知られているソフトウェアを用いて行う。例えばOLIGO 4.06ソフトウェアは、各100ヌクレオチドまでのPCRプライマー対の選択に有用であり、オリゴヌクレオチド及

10

20

30

40

50

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

10

20

30

【0160】

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

【0161】

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

【0162】

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

40

【0163】

「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部分である。レポーター分子には、放射性核種、酵素、蛍光剤、化学発光剤、発色剤、基質、補助因子、阻害因子、磁気粒子及びその他の当分野で既知の成分がある。

【0164】

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

【0165】

50

用語「サンプル」は、その最も広い意味で用いられている。REMAP、REMAP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。

【0166】

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

10

【0167】

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

【0168】

「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。

【0169】

用語「基板」は、任意の好適な固体或いは半固体の支持物を指し、膜及びフィルター、チップ、スライド、ウエハ、ファイバー、磁気または非磁気ビード、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。基板は、壁、溝、ピン、チャネル、孔等、様々な表面形態を有することができ、基板表面にはポリヌクレオチドやポリペプチドが結合する。

20

【0170】

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

【0171】

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

30

【0172】

ここで用いる「遺伝形質転換体」とは任意の生物体であり、限定するものではないが動植物を含み、生物体の1個若しくは数個の細胞が、ヒトの関与によって、例えば本技術分野でよく知られている形質転換技術によって導入された異種核酸を有する。核酸の細胞への導入は、直接または間接的に、細胞の前駆物質に導入することによって、計画的な遺伝子操作によって、例えば微量注射法によって或いは組換えウイルスの導入によって行う。遺伝子操作の語は、古典的な交雑育種或いは試験管内受精を指すものではなく、組換えDNA分子の導入を指すものである。本発明に基づいて予期される遺伝形質転換体には、バクテリア、シアノバクテリア、真菌及び動植物がある。本発明の単離されたDNAは、本技術分野で知られている方法、例えば感染、形質移入、形質転換またはトランス接合（transconjugation）によって宿主に導入することができる。本発明のDNAをこのような生物体に移入する技術はよく知られており、前出のSambrook 他（1989）等の参考文献に与えられて

40

50

いる。

【0173】

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

10

【0174】

特定のポリペプチド配列の「変異体」は、ポリペプチド配列の1本の長さ全体で特定のポリペプチド配列に対して少なくとも40%の相同性を有するポリペプチド配列として画定される。ここで、デフォルトパラメータに設定した「BLAST 2 Sequences」ツールVersion 2.0.9(1999年5月7日)を用いてblastpを実行する。このようなポリペプチド対は、所定の長さに対して、例えば少なくとも50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の配列同一性を示し得る。

20

【0175】

(発明)

本発明は新規の受容体および膜結合タンパク質(REMAP)の発見、REMAPをコードするポリヌクレオチド、および細胞増殖異常、自己免疫/炎症疾患神経疾患、代謝疾患、発達障害および内分泌疾患の診断、治療、ならびに予防に対するこれらの組成の使用に基づいている。

30

【0176】

表1は、本発明の完全長ポリヌクレオチド配列及びポリペプチド配列の命名の概略である。各ポリヌクレオチド及びその対応するポリペプチドは、1つのIncyteプロジェクト識別番号(IncyteプロジェクトID)と相関する。各ポリペプチド配列は、ポリペプチド配列識別番号(ポリペプチドSEQ ID NO)とIncyteポリペプチド配列番号(IncyteポリペプチドID)によって表示した。各ポリヌクレオチド配列は、ポリヌクレオチド配列識別番号(ポリヌクレオチドSEQ ID NO)とIncyteポリヌクレオチドコンセンサス配列番号(IncyteポリヌクレオチドID)によって表示した。

40

【0177】

表2は、GenBankタンパク質(genpept)データベースに対するBLAST分析によって同定されたような、本発明のポリペプチドとの相同性を有する配列を示している。列1および列2はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号(ポリペプチドSEQ ID NO:)およびそれに対応するIncyteポリペプチド配列番号(IncyteポリペプチドID)を示す。列3は、GenBankの最も近い相同体のGenBankの識別番号(GenBank ID NO:)を示す。列4は、各ポリペプチドとその相同体との間の一致を表す確率スコアを示す。列5は、該当箇所には適当な引用を示すとともにGenBank相同体1つ以上の注釈(annotation)を示し、これらはすべて本明細書では参考文献に含まれる。

【0178】

50

表 3 は、本発明のポリペプチドの様々な構造的特徴を示す。列 1 および列 2 はそれぞれ、本発明の各ポリペプチドのポリペプチド配列識別番号 (SEQ ID NO :) およびそれに対応する Incyte ポリペプチド配列番号 (Incyte ポリペプチド ID) を示す。列 3 は、各ポリペプチドのアミノ酸残基数を示す。列 4 および列 5 はそれぞれ、GCG配列分析ソフトウェアパッケージの MOTIFS プログラム (Genetics Computer Group, Madison WI) によって決定された、リン酸化およびグリコシル化の可能性のある部位を示す。列 6 は、シグネチャ配列、ドメイン、およびモチーフを含むアミノ酸残基を示す。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。

【 0 1 7 9 】

表 2 及び表 3 は共に、本発明の各々のポリペプチドの特性を要約しており、それらの特性が請求の範囲に記載されたポリペプチドが受容体および膜結合タンパク質であることを確立している。例えば、SEQ ID NO:1 は、ラットレチノイン酸受容体 2 イソ型 (GenBank ID g3213188) と 9 7 % の同一性があり、これは Basic Local Alignment Search Tool (BLAST) で確認された (表 2 参照) 。 BLAST 確率スコアは $8.5e-245$ であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。 SEQ ID NO:1 はまた、核ホルモン受容体リガンド結合ドメインと C 4 タイプ亜鉛フィンガーを有するが、これは、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリードメインの PFAM データベースにおいて、統計的に有意な一致を検索して決定された。 (表 3 参照) BLIMPS、MOTIFS、及び PROFILESCAN 解析よりのデータは、SEQ ID NO:1 がレチノイン酸受容体である、さらに実証的な証拠を提供する。別の実施例において、SEQ ID NO:2 はヒト多数回膜貫通受容体 TRC 8 (GenBank ID g3395787) に残基 A3 から残基 H575 まで 3 1 % の同一性を有することが $9.7e-61$ の確率スコアで BLAST によって確認された (表 2 参照) 。 BLAST 解析よりのデータは、SEQ ID NO:2 が多数回膜貫通受容体である、さらに実証的な証拠を提供する。別の実施例において、SEQ ID NO:5 はラットの潜在的リガンド結合タンパク質 (GenBank ID g57734) に残基 V4 から残基 A4 7 9 までに、7 6 % の同一性を有することが、 $6.3e-187$ の確率スコアで BLAST によって確認された (表 2 参照) 。 さらに、BLAST 解析よりのデータは、SEQ ID NO:5 が嗅覚リガンド結合タンパク質である、さらに実証的な証拠を提供する。別の実施例において、SEQ ID NO:11 はイヌの嗅覚受容体 Cf0LF2 (GenBank ID g1314663) に残基 M1 から F310 までに、77% の同一性を有することが、 $9.7e-129$ の確率スコアで BLAST によって確認された (表 2 参照) 。 SEQ ID NO:11 はまた、膜 7 回貫通受容体ドメイン (ロドプシンファミリー) を含むが、これは HMM に基づく PFAM データベースの統計的に有意な一致を検して決定された。 (表 3 参照) BLIMPS、MOTIFS 及び PROFILESCAN 分析から得たデータは、SEQ ID NO:11 が G タンパク質共役受容体であることを裏づける証拠を更に提供する。別の実施例において SEQ ID NO:15 は潜在的なりガンド結合タンパク質 RYA3 (GenBank ID g11877275) のオルソログであるヒトタンパク質に対して残基 M5 から残基 M328 までに 99% の同一性を有し、また残基 V313 から残基 E410 までに 89% に同一性を有することが $3.3e-207$ の確率スコアで BLAST によって確認された (表 2 参照) 。 BLIMPS 及び追加の BLAST 解析よりのデータは、SEQ ID NO:15 がリガンド結合タンパク質である、さらに実証的な証拠を提供する。 SEQ ID NO:3 - 4、SEQ ID NO:6 - 10、および SEQ ID NO:12 - 14 は同じような方法で解析し、注釈を付けた。 SEQ ID NO:1-15 の解析のためのアルゴリズム及びパラメータを表 7 に記述する。

【 0 1 8 0 】

表 4 に示すように、本発明の完全長ポリヌクレオチド配列は、cDNA 配列、またはゲノム DNA 由来のコード (エキソン) 配列、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 は本発明の各ポリヌクレオチドに対するポリヌクレオチド配列識別番号 (ポリヌクレオチド SEQ ID NO) および対応する Incyte ポリヌクレオチドコンセンサス配列番号 (Incyte ID)、および塩基対の各ポリヌクレオチド配列の長さを示している。列 2 は、本発明の完全長ポリヌクレオチド配列の構築、並びに、例えば、SEQ ID NO:16-30 を同定する、あるいは SEQ ID NO:16-30 と関連するポリヌクレオチド配列を区別す

10

20

30

40

50

る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のcDNA配列および/またはゲノム配列の開始ヌクレオチド(5')位置および終了ヌクレオチド(3')位置を示す。

【0181】

表4の列2で記述されたポリヌクレオチド断片は、具体的にはたとえば組織特異的なcDNAライブラリまたはプールされたcDNAライブラリに由来するIncyte cDNAを指す場合がある。或いは列2の識別番号は、完全長ポリヌクレオチド配列のアセンブリに寄与するGenBank cDNAまたはESTを指す場合もある。さらに、列2のポリヌクレオチド断片は、ENSEMBL(The Sanger Centre、英国ケンブリッジ)データベースから由来した配列(即ち「ENST」命名を含む配列)を同定し得る。或いは、列2で記述されたポリヌクレオチド断片は、NCBI RefSeq Nucleotide Sequence Records データベースから由来する場合もあり(即ち「NM」または「NT」の命名を含む配列)、またNCBI RefSeq Protein Sequence Recordsから由来する場合もある(即ち「NP」の命名を含む配列)。または列で記述されたポリヌクレオチド断片は、「エキソンスティッチング(exon-stitching)」アルゴリズムにより結び合わせたcDNA及びGenscan予想エキソンの両方からなる群を意味する場合がある。例えば、FL_XXXXXX_N₁_N₂_YYYYY_N₃_N₄ として同定されるポリヌクレオチド配列はアルゴリズムが適用される配列のクラスターの識別番号がXXXXXXであり、アルゴリズムにより生成される予測の番号がYYYYYであり、(もし存在すれば)N_{1,2,3}が解析中に手で編集された可能性のある特定のエキソンであるような「縫合された」配列である(実施例5参照)。または、列2のポリヌクレオチド断片は「エキソンストレッチング(exon-stretching)」アルゴリズムにより結び合わせたエキソンの集合を指す場合もある。例えば、FLXXXXXX_gAAAAA_gBBBBB_1_N として同定されるポリヌクレオチド配列は「ストレッチ」配列である。ここでXXXXXXはIncyteプロジェクト識別番号、gAAAAAは「エキソンストレッチング」アルゴリズムを適用したヒトゲノム配列のGenBank識別番号、gBBBBBは一番近いGenBankタンパク質相同体のGenBank識別番号またはNCBI RefSeq 識別番号である。(実施例5を参照。)RefSeq 配列が「エキソンストレッチング」アルゴリズムのタンパク質相同体として使われた場合は、RefSeq 識別子(「NM」、「NP」または「NT」と表示された)はGenBank 識別子(すなわち、gBBBBB)の代わりに使われる場合もある。

【0182】

あるいは、接頭コードは手で編集されたか、ゲノムDNA配列から予測されたか、または配列解析方法の組み合わせから由来している構成配列を識別する。次の表は構成配列の接頭コードと、接頭コードと関連する同じ配列の分析方法の例を列記する(実施例4と5を参照)。

接頭コード	解析のタイプまたはプログラムの例
GNN、GFG、ENST	例えば、GENSCAN (Stanford University, CA, USA) または FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK)を用いたゲノム配列からのエキソン予測
GBI	手で編集されたゲノム配列の解析
FL	スティッチまたはストレッチゲノム配列(実施例5参照)
INCY	ゲノムへのEST配列のマッピングからの完全長転写とエキソン予測エキソンと転写を予想するために、ゲノム位置とEST構成データが組み合わせられる。

【0183】

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

【0184】

表5はIncyte cDNA配列を用いて構築された完全長ポリヌクレオチド配列のための代表的なcDNAライブラリを示している。代表的なcDNAライブラリは、上記のポリヌクレオチド配列を構築及び確認するために用いられるIncyte cDNA配列によって最も頻繁に代表されるIncyte cDNAライブラリである。cDNAライブラリを作製するために用いた組織及びベクターを表5に示し、表6で説明している。

【0185】

本発明はまた、REMAPの変異体も含む。好適なREMAP変異体は、REMAPの機能的或いは構造的特徴の少なくとも1つを有し、かつ、REMAP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変異体である。

【0186】

本発明はまた、REMAPをコードするポリヌクレオチド群をも含む。特定の実施例において、本発明は、REMAPをコードするSEQ ID NO:16-30からなる群から選択した配列を含むポリヌクレオチド配列を含む。SEQ ID NO:16-30のポリヌクレオチド配列には、配列表に示したように等価RNA配列をも含むが、そこでは窒素塩基チミンの出現はウラシルに置換され、糖のバックボーンはデオキシリボースではなくリボースで構成される。

【0187】

本発明はまた、REMAPをコードするポリヌクレオチド配列の変異配列を含む。詳細には、このようなポリヌクレオチド配列の変異配列は、REMAPをコードするポリヌクレオチド配列に対して少なくとも約70%のポリヌクレオチド配列同一性、或いは少なくとも約85%のポリヌクレオチド配列同一性、更には少なくとも約95%ものポリヌクレオチド配列同一性を有する。本発明の特定の実施形態は、SEQ ID NO:16-30からなる群から選択された核酸配列と少なくとも約70%のポリヌクレオチド配列同一性、或いは少なくとも約85%のポリヌクレオチド配列同一性、更には少なくとも約95%ものポリヌクレオチド配列同一性を有するSEQ ID NO:16-30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配列を提供する。上記のポリヌクレオチド変異配列は何れも、REMAP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し得る。

【0188】

更に別の例では、本発明の或るポリヌクレオチド変異配列は、REMAPをコードするポリヌクレオチド配列のスプライス変異配列である。スプライス変異体はREMAPをコードするポリヌクレオチド配列と有意な相同性を有し得るが、mRNAプロセッシング中のエキソンの選択的スプライシングによる配列ブロックの追加または削除により、通常、ポリヌクレオチドがより多くまたはより少数の塩基を有することになる。或るスプライス変異配列には、約70%未満、または約60%未満、あるいは約50%未満のポリヌクレオチド配列同一性が、REMAPをコードするポリヌクレオチド配列との間で全長に渡って見られるが、このスプライス変異配列の部分は、REMAPをコードするポリヌクレオチド配列の部分と、少なくとも約70%、あるいは少なくとも約85%、または少なくとも約95%、なおまたは100%の、ポリヌクレオチド配列同一性を有することとなる。例えば、SEQ ID:30の配列を含むポリヌクレオチドはSEQ ID:20の配列を含むポリヌクレオチドのスプライス変異体である。前記のスプライス変異体のいずれもが、REMAPの少なくとも一つの機能的特徴か、または構造的特徴を含むアミノ酸配列をコードできる。

【0189】

遺伝暗号の縮重により作り出され得るREMAPをコードする種々のポリヌクレオチド配列には、自然発生する任意の既知の遺伝子のポリヌクレオチド配列と最小の類似性しか有しないものも含まれることを、当業者は理解するであろう。したがって本発明は、可能コドン選択に基づく組合せの選択によって産出し得るようなありとあらゆる可能性のあるポリヌ

10

20

30

40

50

クレオチド配列変異体を網羅し得る。これらの組み合わせは、天然のREMAP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。

【0190】

REMAP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、適切に選択されたストリンジェントな条件下で、天然のREMAPのヌクレオチド配列とハイブリダイズ可能であるが、例えば、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REMAP或いはその誘導体をコードするヌクレオチド配列を作るとは有利となり得る。宿主が特定のコードンを利用する頻度に基づいて、特定の真核又は原核宿主に発生するペプチドの発現率を高めるようにコードンを選択することが可能である。コードされたアミノ酸配列を変えないで、REMAP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

【0191】

本発明はまた、REMAP及びその誘導体をコードするDNA配列またはそれらの断片を完全に合成化学によって作り出すことも含む。作製後、当分野で周知の試薬を用いて、この合成配列を任意の様々な入手可能な発現ベクター及び細胞系中に挿入し得る。更に、合成化学を用いてREMAPまたはその任意の断片をコードする或る配列に突然変異を誘導し得る。

【0192】

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:16-30及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる(例えば、Wahl, G.M.及びS.L. Berger (1987) *Methods Enzymol.* 152:399-407、Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511等を参照)。アニーリング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

【0193】

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

30

40

【0194】

当分野で周知のPCR法をベースにした種々の方法を用いて部分的なヌクレオチド配列を利用して、REMAPをコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば、使用し得る方法の1つである制限部位PCR法は、ユニバーサルプライマー及びネステッドプライマーを用いてクローニングベクター内のゲノムDNAから未知の配列を増幅する方法である(例えば、Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322を参照。)別の方法に逆PCR法があり、これは広範な方向に伸長させたプライマーを用いて環状化した鑄型から未知の配列を増幅する方法である。鑄型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限酵素断片から得る(例えば、Triglia, T. 他

50

(1988) *Nucleic Acids Res.* 16:8186。)第3の方法としてキャプチャPCR法があり、これはヒト及び酵母菌人工染色体DNAの既知の配列に隣接するDNA断片をPCR増幅する方法に関与している。(Lagerstrom, M.他(1991) *PCR Methods Applic* 1:111-119等を参照)。この方法では、PCRを行う前に複数の制限酵素の消化及びライゲーション反応を用いて未知の配列領域内に組換え二本鎖配列を挿入することが可能である。また、未知の配列を検索するために用い得る別の方法については当分野で知られている(Parker, J.D.他(1991) *Nucleic Acids Res.* 19:3055-3060等を参照)。更に、PCR、ネステッドプライマー及びPromoterFinder(商標)ライブラリ(Clontech, Palo Alto CA)を用いてゲノムDNAをウォーキングすることができる。この手順は、ライブラリをスクリーニングする必要がなく、イントロン/エキソン接合部を見付けるのに有用である。全てのPCRベースの方法に対して、市販されているソフトウェア、例えばOLIGO 4.06プライマー分析ソフトウェア(National Biosciences, Plymouth MN)或いは別の好適なプログラムを用いて、長さが約22~30ヌクレオチド、GC含有率が約50%以上、温度約68~72で鋳型に対してアニーリングするようにプライマーを設計し得る。

10

【0195】

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

20

【0196】

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

30

【0197】

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

【0198】

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

40

【0199】

本発明のヌクレオチドは、MOLECULARBREEDING(Maxygen Inc., Santa Clara CA。米国特許第5,837,458号、Chang, C.-C.他(1999) *Nat. Biotechnol.* 17:793-797、Christians, F.C.他(1999) *Nat. Biotechnol.* 17:259-264、Cramer, A.他(1996) *Nat. Biotechnol.* 14:315-319に記載)等のDNAシャッフリング技術の対象となり、REMAPの生物学的特性、

50

例えば生物活性、酵素力、或いは他の分子や化合物との結合力等を変更または向上させ得る。DNAシャッフリングは、遺伝子断片のPCRを介する組換えを用いて遺伝子変異体のライブラリを作製するプロセスである。ライブラリはその後、その遺伝子変異体を所望の特性に同定するような選択またはスクリーニングにかける。次にこれらの好適な変異体をプールし、更に反復してDNAシャッフリング及び選択/スクリーニングを行ってもよい。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、ランダムポイント突然変異を有する単一の遺伝子の断片を組み換えて、スクリーニングし、その後所望の特性が最適化されるまでシャッフリングすることができる。或いは、所定の遺伝子を同種または異種のいずれかから得た同一遺伝子ファミリーの相同遺伝子と組み換え、それによって天然に存在する複数の遺伝子の遺伝多様性を、指示された制御可能な方法で最大化させることができる。

10

【0200】

別の実施例によれば、REMAPをコードする配列は、当分野で周知の化学的方法を用いて、全体或いは一部が合成可能である（例えば、Caruthers, M.H.ら（1980）Nucl. Acids Res. Symp. Ser 7:215-223; 及びHorn, T.他（1980）Nucl. Acids Res. Symp. Ser.225-232を参照）。別法として、化学的方法を用いて、REMAP 自体またはその断片を合成し得る。例えば、種々の液相または固相技術を用いてペプチド合成を行うことができる（Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, 55-60ページ、Roberge, J.Y.他（1995）Science 269:202204等を参照）。自動合成はABI 431Aペプチドシンセサイザ（Applied Biosystems）を用いて達成し得る。更にREMAP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改変により、および/または他のタンパク質群または任意のその一部からの配列群との組み合わせにより、変異配列ポリペプチドを作製、または、天然ポリペプチドの配列を有するポリペプチドを作製することが可能である。

20

【0201】

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

【0202】

生物学的に活性なREMAPを発現させるために、REMAP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し得る。好適な発現ベクターとは、好適な宿主に挿入されたコーディング配列の転写および翻訳の制御に必要なエレメント群を持つベクターである。これらのエレメントの例には、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、ベクター及びREMAPをコードするポリヌクレオチド配列群における調節配列が含まれる。このような要素は、長さ及び特異性が様々である。特定の開始シグナルによって、REMAP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開始コドンと、コザック配列などの近傍の配列が含まれる。REMAP をコードする配列、およびその開始コドンや、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、更なる転写制御シグナルや翻訳制御シグナルは必要ないこともある。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含まれるようにすべきである。外来性の翻訳要素及び開始コドンは、様々な天然物及び合成物を起源とし得る。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である（例えば、Scharf, D. 他（1994）Results Probl. Cell Differ. 201 - 18-162.を参照）。

30

40

【0203】

当業者に周知の方法を用いて、REMAP 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.

50

他。(1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, 4章及び8章, および16-17章; およびAusubel, F.M. 他。(1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9章、13章および16章を参照)。

【0204】

種々の発現ベクター/宿主系を利用して、REMAPをコードする配列の保持及び発現が可能である。限定するものではないがこのような発現ベクター/宿主系には、組換えバクテリオファージ、プラスミドまたはコスミドDNA発現ベクターで形質転換させた細菌などの微生物等や、酵母菌発現ベクターで形質転換させた酵母菌や、ウイルス発現ベクター(例えばバキュロウイルス)に感染した昆虫細胞系や、ウイルス発現ベクター(例えばカリフラワーモザイクウイルス、CaMVまたはタバコモザイクウイルスTMV)または細菌発現ベクター(例えばTiまたはpBR322プラスミド)で形質転換させた植物細胞系、動物細胞系がある。(前出のSambrook、前出のAusubel、Van Heeke, G. 及び S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509、; Engelhard, E.K. 他 (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227、Sandig, V. 他 (1996) Hum. Gene Ther. 7:1937-1945、Takamatsu, N. (1987) EMBOJ. 6:307-311、『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY、191-196ページ、Logan, J. 及びT. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659、Harrington, J.J. 他 (1997) Nat. Genet. 15:345-355等を参照)レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ輸送することができる(Di Nicola, M. 他 (1998) Cancer Gen. Ther. 5(6):350-356、Yu, M. 他 (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344、Buller, R.M. 他 (1985) Nature 317(6040):813-815; McGregor, D.P. 他 (1994) Mol. Immunol. 31(3):219-226、Verma, I.M. 及び N. Somia (1997) Nature 389:239-242等を参照)。本発明は使用される宿主細胞によって限定されるものではない。

10

20

【0205】

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、REMAP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、REMAPをコードするポリヌクレオチド配列の慣例的なクローニング、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 (Life Technologies) などの多機能の大腸菌ベクターを用いることができる。ベクターのマルチクローニング部位にREMAPをコードする配列をライゲーションするとlacZ 遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。更にこれらのベクターは、クローニングされた配列における *in vitro* 転写、ジデオキシのシーケンシング、ヘルパーファージによる一本鎖のレスキュー、入れ子状態の欠失の生成にも有用であろう(例えば、Van Heeke, G. および S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509を参照)。例えば抗体類の産生などに多量のREMAPが必要な場合は、高度レベルのREMAPの発現を誘導するベクター類が使用できる。例えば、強力な誘導SP6バクテリオファージプロモーターまたは誘導T7バクテリオファージプロモーターを含むベクターが使用できる。

30

40

【0206】

酵母の発現系を使用してREMAPを産出することもできる。因子、アルコールオキシダーゼ、PGHプロモーター等の構成型或いは誘導型のプロモーターを含む多数のベクターが、出芽酵母菌 (*Saccharomyces cerevisiae*) またはピキア酵母 (*Pichia pastoris*) に使用可能である。更に、このようなベクターは、発現したタンパク質の分泌か細胞内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. 他 (1987) Methods Enzymol. 153:516-544、及び Scorer, C. A. 他 (1994) Bio/Technology 121 - 181-184.を参照)。

【0207】

50

植物系もREMAPの発現に使用可能である。REMAP をコードする配列の転写は、ウイルスプロモーター、例えば単独あるいはTMV由来のオメガリーダー配列と組み合わせて用いられるようなCaMV由来の35Sおよび19Sプロモーターによって促進し得る (Takamatsu, N. (1987) EMBO J 6:307-311)。或いは、RUBISCOの小サブユニット等の植物プロモーターまたは熱ショックプロモーターを用いてもよい (例えば、Coruzzi, G. 他 (1984) EMBO J. 3 : 1671-1680 ; Broglie, R. 他 (1984) Science 224 : 838-843 ; および Winter, J. 他 (1991) Results Probl. Cell Differ. 17 : 85-105を参照) これらの構成物は、直接DNA形質転換または病原体を媒介とする形質移入によって、植物細胞内に導入可能である。(『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, 191-196ページ等を参照)。

10

【0208】

哺乳動物細胞においては、多数のウイルスベースの発現系を利用し得る。アデノウイルスが発現ベクターとして用いられる場合、後発プロモーター及び3連リーダー配列からなるアデノウイルス転写物/翻訳複合体にREMAP をコードする配列を結合し得る。アデノウイルスゲノムの非必須のE1またはE3領域への挿入を用いれば、宿主細胞内でREMAP を発現する感染ウイルスを得ることができる (例えば、Logan, J. および T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659を参照)。更に、ラウス肉腫ウイルス (RSV) エンハンサー等の転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させ得る。SV 40またはEBVをベースにしたベクターを用いてタンパク質を高レベルで発現させることもできる。

20

【0209】

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

【0210】

長期にわたり哺乳動物系内で組換えタンパク質を生成するためには、細胞株内のREMAPの安定発現が望ましい。例えば、REMAP をコードする配列を細胞株に形質転換するために、発現ベクター類と、同じ或いは別のベクター上の選択可能マーカー遺伝子とを用い得る。用いる発現ベクターは、複製のウイルス起源、及び/または内在性の発現エレメントを持ち得る。ベクターの導入後、選択培地に移す前に強化培地で約1 ~ 2日間細胞を増殖させることができる。選択可能マーカーの目的は選択培地への抵抗性を与えることであり、選択可能マーカーが存在することにより、導入された配列をうまく発現するような細胞の成長及び回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に適した組織培養技術を用いて増殖可能である。

30

【0211】

任意の数の選択系を用いて、形質転換細胞株を回収できる。限定するものではないがこのような選択系には、tk⁻細胞のために用いられる単純ヘルペスウイルスのチミジンキナーゼ遺伝子と、apr⁻細胞のために用いられる単純ヘルペスウイルスのアデニンホスホリボシルトランスフェラーゼ遺伝子がある (例えば、Wigler, M. 他 (1977) Cell 11:223-232; 及びLowy, I. 他 (1980) Cell 22:817-823を参照)。また、選択の基礎として代謝拮抗物質、抗生物質或いは除草剤への耐性を用いることができる。例えばdhfrはメトトレキ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はクロルスルフロンに対する耐性を、patはホスフィノトリシンアセチルトランスフェラーゼに対する耐性を各々与える (Wigler, M. 他 (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. 他 (1981) J. Mol. Biol. 150:114 等を参照。) この他の選択可能な遺伝子、例えば、代謝のための細胞の必要条件を変えるtrpB及びhisDは、文献に記載されている (例えば、Hartman, S.C. 及びR.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051参照。) 可視マーカー、例えばアントシアニン、緑色蛍光タンパク質 (GFP; Clontech)、グルクロニダーゼ及びその基質 グルクロニド、

40

50

またはルシフェラーゼ及びその基質ルシフェリン等を用いてもよい。これらのマーカーを用いて、トランスフォーマントを特定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することが可能である (Rhodes, C.A. (1995) *Methods Mol. Biol.* 55:121-131等を参照)。

【0212】

マーカー遺伝子発現の存在 / 不存在によって目的の遺伝子の存在が示唆されても、その遺伝子の存在及び発現の確認が必要な場合もある。例えば、REMAP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、REMAP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。または単一プロモーター制御下で、REMAPをコードする配列とタンデムにマーカー遺伝子を配置することも可能である。誘導または選択に応答したマーカー遺伝子の発現は通常、タンデム遺伝子の発現も示す。

10

【0213】

一般に、REMAPをコードする核酸配列を持ち、REMAPを発現する宿主細胞は、当業者に周知の種々の方法を用いて同定し得る。これらの方法には、DNA-DNAあるいはDNA-RNAハイブリダイゼーションや、PCR法、核酸配列あるいはタンパク質配列の検出および / または数量化のための膜系、溶液ベース、ある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。

【0214】

特異的ポリクローナル抗体または特異的モノクローナル抗体を用いてREMAPの発現の検出及び計測を行うための免疫学的方法は、当分野で公知である。このような技術の例としては、酵素に結合したイムノソルベントアッセイ (ELISA)、ラジオイムノアッセイ (RIA)、蛍光活性化細胞選別 (FACS) などが挙げられる。REMAP上の2つの非干渉エピトープに反応するモノクローナル抗体を用いた、2部位のモノクローナルベースイムノアッセイ (two-site, monoclonal-based immunoassay) が好ましいが、競合の結合アッセイも用いることもできる。これらのアッセイ及びこれ以外のアッセイは、当分野で公知である (Hampton, R. 他 (1990) *Serological Methods, a Laboratory Manual*, APS Press, St Paul, MN, Sect. IV、Coligan, J. E. 他 (1997) *Current Protocols in Immunology*, Greene Pub. Associates and Wiley-Interscience, New York NY、Pound, J.D. (1998) *Immunochemical Protocols*, Humana Press, Totowa NJ等を参照)。

20

【0215】

多岐にわたる標識方法及び抱合方法が、当業者に知られており、様々な核酸アッセイおよびアミノ酸アッセイにこれらの方法を用い得る。REMAP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、REMAPをコードする配列、またはその任意の断片をmRNAプローブを生成するためのベクターにクローニングすることも可能である。このようなベクターは、当分野において知られており、市販もされており、T7、T3またはSP6等の好適なRNAポリメラーゼ及び標識されたヌクレオチドを加えて、*in vitro*でRNAプローブの合成に用いることができる。このような方法は、例えばAmersham Pharmacia Biotech、Promega (Madison WI)、U.S. Biochemical等から市販されている種々のキットを用いて実行することができる。検出を容易にするために用い得る好適なレポーター分子或いは標識には、基質、補助因子、インヒビター、磁気粒子のほか、放射性核種、酵素、蛍光剤、化学発光剤、発色剤等がある。

30

40

【0216】

REMAPをコードするヌクレオチド配列で形質転換された宿主細胞は、細胞培地でのこのタンパク質の発現及び回収に好適な条件下で培養される。形質転換細胞から製造されたタンパク質が分泌されるか細胞内に留まるかは、使用される配列、ベクター、或いはその両者に依存する。REMAP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PPの分泌を誘導するシグナル配列を含むように設計できることは、当業者には理解されよう。

50

【0217】

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

10

【0218】

本発明の別の実施例では、REMAP をコードする自然或いは変更された、または組換えの核酸配列を上記した任意の宿主系の融合タンパク質の翻訳となる異種配列にライゲーションさせる。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REMAPタンパク質が、REMAP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容易にする。また、異種タンパク質部分及び異種ペプチド部分も、市販されている親和性基質を用いて融合タンパク質の精製を促進し得る。限定されるものではないがこのような部分には、グルタチオンSトランスフェラーゼ (GST)、マルトース結合タンパク質 (MBP)、チオレドキシン (Trx)、カルモジュリン結合ペプチド (CBP)、6-His、FLAG、c-myc、赤血球凝集素 (HA) がある。GSTは固定化グルタチオン上で、MBPはマルトース上で、Trxはフェニルアルシンオキシド上で、CBPはカルモジュリン上で、そして6-Hisは金属キレート樹脂上で、同族の融合タンパク質の精製を可能にする。FLAG、c-myc及び赤血球凝集素 (HA) は、これらのエピトープ標識を特異的に認識する市販されているモノクローナル抗体及びポリクローナル抗体を用いて、融合タンパク質の免疫親和性精製を可能にする。また、融合タンパク質が、REMAP をコードする配列と異種タンパク質配列との間にタンパク質分解切断部位を持つように遺伝子操作すると、精製後にREMAP が異種部分から切断され得る。融合タンパク質の発現及び精製方法は、前出のAusubel (1995) 10章に記載されている。市販されている種々のキットを用いて融合タンパク質の発現及び精製を促進することもできる。

20

【0219】

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REMAPの合成が可能である。これらの系は、T7、T3またはSP6プロモーターと機能的に連結したタンパク質コード配列の転写及び翻訳をカップルさせる。翻訳は、例えば³⁵Sメチオニンのような放射能標識したアミノ酸前駆体の存在下で起こる。

30

【0220】

本発明のREMAPまたはその断片を用いて、REMAP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、REMAPへの特異的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。

40

【0221】

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REMAPの天然のリガンド、または天然の基質、構造的または機能的な擬態性または自然結合パートナーに密接に関連している (Coligan, J.E. 他 (1991) *Current Protocols in Immunology* 1(2)の5章等を参照)。同様に、化合物は、REMAP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。或る実施例では、これら化合物のためのスクリーニングには、分泌タンパク質として、あるいは細胞膜上でREMAPを発現する、好適な細胞群の作製が含まれる。好適な細胞には、哺乳動物、酵母、ショウジョウバエ、または大腸菌からの細胞が含まれる。PKINを発現する細胞またはREMA

50

P を含有する細胞膜断片を試験化合物と接触させて、REMAP 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。

【0222】

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

10

【0223】

本発明のREMAP またはその断片を用いて、REMAP 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、部分的アゴニスト、または逆アゴニスト等が含まれる。一実施例では、REMAP の活性が許容される条件下でアッセイを実施し、そのアッセイでは少なくとも1つの試験化合物とREMAPを混合し、試験化合物の存在下でのREMAP の活性が試験化合物不在下でのREMAP の活性と比較する。試験化合物の存在下でのREMAP の活性の変化は、REMAP の活性を調整する化合物の存在を示唆する。別の実施態様において、試験化合物をREMAP の活性に適した条件下でREMAP を含む *in vitro* または無細胞再構成系と混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、REMAP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はない。少なくとも1つ、または複数の試験化合物をスクリーニングすることができる。

20

【0224】

別の実施態様では、胚性幹細胞（ES細胞）における相同組み換えを用いて動物モデル系内で、REMAP またはその哺乳動物相同体をコードするポリヌクレオチドを「ロックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの産生に有用である（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ES細胞は初期のマウス胚に由来し、培地で増殖させることができる。このES細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（neo: Capecchi, M.R. (1989) *Science* 244:1288-1292）等のマーカー遺伝子で破壊された目的の遺伝子を含むベクターで形質転換する。このベクターは、相同組換えにより宿主ゲノムの対応する領域に組み込まれる。別法では、Cre-loxP系を用いて相同組換えを行い、組織特異的または発生段階特異的に目的遺伝子をロックアウトする（Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. 他 (1997) *Nucleic Acids Res.* 25:4323-4330）。形質転換したES細胞を同定し、例えばC57BL/6マウス系等から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を決め、これを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、可能性のある治療薬や毒性薬剤で検査することができる。

30

【0225】

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

40

【0226】

REMAP をコードするポリヌクレオチドを用いて、ヒト疾患をモデルとした「ロックイン」ヒト化動物（ブタ）または遺伝子組み換え動物（マウスまたはラット）を作製することが可能である。ロックイン技術を用いて、REMAP をコードするポリヌクレオチドのある領域を動物ES細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細胞を

50

胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について研究し、可能性のある医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばREMAPを乳汁内に分泌するなどREMAPを過剰に発現する哺乳動物近交系は、便利なタンパク質源となり得る (Janne, J. 他 (1998) *Biotechnol. Annu. Rev.* 4:55-74)。

【0227】

(治療)

REMAP領域と受容体および膜結合タンパク質領域との間に、例えば配列及びモチーフの内容における化学的及び構造的類似性が存在する。REMAPを発現する組織はまた、表6にも見られる。したがって、REMAPは細胞増殖異常、自己免疫/炎症、神経疾患、代謝疾患、発生障害および内分泌疾患に関与していると考えられる。REMAP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、REMAPの発現または活性を低下させることが望ましい。REMAP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、REMAPの発現または活性を増大させることが望ましい。

10

【0228】

従って、或る実施例において、REMAP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REMAPまたはその断片や誘導体を投与することが可能である。限定されるものではないが、そのような疾患としては日光性角化症、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、子宮の癌が含まれ、自己免疫/炎症性疾患には、後天性免疫不全症候群 (AIDS)、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、貧血症、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、アトピー皮膚炎、皮膚筋炎、真性糖尿病、肺気腫、リンパ球毒素性一時的リンパ球減少症、胎児赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、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、慢性関節リウマチ炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、神経の疾患の中には、癲癇、虚血性脳血管障害、脳卒中、大脳新生物、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病及びその他の錐体外路障害、筋萎縮性側索硬化及びその他の運動ニューロン障害、進行性神経性筋萎縮症、色素性網膜炎、遺伝性運動失調、多発性硬化症及び他の脱髄疾患、細菌性及びウイルス性髄膜炎、脳膿瘍、硬膜下蓄膿症、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎及び神経根炎、ウイルス性中枢神経系疾患と、クールー及びクローイツフェルト ヤコブ病、Gerstmann-Straussler-Scheinker症候群を含むプリオン病と、致死性家族性不眠症、神経系性栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜血管芽腫 (cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、ダウン症を含む中枢神経系性精神遅滞及び他の発達障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経疾患、脊髄疾患筋ジストロフィー、および他の神経筋障害、末梢神経系疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性、及び中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分性及び不安性の障害、分裂病性疾患を含む精神病と、季節性障害 (SAD) と、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群 (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、及びトゥーレット病と、進行性核上麻痺、大脳皮質基底核変性症及び

20

30

40

50

家族性前頭側頭性健忘症が含まれ、限定するものではないが、代謝系の疾患の中には、ア
 ジソン病、脳髄黄色腫症、先天性副腎過形成、クマリン耐性 (coumarin resistance)、
 嚢胞性繊維症、糖尿病、脂肪性肝硬変、フルクトース-1、6 - ジホスファターゼ欠乏症
 、ガラクトース血症、甲状腺腫、グルカゴノーマ、糖原病、遺伝性果糖不耐症、副腎皮質
 機能亢進症 (Hyperadrenalism)、副甲状腺機能亢進症、副甲状腺機能低下症、高コレス
 テロール血症、甲状腺機能亢進症、低血糖症、甲状腺機能低下症、脂肪過剰血症、脂質ミ
 オパシー、脂肪ジストロフィー症、リソソーム蓄積症、マノシドーシス、ノイラミニナー
 ゼ欠損症、肥満、骨粗しょう症、フェニルケトン尿症、プソイドビタミンD欠損くる病、
 炭水化物代謝障害 (例えば、先天性II型異常赤血球産生症性貧血、糖尿病、インスリン依
 存型真性糖尿病、非インスリン依存型真性糖尿病、ガラクトースエピメラーゼ欠乏症、糖
 10
 原病、リソソーム蓄積症、果糖尿、五炭糖尿症)、およびピルビン酸代謝の先天性異常、
 脂肪肝のような脂質代謝異常、胆汁うっ滞、原発性胆汁性肝硬変、カルニチン欠乏症、カ
 ルニチンパルミトイルトランスフェラーゼ欠乏症、ミオアデニレートデミネラーゼ欠乏症、
 高トリグリセリド血症、ファブリー病などの脂質貯蔵病、ゴーシェ病、ニーマン ピック
 病、変染色性白質ジストロフィー、副腎性白質ジストロフィー、 G_{M2} にガングリオシドー
 シス、セロイドリポフスチン症、無 リポ蛋白血症、タンジアー病、リポ蛋白過剰血症
 、脂肪異栄養症、脂肪腫症、急性皮下脂肪組織炎、播種性脂肪組織壊死症、有痛脂肪症、
 リポイド副腎過形成、最小変化症、脂肪腫、アテローム性動脈硬化症、高コレステロール
 20
 血症、高トリグリセリド血症を伴った高コレステロール血症、原発性低 リポ蛋白血症、
 甲状腺機能低下症、腎臓病、肝疾患、レシチン-コレステロールアシルトランスフェラ
 ーゼ欠乏症、脳髄黄色腫症、シトステロール血症、低コレステロール血症、テイ サック
 ス病、サンドホフ病、高脂血症、脂肪過剰血症、脂質筋障害、およびMenke病のような銅
 代謝疾患、ウイルソン病、およびエーレルス ダンロー症候群IX型糖尿病が含まれ、
 また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ー、癩癩、性腺形成異常、WAGR症候群
 (ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神薄弱)、スミス マジェニス症候群 (S
 mith- 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 ツース病及び神経線維腫症などの遺伝性神経病、甲状腺機能低下
 30
 症、水頭症、Sydenham舞蹈病 (Sydenham's chorea)及び脳性小児麻痺などの発作障害、
 脊髄二分裂、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感覚神経性聴力損失が含ま
 れ、内分泌疾患の中には原発脳腫瘍及び腺腫、妊娠性梗塞、下垂体切除、動脈瘤、血管奇
 形、血栓症、感染症、免疫異常、頭部外傷による合併症などの病変から起こる視床下部及
 び下垂体の障害と、性機能低下及びシーハン症候群、尿崩症、カルマン病、ハンド シュ
 ラークリスチャン病、レトラ シヴェ病、サルコイドーシス、エンブティセラ症候群、小
 人症を含む下垂体低下に関連した障害と、良性線種によって発生しやすい不适当抗利尿ホ
 ルモン (ADH) 分泌症候群 (SIADH) 及び先端巨大症、巨人症を含む下垂体亢進に関連した
 障害と、甲状腺腫及び粘液水腫、細菌感染性急性甲状腺炎、ウイルス感染性亜急性甲状腺
 40
 炎、自己免疫性甲状腺炎 (橋本病)、クレチン病を含む甲状腺機能低下症に関連した障害
 と、甲状腺中毒症及びその様々な型、グレーブス病、前脛骨粘液水腫、中毒性多結節性甲
 状腺腫、甲状腺癌、プランマー病を含む甲状腺機能亢進症と、Conn病 (chronic hypercal
 emia) を含む副甲状腺機能亢進症と、I型及びII型糖尿病及び合併症などの膵臓疾患と、
 過形成及び副腎皮質の癌腫や腺腫、アルカローシスに関連した高血圧、アミロイド症、低
 カリウム血、クッシング病、リドル症候群、Arnold-Healy-Gordon症候群、褐色細胞腫瘍
 、アジソン病、副腎機能不全などの副腎に関連した障害と、女性の異常プロラクチン産生
 及び不妊症、子宮内膜症、月経周期の摂動、多嚢胞性卵巣疾患、高プロラクチン血症、選
 択的性腺刺激ホルモン不全 (isolated gonadotropin deficiency)、無月経、乳汁漏出症
 、半陰陽、多毛症及び男性化、乳癌、閉経期後の骨粗鬆症、男性のライジッヒ細胞過形成
 、男性更年期、生殖細胞無形成症、ライジッヒ細胞腫瘍に関連した性機能亢進、アンドロ
 ゲン受容体の欠如に関連したアンドロゲン耐性、5 - 還元酵素症候群、女性化乳房症な
 どの生殖腺ステロイドホルモンに関連した疾患とが含まれる。

【0229】

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

【0230】

更に別の実施態様では、限定するものではないが上記した疾患を含む、REMAP 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REMAP を含む組成物を好適な医薬用キャリアと共に患者に投与することも可能である。

【0231】

更に別の実施例では、REMAP の活性を調節するアゴニストを患者に投与して、限定するものではないが上記した疾患を含むREMAP の発現または活性の低下に関連した疾患を治療または予防することも可能である。 10

【0232】

更なる実施例では、REMAP 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REMAP 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、上記の細胞増殖異常、自己免疫/炎症性疾患、神経疾患、代謝疾患、発達障害および、内分泌疾患が含まれる。一実施態様では、REMAP 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REMAP を発現する細胞または組織に薬剤を運ぶターゲティング或いは送達機構として間接的に用いられ得る。

【0233】

別の実施例では、REMAP をコードするポリヌクレオチドの相補体を発現するベクターを患者に投与して、限定するものではないが上記した疾患を含むREMAP の発現または活性の増大に関連した疾患を治療または予防することも可能である。 20

【0234】

別の実施例では、本発明の任意のタンパク質、アンタゴニスト、抗体、アゴニスト、相補配列、またはベクターを、別の好適な治療薬と組み合わせて投与することもできる。併用療法で用いる好適な治療薬は、当業者が従来の医薬原理に従って選択し得る。治療薬と組み合わせることにより、上記した種々の疾患の治療または予防に相乗効果をもたらし得る。この方法を用いることにより少量の各薬剤で医薬効果をあげることが可能となり、それによって副作用の可能性を低減し得る。 30

【0235】

REMAP のアンタゴニストは、本技術分野で一般的に知られている方法を用いて製造し得る。具体的には、精製されたREMAP を用いて抗体を作るか、治療薬のライブラリをスクリーニングして、TMP と特異結合するものを同定することが可能である。REMAP 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、Fab断片、及びFab発現ライブラリによって作られた断片が含まれる。但し、これらに限定されるものではない。中和抗体（即ち二量体の形成を阻害する抗体）は通常、治療用に好適である。

【0236】

抗体の産生のためには、ヤギ、ウサギ、ラット、マウス、ヒトその他など種々の宿主が、 40
REMAP または、免疫原性の特性を備えるその任意の断片またはオリゴペプチドの注入によって免疫化され得る。宿主の種に応じて、種々のアジュバントを用いて免疫応答を高めることもできる。限定するものではないがこのようなアジュバントには、フロイントアジュバントと、水酸化アルミニウム等のミネラルゲルアジュバントと、リゾレシチン、プルロニックポリオール、ポリアニオン、ペプチド、油性乳剤、スカシガイのヘモシニアン、ジニトロフェノール等の界面活性剤とがある。ヒトに用いられるアジュバントの中では、BCG（カルメット ゲラン杆菌）及びコリネバクテリウム パルバム（*Corynebacterium parvum*）が特に好ましい。

【0237】

REMAP に対する抗体類を誘導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 50

片は、少なくとも約5個のアミノ酸からなるアミノ酸配列を持つものが好ましく、一般的には約10個以上のアミノ酸からなる。これらのオリゴペプチド、ペプチドまたは断片は、天然のタンパク質のアミノ酸配列の一部と同一であることが望ましい。REMAP アミノ酸の短いストレッチは、KLHなど別のタンパク質の配列と融合されることができ、このキメラ分子に対する抗体群が産生され得る。

【0238】

REMAP に対するモノクローナル抗体は、抗体分子を産生する任意の技術を用いて、培地内の連続した細胞株によって作製し得る。限定するものではないがこのような技術には、ハイブリドーマ技術、ヒトB細胞ハイブリドーマ技術及びEBV-ハイブリドーマ技術がある (Kohler, G. 他. (1975) Nature 256:495-497、Kozbor, D. 他 (1985) .J. Immunol. Meth 10
ods 81:31-42、Cote, R.J. 他 (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030、Cole, S.P. 他 (1984) Mol. Cell Biol. 62:109-120等を参照)。

【0239】

更に、ヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの「キメラ抗体」作製のために開発した技術が、好適な抗原特異性及び生物学的活性を備える分子を得るために用いられる (例えば、Morrison, S.L.他. (1984) Proc. Natl. Acad. Sci. 81:6851685
5; Neuberger, M.S.他. (1984) Nature 312:604-608; Takeda, S.他. (1985) Nature 314:452-454を参照)。別法では、当分野で周知の方法を用いて、一本鎖抗体の産生のための記載された技術を適用して、REMAP特異的一本鎖抗体を生成する。関連特異性を有するが
20
イデオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリからチェーンシャッフリングによって産生することもできる (Burton D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137等を参照)。

【0240】

抗体の産生は、リンパ球集団における *in vivo* 産生の誘導によって、或いは文献に開示されているように非常に特異的な結合試薬の免疫グロブリンのライブラリまたはパネルのスクリーニングによっても行い得る。(Orlandi, R. 他 (1989) Proc. Natl. Acad. Sci. U
SA 86: 3833-3837、Winter, G. 他 (1991) Nature 349:293-299等を参照)。

【0241】

REMAP のための特異結合部位を有する抗体断片を産生することもできる。例えば、限定するものではないが、このような断片には、抗体分子のペプシン消化によって作製される F(ab')₂ 断片と、F(ab')₂ 断片のジスルフィド架橋を還元することによって作製される Fab 断片とがある。あるいは、Fab発現ライブラリを作製することによって、所望の特異性を持つモノクローナルFab断片を迅速且つ容易に同定することが可能となる (Huse, W.D. 他
30
(1989) Science 246:1275-1281等を参照)。

【0242】

種々のイムノアッセイを用いてスクリーニングし、所望の特異性を有する抗体を同定することができる。隔離された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる競合的な結合、または免疫放射線活性のための数々のプロトコルが、当分野では周知である。通常このようなイムノアッセイには、REMAPとその特異性抗体との間の複合体形成の計測が含まれる。二つの非干渉性REMAP エピトープに対して反応性のモノクローナル抗体を用いる、2部位モノクローナルベースのイムノアッセイが一般に利用されるが、競合的結合アッセイも利用することができる (Pound、前出)。
40

【0243】

ラジオイムノアッセイ技術と共にScatchard分析などの様々な方法を用いて、REMAPに対する抗体の親和性を評価する。親和性を結合定数Kaで表すが、このKaは、平衡状態の下でREMAP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。ポリクローナル抗体類は多数のREMAP エピトープに対して親和性が不均一であり、或るポリクローナル抗体試薬に関して判定したKaは、REMAP に対する抗体群の平均の親和性または結合活性を表す。特定のREMAP エピトープに単一特異的なモノクローナル抗体試薬のKaは、親和性の真の測定値を表す。Ka値が $10^9 \sim 10^{12}$ liter/mol の高親和性抗
50

体試薬は、REMAP抗体複合体が過酷な処理に耐えなければならないイムノアッセイに用いるのが好ましい。Ka値が $10^6 \sim 10^7$ liter/mo lの低親和性抗体試薬は、REMAPが抗体から最終的に活性化状態で解離する必要がある免疫精製及び類似の処理に用いるのが好ましい(Catty, D. (1988) Antibodies, Volume I: A Practical Approach. IRL Press, Washington, DC; Liddell, J. E. 及び Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY)。

【0244】

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

【0245】

本発明の別の実施例では、REMAP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を、治療目的で使用できる。ある実施態様では、REMAP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、アンチセンスオリゴヌクレオチドまたは大きな断片を、REMAPをコードする配列の制御領域、またはコード領域に沿ったさまざまな位置から設計可能である(Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totowa NJを参照。)治療に用いる場合、アンチセンス配列を好適な標的細胞に導入するのに好適な任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を作製する発現プラスミドの形で、細胞内に送達し得る(Slater, J.E. 他 (1998) *J. Allergy Clin. Immunol.* 102(3):469-475 及び Scanlon, K.J. 他 (1995)9(13):1288-1296.等を参照)アンチセンス配列はまた、例えばレトロウイルスやアデノ随伴ウイルスベクターなどのウイルスベクターを用いて細胞内に導入することもできる(Miller, A.D. (1990) *Blood* 76:271、前出の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公知のその他のシステムが含まれる(Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R. J.他 (1998) *J. Pharm. Sci.* 87(11):1308-1315、Morris, M.C. 他 (1997) *Nucleic Acid s Res.* 25(14):2730-2736.等を参照)。

【0246】

本発明の別の実施例では、REMAP 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遺伝子治療に用いることが可能である。遺伝子治療を行うことにより、(i) 遺伝子欠損症(例えばX染色体鎖遺伝(Cavazzana-Calvo, M. ら (2000) *Science* 288:669-672)により特徴付けられる重度の複合型免疫欠損(SCID)-X1の場合)、先天性アデノシンデアミナーゼ(ADA)欠損症に関連する重度の複合型免疫欠損(Blaese, R.M. ら (1995) *Science* 270:475-480、Bordignon, C. ら (1995) *Science* 270:470-475)、嚢胞性繊維症(Zabner, J. ら (1993) *Cell* 75:207-216; 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(thalassamiasis)、家族性高コレステロール血症、第VIII因子若しくは第IX因子欠損に起因する血友病(Crystal, R.G. (1995) *Science* 270:404-410、Verma, I.M. および N. Somia (1997) *Nature* 389:239-242)を治療し、(ii) 条件的致死性遺伝子産物を発現させ(例えば制御不能な細胞増殖に起因する癌の場合)、(iii) 細胞内の寄生虫(例えばヒト免疫不全ウイルス(HIV)(Baltimore, D. (1988) *Nature* 335:395-396、Poesebla, E. ら (1996) *Proc. Natl. Acad. Sci. USA.* 93:11395-11399)、B型若しくはC型肝炎ウイルス(HBV、HCV)、Candida albicans及びParacoccidioides brasiliensis等の真菌寄生

虫、並びに *Plasmodium falciparum* 及び *Trypanosoma cruzi* 等の原虫寄生体に対する防御機能を有するタンパク質を発現させることができる。REMAP 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 REMAP 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。

【0247】

本発明の更なる実施例では、REMAPの欠損による疾患や障害は、REMAP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REMAP欠損細胞に導入することによって治療される。*in vivo* 或いは *ex vitro* の細胞に用いる機械的導入技術には、(i) 個々の細胞内への直接的なDNA微量注射法、(ii) 遺伝子銃、(iii) リポソームを介した形質移入、(iv) 受容体を介した遺伝子導入、及び(v) DNAトランスポソンの使用 (Morgan, R.A. および W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217、Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. および H. Recipon (1998) Curr. Opin. Biotechnol. 9:445-450) がある。

10

【0248】

REMAP の発現に有効であり得る発現ベクターの例として、限定するものではないが pcDNA 3.1、EpiTag、pRcCMV2、pREP、pVAX、pCRII-TOPOTAベクター (Invitrogen, Carlsbad CA)、pCMV-Script、pCMV-Tag、PEGSH/PERV (Stratagene, La Jolla CA) および PTET-OFF、PTET-ON、PTRE2、PTRE2-LUC、PTK-HYG (Clontech, Palo Alto CA) が挙げられる。REMAP を発現させるために、(i) 恒常的に活性なプロモーター (例えば、サイトメガロウイルス (CMV)、ラウス肉腫ウイルス (RSV)、SV40ウイルス、チミジンキナーゼ (TK)、若しくは - アクチン遺伝子等)、(ii) 誘導性プロモーター (例えば、市販の 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. 他 (1995) Science 268:1766-1769; Rossi, F.M.V. 及び H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456)、エクジソン誘導性プロモーター (市販のプラスミド 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 RU 486/ミフェプリストー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 REMAP 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。

20

【0249】

市販のリポソーム形質転換キット (例えば Invitrogen 社の PerFect Lipid Transfection Kit) を用いれば、当業者は経験にそれほど頼らないでもポリヌクレオチドを培養中の標的細胞に導入することが可能になる。別法では、リン酸カルシウム法 (Graham, F.L. 及び A.J. Eb (1973) Virology 52:456-467) 若しくは電気穿孔法 (Neumann, B. 他 (1982) EMBO J. 1:841-845) を用いて形質転換を行う。初代培養細胞にDNAを導入するためには、標準化された哺乳動物の形質移入プロトコルの修飾が必要である。

30

【0250】

本発明の別の実施態様では、REMAP の発現に関連する種々の遺伝子欠損によって起こる疾患や障害を、(i) レトロウイルス長末端反復配列 (LTR) プロモーター若しくは或る独立プロモーターのコントロール下で REMAP 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群と、(iii) 効率的なベクター増殖に必要なコーディング配列群と追加のレトロウイルス・シス作用性 RNA 配列群とを伴う 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し得る。レトロウイルスベクター (例えば PFB 及び PFBNE0) は Stratagene 社から市販されており、刊行データ (Riviere, I. 他 (1995) Proc. Natl. Acad. Sci. U.S.A. 92:6733-6737) に基づいている。上記データを引用することをもって本明細書の一部とする。ベクターは、好適なベクター産生細胞株 (VPCL) において増殖され、VPCL は、標的細胞上の受容体に対する親和性を有するエンペロープ遺伝子または VSVg 等の汎親和性エンペロープタンパク質を発現する (Armentano, D. 他 (1987) J. Virol. 61:1647-1650、Bender, M.A. 他 (1987) J. Virol. 61:1639-1646、Adam, M.A. および A.D. Miller (1988) J. Virol. 62:3802-3806、Dull, T. 他 (1

40

50

998) J. Virol. 72:8463-8471、Zufferey, R. 他 (1998) J. Virol. 72:9873-9880)。RIGGに付与された米国特許第5,910,434号(「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」)において、レトロウイルスパッケージング細胞系を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、細胞集団(例えばCD4⁺ T細胞)の形質導入、及び形質導入した細胞の患者への戻しは、遺伝子治療の分野では当業者に公知の方法であり、多数の文献に記載されている(Ranga, U. 他 (1997) J. Virol. 71:7020-7029、Bauer, G. 他 (1997) Blood 89:2259-2267、Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716、Ranga, U. 他 (1998) Proc. Natl. Acad. Sci. U.S.A. 95:1201-1206、Su, L. (1997) Blood 89:2283-2290)。

10

【0251】

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、REMAPの発現に関連する1つ以上の遺伝子異常を有する細胞にREMAPをコードするポリヌクレオチドを送達する。アデノウイルス系ベクターの作製及びパッケージングについては、当業者に公知である。複製欠損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を臍臓の無損傷の臍島内に導入するために融通のきくことが証明された(Csete, M.E. 他 (1995) Transplantation 27:263-268)。使用できる可能性のあるアデノウイルスベクターは、Armentanoに付与された米国特許第5,707,618号(「Adenovirus vectors for gene therapy」)に記載されており、引用することをもって本明細書の一部とする。アデノウイルスベクターについては、Antinozzi, P.A. 他 (1999) Annu. Rev. Nutr. 19:511-544 及び Verma, I.M. 及び N. Somia (1997) Nature 18:389:239-242も参照されたい。両文献は、引用することをもって本明細書の一部とする。

20

【0252】

別法では、ヘルペス系遺伝子治療の或る送達系を用いて、REMAPの発現に関連する1つ以上の遺伝子異常を持つ標的細胞群に、REMAPをコードするポリヌクレオチド群を送達する。単純ヘルペスウイルス(HSV)系ベクター類は、HSVが親和性を持つ中枢神経系の細胞にREMAPを導入する際に、特に有用であろう。ヘルペス系ベクターの作製及びパッケージングは、当業者に公知である。複製適格性単純ヘルペスウイルス(HSV)I型系のベクターは、レポーター遺伝子を霊長類の眼に送達するために用いられてきた(Liu, X. 他 (1999) Exp. Eye Res. 169:385-395)。HSV-1ウイルスベクターの作製についても、DeLucaに付与された米国特許第5,804,413号(「Herpes simplex virus swains for gene transfer」)が開示されており、該特許の引用をもって本明細書の一部とする。米国特許第5,804,413号には、ヒト遺伝子治療を含む目的のために好適なプロモーターの制御下において細胞に導入される少なくとも1つの外在性遺伝子を有するゲノムを含む組換えHSV d92についての記載がある。上記特許はまた、ICP4、ICP27及びICP22のために除去される組換えHSV系統の作製及び使用について開示している。HSVベクターについては、Goins, W.F. 他 (1999) J. Virol. 73:519-532 及び Xu, H. 他 (1994) Dev. Biol. 163:152-161も参照されたい。両文献は、引用をもって本明細書の一部とする。クローン化ヘルペスウイルス配列の操作、巨大ヘルペスウイルスのゲノムの異なった部分を含む多数のプラスミドを形質移入した後の組換えウイルスの産生、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は、当業者に公知の技術である。

30

40

【0253】

別法では、ウイルス(正の一本鎖RNAウイルス)ベクターを用いてREMAPをコードするポリヌクレオチドを標的細胞に送達する。プロトタイプのウイルスであるセムリキ森林熱ウイルス(Semliki Forest Virus, SFV)の生物学的研究が広範に行われており、遺伝子導入ベクターがSFVゲノムに基づいていることが分かった(Garoff, H. 及び K.-J. Li (1998) Cun. Opin. Biotech. 9:464-469)。ウイルスRNAの複製中に、通常はウイルスのキャプシッドタンパク質をコードするサブゲノムRNAが作り出される。このサブゲノムRNAは、完全長のゲノムRNAより高いレベルに複製されるため、酵素活性(例えばプロテアーゼ及びポリメラーゼ)を有するウイルスタンパク質に比べてキャプシッドタンパク質が過

50

剥產生される。同様に、REMAP をコードする配列をキャプシッドをコードする領域の代わりに ウイルスゲノム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REMAP をコードするRNAが産生され、高いレベルでREMAP が合成される。通常は ウイルスの感染が数日以内での細胞溶解に関係する一方で、シンドピスウイルス (SIN) の変異体を有するハムスター正常腎臓細胞 (BHK-21) の持続的な感染を確立する能力は、 ウイルスの溶解複製を遺伝子治療に適用できるように好適に変更可能であることを示唆している (Dryga, S.A. 他 (1997) *Virology* 228 :74-83)。様々な宿主に ウイルスを導入できることから、様々なタイプの細胞にREMAP を導入することができる。或る集団におけるサブセットの細胞の特定形質導入は、形質導入前に細胞の選別を必要とし得る。 ウイルスの感染性cDNAクローンの処置方法、 ウイルスのcDNA及びRNAの形質移入方法及び ウイルスの感染方法は、当業者に公知である。 10

【0254】

転写開始部位由来のオリゴヌクレオチドを用いて遺伝子発現を阻害することも可能である。転写開始部位とは例えば開始部位から数えて約 - 10 と約 + 10 の間である。同様に、三重らせん塩基対の形成方法を用いて阻害が可能となる。三重らせん塩基対形成は、ポリメラーゼ、転写因子または調節分子の結合のために十分に開くような二重らせんの能力を阻害するので、三重らせん塩基対形成は有用である。三重らせんDNAを用いる最近の治療の進歩については文献に記載がある (Gee, J.E. 他 (1994) in: Huber, B.E.及びB.I. Carr, *Molecular and Immunologic Approaches*, Futura Publishing Co., Mt. Kisco, NY, 163-177ページ等を参照)。相補配列またはアンチセンス分子もまた、転写物がリボソームに結合するのを阻止することによってmRNAの翻訳を阻止するべく設計することができる。 20

【0255】

リボザイムは酵素性RNA分子であり、RNAの特異的切断を触媒するためにリボザイムを用いることもできる。リボザイム作用のメカニズムは、相補的標的RNAへのリボザイム分子の配列特異的ハイブリダイゼーションと、その後起こる内ヌクレオチド鎖切断に関与している。例えば、遺伝子操作で作られたハンマーヘッド型リボザイム分子は、REMAPをコードする配列の内ヌクレオチド鎖分解性の切断を特異的且つ効果的に触媒する可能性がある。 30

【0256】

任意のRNA標的内の特異的リボザイム切断部位は、GUA、GUU、GUC配列を含めたリボザイム切断部位に対して標的分子をスキャンすることによって先ず同定される。一度同定されると、切断部位を含む標的遺伝子の領域に対応する15~20リボヌクレオチドの短いRNA配列が、そのオリゴヌクレオチドを機能不全にするような2次構造の特徴をもっていないかを評価することが可能になる。候補標的の適合性の評価も、リボヌクレアーゼ保護アッセイを用いて相補的オリゴヌクレオチドとのハイブリダイゼーションのアクセス可能性をテストすることによって行うことができる。 40

【0257】

本発明の相補リボ核酸分子及びリボザイムは、核酸分子合成のために当分野でよく知られている任意の方法を用いて作製し得る。任意の方法には、固相フォスフォアミダイト化学合成等のオリゴヌクレオチドを化学的に合成する方法がある。或いは、REMAPをコードするDNA配列の *in vitro* 及び *in vivo* 転写によってRNA分子を産出し得る。このようなDNA配列は、T7やSP6等の好適なRNAポリメラーゼプロモーターを用いて多様なベクター内に取り込むことが可能である。或いは、相補的RNAを構成的或いは誘導的に合成するようなこれらcDNA産物を、細胞系、細胞または組織内に導入することができる。 40

【0258】

細胞内の安定性を高め、半減期を長くするためにRNA分子を修飾することができる。限定するものではないが可能な修飾には、分子の5'末端、3'末端、あるいはその両方においてフランキング配列を追加したり、分子の主鎖内においてホスホジエステラーゼ結合ではなくホスホリボチオネートまたは2'-O-メチルを使用したりすることが含まれる。この概念 50

は、PNAの産出に固有のものであり、これら全ての分子に拡大することができる。それには、内因性エンドヌクレアーゼによって容易には認識されないアデニン、シチジン、グアニン、チミン、及びウリジンにアセチル -、メチル -、チオ - 及び同様の修飾をしたものの他、非従来型塩基、例えばイノシン、クエオシン (queosine)、ワイプトシン (wybutosine) 等を加えることでできる。

【0259】

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

【0260】

特異ポリヌクレオチドの変異発現における有効性に対して、少なくとも1個から複数個の試験化合物をスクリーニングし得る。試験化合物は、当分野で通常知られている任意の方法により得られる。このような方法には、ポリヌクレオチドの発現を変異させる場合と、既存の、市販のまたは専売の、天然または非天然の化合物ライブラリから選択する場合と、標的ポリヌクレオチドの化学的及び/または構造的特性に基づく化合物を合理的にデザインする場合と、組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合に有効であることが知られているような化合物の化学修飾がある。REMAP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、透過処理した細胞、無細胞再構成系または再構成生化学系があり得る。REMAPをコードするポリヌクレオチドの発現における変化は、当分野で通常知られている任意の方法でアッセイする。通常、REMAPをコードするポリヌクレオチドの配列に相補的なヌクレオチド配列を有するプローブを用いたハイブリダイゼーションにより、特異ヌクレオチドの発現を検出する。ハイブリダイゼーション量を定量し、それによって1つ若しく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対する基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現における変化の検出は、ポリヌクレオチドの発現を変異する際に試験化合物が有効であることを示している。特異ポリヌクレオチドの変異発現に有効な化合物に対して、例えば分裂酵母 (*Schizosaccharomyces pombe*) 遺伝子発現系 (Atkins, D. 他 (1999) 米国特許第5,932,435号、Arndt, G.M. 他 (2000) *Nucleic Acids Res.* 28:E15) またはHeLa細胞等のヒト細胞系 (Clarke, M.L. 他 (2000) *Biochem. Biophys. Res. Commun.* 268:8-13) を用いてスクリーニングを実行する。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性を調べるためのオリゴヌクレオチド (デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾オリゴヌクレオチド) の組合せライブラリをスクリーニングすることに関与している (Bruce, T.W. 他 (1997) 米国特許第5,686,242号、Bruce, T.W. 他 (2000) 米国特許第6,022,691号)。

【0261】

ベクターを細胞または組織に導入する多数の方法が利用可能であり、*in vivo*、*in vitro* 及び *ex vivo* の使用に対して同程度に適している。*ex vivo* 治療の場合、ベクターを患者から採取した幹細胞内に導入し、クローニング増殖して同一患者に自家移植で戻すことができる。トランスフェクション、リボソーム注入またはポリカチオンアミノポリマーによる送達は、当分野でよく知られている方法を用いて実行することができる (Goldman, C.K. 他 (1997) *Nat. Biotechnol.* 15:462-466. 等を参照)。

10

20

30

40

50

【0262】

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

【0263】

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

10

【0264】

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

【0265】

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

20

【0266】

本発明での使用に適した組成物には、所定の目的を達成するために必要なだけの量の活性成分を含有する成分が含まれる。有効投与量の決定は、当業者の能力の範囲内で行う。

【0267】

特殊形状の成分は、REMAP またはその断片を含む高分子を直接細胞内輸送するために調製される。例えば、細胞不透過性高分子を含むリポソーム製剤は、細胞融合及び高分子の細胞内送達を促進し得る。別法では、REMAP またはその断片をHIV Tat-1タンパク質の陽イオンN末端部に結合することもできる。このようにして生成された融合タンパク質は、マウスモデル系の脳を含む全ての組織の細胞に形質導入することがわかっている(Schwarze, S.R. 他 (1999) Science 285:1569-1572)。

30

【0268】

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

【0269】

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

40

50

【0270】

正確な投与量は、治療が必要な被験者に関する要素を考慮して、現場の医者が決定することになる。効果的なレベルの活性成分を与え、或いは所望の効果を維持するべく、投与量及び投与を調節する。被験者に関する要素としては、疾患の重症度、患者の通常の状態、患者の年齢、体重及び性別、投与の時間及び頻度、薬剤の配合、反応感受性及び治療に対する応答等を考慮する。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よって3～4日毎に1度、1週間に1度、或いは2週間に1度の間隔で投与し得る。

【0271】

通常投与量は、投与の経路にもよるが約0.1～100,000 µgであり、合計で約1g 10
までとする。特定の投与量及び送達方法に関するガイダンスは文献に記載されており、現場の医者は通常それを利用することができる。当業者は、タンパク質またはインヒビターに対する処方とは異なる、ヌクレオチドに対する処方を利用することになる。同様に、ポリヌクレオチドまたはポリペプチドの送達は、特定の細胞、状態、位置等に特異的なものとなる。

【0272】

(診断)

別の実施例では、REMAPの発現によって特徴付けられる疾患の診断のために、或いはREMAP 20
やREMAPのアゴニスト、アンタゴニストまたは阻害剤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おいて、REMAPを特異的に結合する抗体が用いられることがある。診断目的に有用な抗体は、上記の治療の箇所で記載した方法と同じ方法で調合される。REMAPの診断アッセイには、ヒトの体液から、あるいは細胞や組織の抽出物から、抗体および標識を用いてREMAPを検出する方法が含まれる。この抗体は修飾されたものも、されていないものも可能であり、レポーター分子との共有結合または非共有結合で標識化できる。レポーター分子としては広くさまざまな種類が本分野で知られており、また使用可能であるが、そのうちのいくつかは上記で説明されている。

【0273】

REMAPを測定するための、ELISA, RIA, 及びFACSなど種々のプロトコルは、当分野では既知であり、変容した或いは異常なレベルのREMAP発現を診断するための基盤を提供する。30
正常或いは標準的なREMAPの発現の値は、複合体の形成に適した条件下で、正常な哺乳動物、例えばヒトなどの被験者から採取した体液または細胞とREMAPに対する抗体とを結合させることによって決定する。標準複合体形成量は、種々の方法、例えば測光法で定量できる。被験者、対照、及び疾患生検組織からの各サンプルのREMAPの発現の量が基準値と比較される。基準値と被験者との偏差が疾患を診断するパラメータとなる。

【0274】

本発明の別の実施例によれば、REMAPをコードするポリヌクレオチドを診断のために用い 40
ることもできる。用いられることができるポリヌクレオチドには、オリゴヌクレオチド配列、相補的RNA及びDNA分子、そしてPNAが含まれる。このポリヌクレオチドを用いて、疾患と相関し得るREMAP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、REMAPの存在の有無、更に過剰な発現を調べ、治療中のREMAP値の調節を監視する。

【0275】

一実施形態では、REMAP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、REMAPをコードする核酸配列を同定することが可能である。プローブが高度に特異的な領域(例えば5'調節領域)から作成されていようと、あるいは、やや特異性の低い領域(例えば保存されたモチーフ)から作成されていようと、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REMAPをコードする天然の配列のみを同定するか、或いは対立形質変異体や関連配列をコードする自然界の配列のみを同定するかによって決まるであろう。

【0276】

プローブはまた、関連する配列の検出に利用され、REMAP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:16-30の配列、或いはREMAP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。

【0277】

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

【0278】

REMAPをコードするポリヌクレオチド配列を用いて、REMAPの発現に関連する疾患を診断することが可能である。限定されるものではないが、そのような疾患の例として、細胞増殖異常には日光性角化症、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、子宮の癌が含まれ、自己免疫/炎症性疾患には、後天性免疫不全症候群(AIDS)、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、貧血症、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、アトピー皮膚炎、皮膚筋炎、真性糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、胎児赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、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、慢性関節リウマチ炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、神経の疾患の中には、癲癇、虚血性脳血管障害、脳卒中、大脳新生物、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病及びその他の錐体外路障害、筋萎縮性側索硬化及びその他の運動ニューロン障害、進行性神経性筋萎縮症、色素性網膜炎、遺伝性運動失調、多発性硬化症及び他の脱髄疾患、細菌性及びウイルス性髄膜炎、脳膿瘍、硬膜下蓄膿症、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎及び神経根炎、ウイルス性中枢神経系疾患と、クールー及びクロイツフェルト ヤコブ病、Gerstmann-Straussler-Scheinker症候群を含むプリオン病と、致死性家族性不眠症、神経系性栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜血管芽腫(cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、ダウン症を含む中枢神経系性精神遅滞及び他の発達障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経疾患、脊髄疾患筋ジストロフィー、および他の神経筋障害、末梢神経系疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性、及び中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分性及び不安性の障害、分裂病性疾患を含む精神病と、季節性障害(SAD)と、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群(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、及びトウレット病と、進行性核上麻

10

20

30

40

50

痺、大脳皮質基底核変性症及び家族性前頭側頭性健忘症が含まれ、代謝系の疾患の中には、アジソン病、脳髄黄色腫症、先天性副腎過形成、クマリン耐性 (coumarin resistance)、嚢胞性繊維症、糖尿病、脂肪性肝硬変、フルクトース-1、6 - ジホスファターゼ欠乏症、ガラクトース血症、甲状腺腫、グルカゴノーマ、糖原病、遺伝性果糖不耐症、副腎皮質機能亢進症 (Hyperadrenalism)、副甲状腺機能亢進症、副甲状腺機能低下症、高コレステロール血症、甲状腺機能亢進症、低血糖症、甲状腺機能低下症、脂肪過剰血症、脂質ミオパシー、脂肪ジストロフィー症、リソソーム蓄積症、マノシドーシス、ノイラミニナーゼ欠損症、肥満、骨粗しょう症、フェニルケトン尿症、プソイドビタミンD欠損くる病、炭水化物代謝障害(例えば、先天性II型異常赤血球産生性貧血、糖尿病、インスリン依存型真性糖尿病、非インスリン依存型真性糖尿病、ガラクトースエピメラーゼ欠乏症、糖原病、リソソーム蓄積症、果糖尿、五炭糖尿症)、およびピルビン酸代謝の先天性異常、脂肪肝のような脂質代謝異常、胆汁うっ滞、原発性胆汁性肝硬変、カルニチン欠乏症、カルニチンパルミトイルトランスフェラーゼ欠乏症、ミオアデニレートデミネラーゼ欠乏症、高トリグリセリド血症、ファブリー病などの脂質貯蔵病、ゴーシェ病、ニーマンピック病、変染色性白質ジストロフィー、副腎性白質ジストロフィー、 G_{M2} にガングリオシドーシス、セロイドリポフスチン症、無リポ蛋白血症、タンジアー病、リポ蛋白過剰血症、脂肪異栄養症、脂肪腫症、急性皮下脂肪組織炎、播種性脂肪組織壊死症、有痛脂肪症、リポイド副腎過形成、最小変化症、脂肪腫、アテローム性動脈硬化症、高コレステロール血症、高トリグリセリド血症を伴った高コレステロール血症、原発性低リポ蛋白血症、甲状腺機能低下症、腎臓病、肝疾患、レシチン-コレステロールアシルトランスフェラーゼ欠乏症、脳髄黄色腫症、シトステロール血症、低コレステロール血症、テイサク病、サンドホフ病、高脂血症、脂肪過剰血症、脂質筋障害、およびMenke病のような銅代謝疾患、ウイルソン病、およびエーレルスダンロー症候群IX型糖尿病が含まれ、また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ー、癩癩、性腺形成異常、WAGR症候群(ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神薄弱)、スミスマジエニス症候群(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、シャルコーマリーツース病及び神経線維腫症などの遺伝性神経病、甲状腺機能低下症、水頭症、Sydenham舞蹈病(Sydenham's chorea)及び脳性小児麻痺などの発作障害、脊髄二分裂、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感覚神経性聴力損失が含まれ、内分泌疾患の中には原発脳腫瘍及び腺腫、妊娠性梗塞、下垂体切除、動脈瘤、血管奇形、血栓症、感染症、免疫異常、頭部外傷による合併症などの病変から起こる視床下部及び下垂体の障害と、性機能低下及びシーハン症候群、尿崩症、カルマン病、ハンドシュラークリスチャン病、レトラシヴェ病、サルコイドーシス、エンブティセラ症候群、小人症を含む下垂体低下に関連した障害と、良性線種によって発生しやすい不适当抗利尿ホルモン(ADH)分泌症候群(SIADH)及び先端巨大症、巨人症を含む下垂体亢進に関連した障害と、甲状腺腫及び粘液水腫、細菌感染性急性甲状腺炎、ウイルス感染性亜急性甲状腺炎、自己免疫性甲状腺炎(橋本病)、クレチン病を含む甲状腺機能低下症に関連した障害と、甲状腺中毒症及びその様々な型、グレーブス病、前脛骨粘液水腫、中毒性多結節性甲状腺腫、甲状腺癌、プランマー病を含む甲状腺機能亢進症と、Conn病(chronic hypercalcemia)を含む副甲状腺機能亢進症と、I型及びII型糖尿病及び合併症などの膵臓疾患と、過形成及び副腎皮質の癌腫や腺腫、アルカローシスに関連した高血圧、アミロイド症、低カリウム血、クッシング病、リドル症候群、Arnold-Healy-Gordon症候群、褐色細胞腫瘍、アジソン病、副腎機能不全などの副腎に関連した障害と、女性の異常プロラクチン産生及び不妊症、子宮内膜症、月経周期の摂動、多嚢胞性卵巣疾患、高プロラクチン血症、選択的性腺刺激ホルモン不全(isolated gonadotropin deficiency)、無月経、乳汁漏出症、半陰陽、多毛症及び男性化、乳癌、閉経期後の骨粗鬆症、男性のライジッヒ細胞過形成、男性更年期、生殖細胞無形成症、ライジッヒ細胞腫瘍に関連した性機能亢進、アンドロゲン受容体の欠如に関連したアンドロゲン耐性、5-還元酵素症候群、女性化乳房症などの生殖腺ステロイドホルモンに関連した疾患とが含まれる。REMAPをコードするポリヌ

10

20

30

40

50

クレオチド配列は、変容したREMAP 発現を検出するために患者から採取した体液あるいは組織を利用する、サザン法やノーザン法、ドットプロット法、ある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)、およびマルチフォーマットのELISA式アッセイ、およびマイクロアレイに使用可能である。このような定性方法または定量方法は、当分野で公知である。

【0279】

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

10

【0280】

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

20

【0281】

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

30

【0282】

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

【0283】

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

40

【0284】

或る実施態様においては、REMAP をコードするポリヌクレオチド配列群に由来するオリゴヌクレオチドプライマー類を用いて、一塩基多型 (SNP) を検出し得る。SNP は、多くの場合にヒトの先天性または後天性遺伝病の原因となるような置換、挿入及び欠失である。限定するものではないがSNPの検出方法には、SSCP (single-stranded conformation polymorphism) 及び蛍光SSCP (fSSCP) がある。REMAP では、REMAPをコー

50

ドするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いたポリメラーゼ連鎖反応 (PCR) でDNAを増幅する。DNAは例えば、病変組織または正常組織、生検サンプル、体液その他に由来し得る。DNA内のSNPは、一本鎖形状のPCR生成物の2次及び3次構造に差異を生じさせる。差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。fSCCPでは、オリゴヌクレオチドプライマーを蛍光性に標識する。それによってDNAシーケンシング機などの高処理機器でアンプリマー (amplimer) の検出が可能になる。更に、インシリコSNP (in silico SNP, isSNP) と呼ばれる配列データベース分析法は、一般的なコンセンサス配列に配列されるような個々のオーバーラップするDNA断片の配列を比較することにより、多型性を同定し得る。これらのコンピュータベースの方法は、DNAの実験室での調整及び統計モデル及びDNA配列クロマトグラムの自動分析を用いたシーケンシングのエラーに起因する配列の変異をフィルタリングして除去する。別の態様では、例えば高処理MASSARRAYシステム (Sequenom, Inc., San Diego CA) を用いた質量分析によりSNPを検出し、特徴付ける。

10

【0285】

SNPはヒト疾患の遺伝的根拠の研究に用いられ得る。例えば、少なくとも、16のありふれたSNPは非インスリン型真性糖尿病に関連している。SNPは、例えば、嚢胞性線維症、鎌状赤血球貧血、または慢性肉芽腫症のような単一遺伝子病の疾患の現れ方の差異を調べる場合にも有用である。SNPにはまた生命を脅かす毒性のような薬剤に対する患者の反応に影響を与える遺伝的変異体の同定、薬物ゲノムにおける有用性がある。例えば、N-アセチルトランスフェラーゼにおける変異は抗結核剤、イソニアジドに反応した末梢神経障害の高発生率と関連しているが、ALOX5 遺伝子のコアプロモータの変異は5-リポキシゲナーゼ経路を標的とする抗喘息薬での治療に対する臨床的反応を減少する。異なった集団でのSNPの分布についての分析は遺伝的浮動、突然変異、組み換え体や選択および集団の起源の追跡および集団の移動について調べるために有用である。(Taylor, J.G. 他 (2001) *Trends Mol. Med.* 7:507-512; Kwok, P.-Y. および Z. Gu (1999) *Mol. Med. Today* 5:538-543; Nowotny, P. 他 (2001) *Curr. Opin. Neurobiol.* 11:637-641.)

20

REMAPの発現を定量するために用い得る方法には、ヌクレオチドの放射標識またはビオチン標識、調節核酸の相互増幅 (coamplification)、及び、標準曲線から得た結果の補間もある (例えば、Melby, P.C. 他 (1993) *J. Immunol. Methods* 159:235244; Duplaa, C. 他 (1993) *Anal. Biochem.* 212:229-236を参照。) 目的のオリゴマーが種々の希釈液中に存在し、分光光度法または比色反応によって定量が迅速になるような高処理フォーマットのアッセイを行うことによって、複数のサンプルの定量速度を加速することができる。

30

【0286】

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

40

【0287】

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

【0288】

50

或る実施例は、或る組織または細胞タイプの転写イメージを生成するような本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞タイプにより遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所定の条件下で所定の時間に発現した遺伝子の数及び相対存在量を定量することにより分析される（Seilhamer 他、米国特許第5,840,484号の「Comparative Gene Transcript Analysis」を参照。これらを引用することを以って本明細書の一部とする）。従って、特定の組織または細胞タイプの転写または逆転写の全体に本発明のポリヌクレオチドまたはその補体をハイブリダイズすることにより、転写イメージを生成し得る。或る実施例では、本発明のポリヌクレオチドまたはその補体がマイクロアレイ上のエレメントのサブセットを複数含むような高処理フォーマットでハイブリダイゼーションを発生させる。結果として得られる転写イメージは、遺伝子活性のプロフィールを提供し得る。 10

【0289】

転写イメージは、組織、株化細胞、生検またはその生物学的サンプルから単離した転写物を用いて生成し得る。転写イメージは従って、組織または生検サンプルの場合には *in vivo*、または株化細胞の場合には *in vitro* での遺伝子発現を反映する。

【0290】

本発明のポリヌクレオチドの発現のプロフィールを作製する転写イメージはまた、工業的または天然の環境化合物の毒性試験のみならず、*in vitro*モデル系及び薬剤の前臨床評価と併せて使用し得る。全ての化合物は、作用及び毒性のメカニズムを示す、しばしば分子フィンガープリントまたは毒性シグネチャと称される特徴的な遺伝子発現パターンを惹起する（Nuwaysir, E.F. 他（1999）*Mol. Carcinog.* 24:15 3-159、Steiner, S. 及び N.L. Anderson（2000）*Toxicol. Lett.* 112-113:467-471、該文献は特に引用することを以って本明細書の一部となす）。試験化合物が、既知の毒性を有する化合物のシグネチャと同様のシグネチャを有する場合には、毒性特性を共有している可能性がある。フィンガープリントまたはシグネチャは、多数の遺伝子及び遺伝子ファミリーからの発現情報を含んでいる場合には、最も有用且つ正確である。理想的には、発現のゲノム全域にわたる測定が最高品質のシグネチャを提供する。たとえ、発現が任意の試験された化合物によって変化しない遺伝子があったとしても、それらの遺伝子の発現レベルを残りの発現データを標準化するために使用できるため、それらの遺伝子は重要である。標準化手順は、異なる化合物で処理した後の発現データの比較に有用である。毒性シグネチャの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シグネチャ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日に米国国立環境健康科学研究所（National Institute of Environmental Health Sciences）より発行されたPress Release 00-02を参照されたい。これについては<http://www.niehs.nih.gov/oc/news/toxchip.htm>で入手可能である）。従って、毒性シグネチャを用いる中毒学的スクリーニングの際に、全ての発現した遺伝子配列を含めることは重要且つ望ましいことである。 20 30

【0291】

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

【0292】

別の実施例は、本発明のポリペプチド配列を用いて組織または細胞タイプのプロテオームを分析することに関連する。プロテオームの語は、特定の組織または細胞タイプでのタンパク質発現の包括的パターンを指す。プロテオームの各タンパク質成分は、個々に更に分析の対象とすることができる。プロテオーム発現パターン即ちプロフィールは、所与の条件下で所与の時間に発現したタンパク質の数及び相対存在量を定量することにより分析さ 50

れる。従って細胞のプロテオームのプロフィールは、特定の組織または細胞タイプのポリペプチドを分離及び分析することにより作成し得る。或る実施例では、1次元等電点電気泳動によりサンプルからタンパク質を分離し、2次元ドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に応じて分離するような2次元ゲル電気泳動により分離が達成される(前出のSteiner および Anderson)。タンパク質は、通常はクーマシーブルー、あるいは銀染色液または蛍光染色液などの物質を用いてゲルを染色することにより、分散した、独自の位置にある点としてゲル中で可視化される。各タンパク質スポットの光学密度は、通常サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処理された、または未処理の生物学的サンプルから得られるタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を同定する。スポット内のタンパク質は、例えば化学的または酵素的切断とそれに続く質量分析を用いる標準的な方法を用いて部分的にシーケンシングする。スポット内のタンパク質の同一性は、その部分配列を、好適には少なくとも5個の連続するアミノ酸残基を、本発明のポリペプチド配列と比較することにより決定し得る。場合によっては、決定的なタンパク質同定のための更なる配列が得られる。

10

【0293】

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

20

【0294】

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

30

【0295】

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

40

【0296】

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

【0297】

マイクロアレイは、本技術分野でよく知られている方法を用いて調製し、使用し、そして

50

分析する (Brennan, T.M. 他 (1995) の米国特許第5,474,796号、Sчена, M. 他 (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619、Baldeschweiler 他 (1995) PCT出願第W095/251116号、Shalon, D.他 (1995) PCT出願第W095/35505号、Heller, R.A. 他 (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155、Heller, M.J. 他 (1997) 米国特許第5,605,662号等を参照)。様々なタイプのマイクロアレイが公知であり、詳細については、DN A Microarrays: A Practical Approach, M. Schena, 編集. (1999) Oxford University Press, Londonに記載されている。該文献は、特に引用することを以って本明細書の一部となす。

【0298】

本発明の別の実施例では、天然のゲノム配列をマッピングする際に有効なハイブリダイゼーションプローブを産出するため、REMAP をコードする核酸配列を用いることが可能である。コード配列または非コード配列のいずれかを用いることができ、或る例では、コード配列より非コード配列の方が好ましい。例えば、多重遺伝子ファミリーのメンバー内でのコード配列の保存により、染色体マッピング中に望ましくないクロスハイブリダイゼーションが生じる可能性がある。核酸配列は、特定の染色体、染色体の特定領域または人工形成の染色体、例えば、ヒト人工染色体 (HAC)、酵母人工染色体 (YAC)、細菌人工染色体 (BAC)、細菌P1産物、或いは単一染色体cDNAライブラリに対してマッピングされる (例えば、Harrington, J.J. 他 (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127134; and Trask, B.J. (1991) Trends Genet. 7:149154を参照。) 一度マッピングすると、本発明の核酸配列を用いて例えば病状の遺伝を特定の染色体領域の遺伝または制限酵素断片長多型 (RFLP) と関連するような遺伝子連鎖地図を作製できる。 (例えば、Lander, E.S. 及び D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照。)

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

【0299】

確定した染色体マーカーを用いた連鎖分析等の物理的マッピング技術及び染色体標本原位ハイブリッド形成法を用いて、遺伝地図を拡張することができる。例えばマウスなど別の哺乳動物の染色体上に遺伝子を配置することにより、正確な染色体上の遺伝子座が未知でも、関連するマーカー類をしばしば明らかにし得る。この情報は、位置クローニングその他の遺伝子発見技術を用いて遺伝的疾患を探す研究者にとって価値がある。疾患または症候群に關与する遺伝子が、血管拡張性失調症の11q22-23領域等、特定の遺伝子領域への遺伝的結合によって大まかに位置決めがなされると、該領域に対するいかなる配列マッピングも、更なる調査のための関連遺伝子或いは調節遺伝子を表すことができる (Gatti, R.A. 他 (1988) Nature 336:577-580等を参照) 転座、反転等に起因する、健常者、保有者、感染者の三者間における染色体位置の相違を発見するために本発明のヌクレオチド配列を用いてもよい。

【0300】

本発明の別の実施例では、REMAP、その触媒作用断片或いは免疫原断片またはそのオリゴペプチドを、種々の任意の薬剤スクリーニング技術における化合物のライブラリのスクリーニングに用いることができる。薬剤スクリーニングに用いる断片は、溶液中に遊離しているか、固体支持物に固定されるか、細胞表面上に保持されるか、細胞内に位置することになる。REMAP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。

【0301】

別の薬剤スクリーニング方法は、目的のタンパク質に対して好適な結合親和性を有する化

合物を高い処理能力でスクリーニングするために用いられる (Geysen, 他 (1984) PCT application W084/03564等を参照)。この方法においては、多数の異なる小さな試験用化合物を固体基板上で合成する。試験用化合物は、REMAP、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REMAPが当分野で周知の方法で検出される。精製されたREMAP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捉し、ペプチドを固体支持物に固定することもできる。

【0302】

別の実施例では、REMAPと特異結合可能な中和抗体がREMAPとの結合について試験用化合物競合する、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、REMAP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。

10

【0303】

別の実施例では、発展途上の分子生物学技術にREMAP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。

【0304】

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

【0305】

本明細書において開示した全ての特許、特許出願及び刊行物、特に米国特許第60/262.838号、第60/265.927号、第60/271.196号、第60/274.549号及び第60/334.179号は、言及することをもって本明細書の一部となす。

20

【実施例】

【0306】

1 cDNAライブラリの作製

Incyte cDNAは、LIFESEQ GOLDデータベース (Incyte Genomics, Palo Alto CA) に記載されたcDNAライブラリに由来するものである。ホモジナイズしてグアニジニウムイソチオシアネート溶液に溶解した組織もあり、また、ホモジナイズしてフェノールまたは好適な変性剤の混合液に溶解した組織もある。変性剤の混合液は、例えばフェノールとグアニジニウムイソチオシアネートの単相溶液であるTRIZOL (Life Technologies) 等である。結果として得られた溶解物は、塩化セシウムにおいて遠心分離するかクロロホルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方、或いは別の方法を用いて、溶解物からRNAを沈殿させた。

30

【0307】

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

40

【0308】

場合によってはStratagene社へのRNA提供を行い、対応するcDNAライブラリをStratagene社が作製することもあった。そうでない場合は、UNIZAPベクターシステム (Stratagene) またはSUPERSCRIPトプラスミドシステム (Life Technologies) を用いて本技術分野で既知の推奨方法または類似の方法でcDNAを合成し、cDNAライブラリを作製した (前出のAusubel, 1997, 5.1-6.6ユニットなどを参照)。逆転写は、オリゴd(T)またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖cDNAに連結反応させ、好適な制限酵素でcDNAを消化した。殆どのライブラリに対して、cDNAのサイズの選択 (300 ~ 1000 bp) は、SEPHACRYL S1000、SEPHAROSE CL2BまたはSEPHAROSE CL4Bカ

50

ラムクロマトグラフィー (Amersham Pharmacia Biotech)、或いは調製用アガロースゲル電気泳動法を用いて行った。cDNAは好適なプラスミドのポリリンカーの適合する制限酵素部位に結合させた。好適なプラスミドは例えば、PBLUESCRIPT プラスミド (Stratagene)、PSPORT1 プラスミド (Life Technologies)、PCDNA2.1 プラスミド (Invitrogen, Carlsbad CA)、PBK-CMV プラスミド (Stratagene)、PCR2-TOPOTA (Invitrogen)、PCMV-ICIS プラスミド (Stratagene)、pIGEN (Incyte Genomics, Palo Alto CA)、pRARE (Incyte Genomics) または pINCY (Incyte Genomics)、あるいはその誘導体である。組換えプラスミドは、Stratagene社のXL1-Blue、XL1-BlueMRFまたはSOLR、或いはLife Technologies社のDH5、DH10BまたはELECTROMAX DH10Bを含む大腸菌細胞に形質転換した。

【0309】

10

2 cDNAクローンの単離

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

【0310】

20

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

【0311】

3 シークエンシング及び分析

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

30

40

【0312】

Incyte cDNA配列に由来するポリヌクレオチド配列は、ベクター、リンカー及びポリ(A)配列を除去し、あいまいな塩基対をマスクすることによって有効性を確認した。その際、BLAST、動的プログラミング及び隣接ジヌクレオチド頻度分析に基づくアルゴリズム及びプログラムを用いた。次に、Incyte cDNA配列またはそれらの翻訳の問い合わせを、以下のデータベース群より選択した1つに対して行った。すなわち、公共のデータベース群 (例えばGenBankの霊長類およびげっ歯類、哺乳動物、脊椎動物、真核生物のデータベースと

50

、BLOCKS、PRINTS、DOMO、PRODOM)、PROTEOMEの、ヒト、ラット、マウス、線虫 (*Caenorhabditis elegans*)、出芽酵母 (*Saccharomyces cerevisiae*)、分裂酵母 (*Schizosaccharomyces pombe*) および *Candida albicans* (Incyte Genomics, Palo Alto CA) からの配列群を持つデータベース群、および、PFAM等の隠れマルコフモデル (HMM) ベースのタンパク質ファミリーデータベース群である (HMMは、遺伝子ファミリーのコンセンサス1次構造を分析する確率的アプローチである。例えば、Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365を参照のこと)。Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365等を参照) の選択に対するIncyte cDNA配列またはその翻訳を問い合わせた。問合せは、BLAST、FASTA、BLIMPS及びHMMERに基づくプログラムを用いて行った。Incyte cDNA配列は、完全長のポリヌクレオチド配列を産出するように構築した。或いは、GenBank cDNA、GenBank EST、ステッチされた配列、ストレッチされた配列またはGenscan予測コード配列 (実施例4及び5を参照) を用いてIncyte cDNAの集団を完全長まで伸長させた。Phred、Phrap及びConsedに基づくプログラムを用いて構築し、GenMark、BLAST及びFASTAに基づくプログラムを用いてcDNAの集団をオープンリーディングフレームに対してスクリーニングした。対応する完全長ポリペプチド配列を誘導するべく完全長ポリヌクレオチド配列を翻訳した。或いは、本発明のポリペプチドは完全長翻訳ポリペプチドの任意のメチオニン残基で開始し得る。引き続き、GenBankタンパク質データベース (genpept)、SwissProt、PROTEOMEデータベース、BLOCKS、PRINTS、DOMO、PRODOM及びProsites等のデータベース、PFAM等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する問合せによって完全長ポリペプチド配列を分析した。完全長ポリヌクレオチド配列はまた、MACDNASIS PROソフトウェア (日立ソフトウェアエンジニアリング, South San Francisco CA) 及びLASERGENEソフトウェア (DNASTAR) を用いて分析した。ポリヌクレオチド及びポリペプチド配列アラインメントは、アラインメントした配列と配列の一致率も計算するMEGALIGNマルチシークエンスアラインメントプログラム (DNASTAR) に組み込まれているようなCLUSTALアルゴリズムによって特定されるデフォルトパラメータを用いて作製する。

【0313】

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

【0314】

完全長ポリヌクレオチド配列およびポリペプチド配列の組み立て及び分析に用いる上記のプログラムは、SEQ ID NO:16-30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2の列4に示した。

【0315】

4 ゲノムDNAからのコード配列の同定及び編集

推定上の受容体および膜結合タンパク質は、公共のゲノム配列データベース (例えば、gbpriやgbhtg) においてGenscan遺伝子同定プログラムを実行して初めに同定された。Genscanは、様々な生物からゲノムDNA配列を分析する汎用遺伝子同定プログラムである (Burge, C. および S. Karlin (1997) *J. Mol. Biol.* 268:78-94 及びBurge, C. 及び S. Karlin (1998) *Curr. Opin. Struct. Biol.* 8:346-354参照)。プログラムは予測エキソンを連結し、メチオニンから停止コドンに及ぶ構築されたcDNA配列を形成する。Genscanの出力は、ポリヌクレオチド及びポリペプチド配列のFASTAデータベースである。Genscanが一度に分析する配列の最大範囲は、30kbに設定した。これらのGenscan予測cDNA配列の内、どの配列が受容体および膜結合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

10

20

30

40

50

たポリペプチドをPFAMモデルにおいて受容体および膜結合タンパク質について問合せて分析した。潜在的な受容体および膜結合タンパク質は、また受容体および膜結合タンパク質として注釈が付けられたIncyte cDNA配列に対する相同性を基に同定した。こうして選択されたGenscan予測配列は、次にBLAST分析により公共データベースgenpept及びgbpriと比較した。必要であれば、genpeptからのトップのBLASTヒットと比較することによりGenscan予測配列を編集し、余分なまたは取り除かれたエキソンなどのGenscanにより予測された配列のエラーを修正する。BLAST分析はまた、任意のIncyte cDNAまたはGenscan予測配列の公共cDNA適用範囲の発見に用いられるので、転写の証拠を提供する。Incyte cDNA適用範囲が利用できる場合には、この情報を用いてGenscan予測配列を修正または確認した。完全長ポリヌクレオチド配列は、実施例 3に記載された構築プロセスを用いて、Incyte cDNA配列及び/または公共のcDNA配列でGenscan予測コード配列を構築することにより得た。或いは、完全長ポリヌクレオチド配列は編集または非編集のGenscan予測コード配列に完全に由来する。

【0316】

5 cDNA配列データを使ったゲノム配列データの構築

ステッチ配列 (Stitched Sequence)

部分cDNA配列は、実施例 4に記載のGenscan遺伝子同定プログラムにより予測されたエキソンを用いて伸長させた。実施例 3に記載されたように構築された部分cDNAは、ゲノムDNAにマッピングし、関連するcDNA及び1つ若しくは複数のゲノム配列から予測されたGenscanエキソンを含むクラスタに分解した。cDNA及びゲノム情報を統合するべくグラフ理論及び動的プログラミングに基づくアルゴリズムを用いて各クラスタを分析し、引き続いて確認、編集または伸長して完全長配列を産出するような潜在的スプライス変異体を生成した。間隔全体の長さがクラスタ中の2以上の配列に存在するような配列を同定し、そのように同定された間隔は推移により等しいと考えられた。例えば、1つのcDNA及び2つのゲノム配列に間隔が存在する場合、3つの間隔は全て等しいと考えられる。このプロセスは、無関係であるが連続したゲノム配列をcDNA配列により結び合わせて架橋し得る。このようにして同定された区間を、親配列 (parent sequence) に沿って現われるようにステッチアルゴリズムで縫い合わせ、可能な最も長い配列および変異配列を作製する。1種類の親配列に沿って発生した間隔と間隔との連鎖 (cDNA - cDNAまたはゲノム配列 - ゲノム配列) は、親の種類を変える連鎖 (cDNA - ゲノム配列) に優先した。結果として得られるステッチ配列は、BLAST分析により公共データベースgenpept及びgbpriに翻訳されて比較された。Genscanにより予測された不正確なエキソンは、genpeptからのトップのBLASTヒットと比較することにより修正した。必要な場合には、追加cDNA配列を用いるかゲノムDNAの検査により配列を更に伸長させた。

【0317】

ストレッチ配列 (Stretched Sequence)

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

【0318】

6 REMAP 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:16-30を構築するために用いた配列を、BLAST及びSmith-Watermanアルゴリズムを用いて、Incyte LIFESEQデータベース及び公共のドメインデータベースの配列と比較した。SEQ ID NO:16-30と一致するこれらのデータベースの配列を、Phrap(表7)などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。スタンフォード・ヒトゲノムセンター(SHGC)、ホワイトヘッド・ゲノム研究所(WIGR)、Genethon等の公的な情報源から入手可能な放射線ハイブリッド及び遺伝地図データを用いて、クラスター化された配列が前にマッピングされていたかを確認した。マッピングされた配列が或るクラスターに含まれている場合、そのクラスターの全配列が、個々の配列番号と共に、地図上の位置に割り当てられた。

【0319】

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

【0320】

7 ポリヌクレオチド発現の分析

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

BLASTに適用する類似のコンピュータ技術を用いて、GenBankやLifeSeq(Incyte Pharmaceuticals)等のヌクレオチドデータベースにおいて同一または関連分子を検索する。ノーザン分析は、多数膜系ハイブリダイゼーションよりも非常に速い。更に、特定の一致を厳密な一致、或いは類似的な一致として分類するか否かを決定するため、コンピュータ検索の感度を変更することができる。検索の基準は積スコアであり、次式で定義される。

【0321】

【数1】

(BLASTスコア×配列一致率)

5 × (長さ(配列1), 長さ(配列2))の最小値

【0322】

積スコアは、2つの配列間の類似度及び配列が一致する長さの両方を考慮している。積スコアは、0~100の規準化された値であり、次のようにして求める。BLASTスコアにヌクレオチドの配列一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除する。高スコアリングセグメント対(HSP)に一致する各塩基に+5のスコアを割り当て、各不適性塩基対に-4を割り当てることにより、BLASTスコアを計算する。2つの配列は、2以上のHSPを共有し得る(ギャップにより隔離され得る)。2以上のHSPがある場合には、最高BLASTスコアの塩基対を用いて積スコアを計算する。積スコアは、断片的オーバーラップとBLASTアラインメントの質とのバランスを表す。例えば積スコア100は、比較した2つの配列の短い方の長さ全体にわたって100%一致する場合のみ得られる。積スコア70は、一端が100%一致し、70%オーバーラップしているか、他端が88%一致し、100%オーバーラップしているかのいずれかの場合に得られる。積スコア50は、一端が100%一致し、50%オーバーラップしているか、他端が79%一致し、100%

10

20

30

40

50

オーバーラップしているかのいずれかの場合に得られる。

【0323】

或いは、REMAPをコードするポリヌクレオチド配列を、ポリヌクレオチド配列が派生した組織源に関連して分析した。例えば或る完全長配列は、Incyte cDNA配列（実施例3を参照）と少なくとも一部はオーバーラップするように構築される。各cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各ヒト組織は、以下の生物/組織カテゴリー即ち心血管系、結合組織、消化器系、胎芽構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液及び免疫系、肝、筋骨格系、神経系、脾臓、呼吸器系、感覚器、皮膚、顎口腔系、非分類性/混合性または尿路の1つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。同様に、各ヒト組織は、以下の疾患/病状カテゴリー即ち癌、細胞系、発達、炎症、神経性、外傷、心血管、プール、その他の1つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。得られるパーセンテージは、REMAP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、LIFSEQ GOLD データベース（Incyte Genomics, Palo Alto CA）から得ることができる。

10

【0324】

8 REMAPをコードするポリヌクレオチドの伸長

完全長のポリヌクレオチド配列もまた、完全長分子の適切な断片から設計したオリゴヌクレオチドプライマーを用いて該断片を伸長させて生成した。一方のプライマーは既知の断片の5'伸長を開始するべく合成し、他方のプライマーは既知の断片の3'伸長を開始するべく合成した。開始プライマーは、長さが約22~30ヌクレオチド、GC含有率が約50%以上となり、約68~72の温度で標的配列にアニーリングするように、OLIGO 4.06ソフトウェア（National Biosciences）或いは別の適切なプログラムを用いて、cDNAから設計した。ヘアピン構造及びプライマー-プライマー二量体を生ずるようなヌクレオチドの伸長は全て回避した。

20

【0325】

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

30

【0326】

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

40

【0327】

1X TEに溶解したPICOGREEN定量試薬（0.25%（v/v）PICOGREEN、Molecular Probes, Eugene OR）100 μ lと、希釈していないPCR産物0.5 μ lとを不透明な蛍光光度計プレート（Coming Costar, Acton MA）の各穴に分配し、DNAを試薬と結合可能なようにさせることによって各穴内のDNA濃度の測定を行った。サンプルの蛍光を計測してDNAの濃度を定量するためにプレートをFluoroskan II（Labsystems Oy, Helsinki, Finland）でスキャンした。反応混合物のアリコート5~10 μ lを1%アガロースミニゲル上で電気泳動法に

50

よって解析し、どの反応が配列の伸長に成功したかを決定した。

【0328】

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

10

【0329】

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

20

【0330】

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

【0331】

9 REMAP がコードするポリヌクレオチドのSNP(一塩基多型)の同定

SNPとして知られるありふれたDNA配列変異体がLIFESEQ データベース (Incyte Genomics) を用いてSEQ ID NO:16-30 同定された。実施例3に記述されているように、同じ遺伝子からの配列を共にクラスターにして構築し、これによって遺伝子のすべての配列変異体の同定ができた。一連のフィルタからなるアルゴリズムを使って、SNPを他の配列変異体から区別した。予備フィルタは、最小限のPhredクオリティスコア15を要求することにより大多数のベースコールのエラーを除去し配列アライメントエラーや、ベクター配列、キメラおよびスプライス変異体の不適当なトリミングにより生じるエラーを取り除いた。染色体の高度解析の自動化手順により、推定SNPの近傍に最初のクロマトグラムファイルが解析された。クローンのエラーフィルタは統計的に生み出されたアルゴリズムを用いて、逆転写酵素、ポリメラーゼ、または体細胞突然変異によって引き起こされるエラーのような、実験処理時に導入されるエラーを識別した。クラスターエラーフィルタは、統計的に生み出されたアルゴリズムを用いて近接の相同体または偽遺伝子のクラスター化に起因するエラー、または非ヒト配列によるコンタミネーションにより生じたエラーを同定した。最後のフィルタによって複製と、免疫グロブリンまたはT細胞受容体に存在するSNPが除去された。

30

40

【0332】

高速処理MASSARRAY システム (Sequenom, Inc.) を用いて質量分析によってさらに特徴付けるために特定のSNPを選択して、四つの異なったヒト母集団中のSNP部位における対立遺伝子発生頻度を分析した。白人母集団は、ユタ州の83人、フランス人が4人、ベネズエラ

50

、3人およびアーミッシュ派人の2人を含む92人で構成された。アフリカ人母集団はすべてアメリカ黒人である194人(男性97人、女性97人)からなる。ラテンアメリカ人母集団はすべてメキシコ系ラテンアメリカ人の324人(男性162人、女性162人)からなる。アジア人母集団は126人(男性64人、女性62人)からなり、その内訳は中国人43%、日本人31%、韓国人13%、ベトナム人5%およびその他のアジア人8%と報告されている。対立形質の発生頻度は最初に白人母集団において分析し、いくつかの例において、この母集団で対立形質変化を示さなかったSNPは他の三つの部集団においてさらに検査しなかった。

【0333】

10 個々のハイブリダイゼーションプローブの標識及び使用

SEQ ID NO:16-30由来のハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記載するが、より大きなヌクレオチド断片に対しても事実上同一の手順が用いられる。オリゴヌクレオチドは、OLIGO 4.06ソフトウェア(National Biosciences)等の最新ソフトウェアを用いて設計し、各オリゴマー50 pmolと、[$-^{32}\text{P}$]アデノシン3リン酸(Amersham Pharmacia Biotech) 250 μCi と、T4ポリヌクレオチドキナーゼ(DuPont NEN, Boston MA)を混合することにより標識する。標識したオリゴヌクレオチドは、SEPHADEX G-25超細繊維分子サイズ排除デキストランビードカラム(Amersham Pharmacia Biotech)を用いて十分に精製する。Ase I、Bgl II、Eco RI、Pst I、Xba IまたはPvu II(DuPont NEN)のいずれか1つのエンドヌクレアーゼで消化したヒトゲノムDNAの典型的な膜ベースのハイブリダイゼーション解析において、毎分 10^7 カウントの標識されたプローブを含むアリコットを用いる。

【0334】

各消化物から得たDNAは、0.7%アガロースゲル上で分画してナイロン膜(Nytran Plus, Schleicher & Schuell, Durham NH)に移す。ハイブリダイゼーションは、40で16時間行う。非特異的シグナルを除去するため、例えば0.1Mクエン酸ナトリウム食塩水及び0.5%ドデシル硫酸ナトリウムに一致する条件下で、プロットを室温で順次洗浄する。オートラジオグラフィまたはそれに代わるイメージング手段を用いてハイブリダイゼーションパターンを視覚化し、比較する。

【0335】

11 マイクロアレイ

マイクロアレイ上でアレイエレメントの結合または合成は、フォトリソグラフィ、圧電印刷(インクジェット印刷、前出のBalteschweiler等を参照)、機械的マイクロスポッティング技術及びこれらから派生したものをを用いて達成することが可能である。上記各技術において基板は、均一かつ非多孔性の固体とするべきである(Schena (1999) 前出)。推奨する基板には、シリコン、シリカ、スライドガラス、ガラスチップ及びシリコンウエハがある。或いは、ドットプロット法またはスロットプロット法に類似のアレイを利用して、熱的、紫外線的、化学的または機械的結合手順を用いて基板の表面にエレメントを配置及び結合させてもよい。通常のアレイは、当業者に周知の利用可能な方法や機械を用いて作製でき、任意の適正数のエレメントを有し得る(Schena, M. 他 (1995) Science 270:467-470, Shalon, D. 他 (1996) Genome Res. 6:639-645, Marshall, A. および J. Hodgson (1998) Nat. Biotechnol. 16:27-31.を参照)。

【0336】

完全長cDNA、発現配列タグ(EST)、またはその断片またはオリゴマーは、マイクロアレイのエレメントと成り得る。ハイブリダイゼーションに好適な断片またはオリゴマーを、レーザGENEソフトウェア(DNASTAR)等の本技術分野で公知のソフトウェアを用いて選択することが可能である。アレイエレメントは、生物学的サンプル中でポリヌクレオチドを用いてハイブリダイズされる。生物学的サンプル中のポリヌクレオチドは、検出を容易にするために蛍光標識またはその他の分子タグに抱合される。ハイブリダイゼーション後、生物学的サンプルからハイブリダイズされていないヌクレオチドを除去し、蛍光スキャナを用いて各アレイエレメントにおいてハイブリダイゼーションを検出する。或いは、レー

10

20

30

40

50

ザ脱離及び質量スペクトロメトリを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上のエレメントにハイブリダイズする各ポリヌクレオチドの相補性の度合及び相対存在度は、算定し得る。一実施例におけるマイクロアレイの調整及び使用について、以下に詳述する。

【0337】

組織または細胞サンプルの調製

グアニジウムチオシアネート法を用いて組織サンプルから全RNAを単離し、オリゴ(dT)セルロース法を用いてポリ(A)⁺RNAを精製する。各ポリ(A)⁺RNAサンプルは、MMLV逆転写酵素、0.05 pg/ μ lのオリゴ(dT)プライマー(21mer)、1X第1鎖バッファ、0.03 unit/ μ lのRNアーゼ阻害剤、500 μ MのdATP、500 μ MのdGTP、500 μ MのdTTP、40 μ MのdCTP、40 μ MのdCTP-Cy3(BDS)またはdCTP-Cy5(Amersham Pharmacia Biotech)を用いて逆転写する。逆転写反応は、GEMBRIGHTキット(Incyte)を用いてポリ(A)⁺RNA含有の25体積ml内で行う。特異的対照ポリ(A)⁺RNAは、非コード酵母ゲノムDNAから *in vitro* 転写により合成する。37 で2時間インキュベートした後、各反応サンプル(1つはCy3、もう1つはCy5標識)は、2.5 mlの0.5 M水酸化ナトリウムで処理し、85 で20分間インキュベートし、反応を停止させてRNAを分解させる。サンプルは、2つの連続するCHROM A SPIN 30ゲル濾過スピカラム(CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA)を用いて精製する。混合後、2つの反応サンプルは、1 mlのグリコーゲン(1 mg/ml)、60 mlの酢酸ナトリウム及び300 mlの100%エタノールを用いてエタノール沈殿させる。サンプルは次に、SpeedVAC(Savant Instruments Inc., Holbrook NY)を用いて乾燥して仕上げ、14 μ lの5 \times SSC/0.2% SDS中で再懸濁する。

10

20

【0338】

マイクロアレイの調製

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

【0339】

精製したアレイエレメントは、ポリマーコートされたスライドガラス上に固定する。顕微鏡スライドガラス(Corning)は、処理中及び処理後に0.1%のSDS及びアセトン中で超音波をかけ、蒸留水で非常に良く洗って洗浄する。スライドガラスは、4%フッ化水素酸(VWR Scientific Products Corporation (VWR), West Chester PA)中でエッチングし、蒸留水中で広範囲にわたって洗浄し、95%エタノール中で0.05%アミノプロピルシラン(Sigma)を用いてコーティングする。コーティングしたスライドガラスは、110のオープンで硬化させる。

30

【0340】

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/ μ lのアレイエレメントDNA 1 μ lを高速機械装置により開放型キャピラリープリンティングエレメント(open capillary printing element)に充填する。装置はここで、スライド毎に約5 nlのアレイエレメントサンプルを加える。

40

【0341】

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

【0342】

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

10

【0343】

検出

レポーター標識されたハイブリダイゼーション複合体は、Cy3の励起のためには488 nm、Cy5の励起のためには632 nmでスペクトル線を発生し得るInnova 70混合ガス10 Wレーザー(Coherent, Santa Clara CA)を備えた顕微鏡で検出する。20 × 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の焦点を当てる。アレイを含むスライドを顕微鏡のコンピュータ制御のX-Yステージに置き、対物レンズを通してラスタースキャンする。本実施例で用いた1.8 cm × 1.8 cmのアレイは、20 µmの解像度でスキャンした。

【0344】

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

20

【0345】

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

30

【0346】

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

40

【0347】

グリッドが蛍光シグナルイメージ上に重ねられ、それによって各スポットからのシグナルはグリッドの各エレメントに集められる。各エレメント内の蛍光シグナルは統合され、シグナルの平均強度に応じた数値が得られる。シグナル分析に用いるソフトウェアは、GEMT00LS遺伝子発現分析プログラム(Incyte)である。

50

【0348】

1.2 相補的ポリヌクレオチド

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

10

【0349】

1.3 REMAPの発現

REMAP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REMAPが発現するために、抗生物質耐性及びcDNAの転写レベルを高める誘導性のプロモーターを含む好適なベクターにcDNAをサブクローニングする。このようなプロモーターには、lacオペレーター調節因子に関連するT5またはT7バクテリオファージプロモーター及びtrp-lac(tac)ハイブリッドプロモーターが含まれるが、これらに限定するものではない。組換えベクターを、BL21(DE3)等の好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピル-Dチオガラクトピラノシド(IPTG)で誘発されるとREMAPを発現する。真核細胞でのREMAPの発現は、一般にバキュロウイルスとして知られているAutographica californica核多角体病ウイルス(AcMNPV)を昆虫細胞株または哺乳動物細胞株に感染させて行う。バキュロウイルスの非必須の多角体遺伝子を、相同組換え或いは転移プラスミドの媒介を伴う細菌の媒介による遺伝子転移のどちらかによって、REMAPをコードするcDNAと置換する。ウイルスの感染力は維持され、強い多角体プロモーターによって高いレベルのcDNAの転写が行われる。組換えバキュロウイルスは、多くの場合はSpodoptera frugiperda(Sf9)昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる。(Engelhard, E. K.他(1994) Proc. Natl. Acad. Sci. USA 91:3224-3227, Sandig, V. 他(1996) Hum. Gene Ther. 7:1937-1945.等を参照)。

20

30

【0350】

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

40

【0351】

1.4 機能的アッセイ

REMAP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REMAPをコードする配列の発現によって評価する。cDNAを、cDNAを高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにサブクローニングする。選択ベクターには、pCMV SPORTプラスミド(Life Technologies)及びpCR 3.1プラスミド(Invitrogen)が含まれ、どちらもサイトメガロウイルスプロモーターを有する。リボソーム製剤或いは電気穿孔法

50

を用いて、5 ~ 10 μ gの組換えベクターをヒト細胞株、例えば内皮由来または造血由来の細胞株に一過的に形質移入する。更に、標識タンパク質をコードする配列を含む1 ~ 2 μ gのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、cDNAの組換えベクターからの発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質 (GFP; Clontech)、CD64またはCD64-GFP融合タンパク質から選択できる。自動化された、レーザー光学に基づく技術であるフローサイトメトリー (FCM) を用いて、GFPまたはCD64-GFPを発現する形質移入された細胞を同定し、その細胞のアポトーシス状態や他の細胞特性を評価する。FCMは、細胞死に先行するか或いは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、10
プロピジウムヨウ化物によるDNA染色によって計測される核DNA内容物の変化、前方光散乱と90°側方光散乱によって計測される細胞サイズと顆粒状性の変化、プロモデオキシウリジンの取込量の低下によって計測されるDNA合成の下方調節、特異抗体との反応性によって計測される細胞表面及び細胞内におけるタンパク質の発現の変化、及び蛍光複合アネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変化とがある。フローサイトメトリー法については、Ormerod, M. G. (1994) *Flow Cytometry* Oxford, New York, NY. に記述がある。

【0352】

遺伝子発現におけるREMAPの影響は、REMAPをコードする配列とCD64またはCD64-GFPのどちらかが形質移入された高度に精製された細胞集団を用いて評価することができる。CD64またはCD64-GFPは、形質転換された細胞表面で発現し、ヒト免疫グロブリンG (IgG) の保存領域と結合する。形質転換された細胞と形質転換されない細胞とは、ヒトIgGがCD64に対する抗体のどちらかで被覆された磁気ビードを用いて分離することができる (DYNAL, Lake Success, NY)。mRNAは、当分野で周知の方法で細胞から精製することができる。REMAPおよび目的の他の遺伝子をコードするmRNAの発現は、ノーザン分析やマイクロアレイ技術で分析することができる。20

【0353】

1.5 REMAPに特異的な抗体の作製

ポリアクリルアミドゲル電気泳動法 (PAGE; 例えば, Harrington, M.G. (1990) *Methods Enzymol.* 1816 - 3088-495を参照) または他の精製技術で実質的に精製されたREMAP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。30

【0354】

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

【0355】

通常は、長さ約15残基のオリゴペプチドを、Fmocケミストリを用いるABI 431A ペプチドシンセサイザ (Applied Biosystems) を用いて合成し、N-マレイミドベンゾイル-N-ヒドロキシスクシンイミドエステル (MBS) を用いた反応によってKLH (Sigma-Aldrich, St. Louis MO) に結合させて、免疫原性を高める (前出のAusubel, 1995 等を参照)。完全フロイントアジュバントにおいてオリゴペプチド-KLH複合体を用いてウサギを免疫化する。例えば、ペプチドまたはREMAP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と反応させて、得られた抗血清の抗ペプチド活性及び抗REMAP活性を検査する。40

【0356】

1.6 特異的な抗体を用いる天然REMAPの精製

天然REMAP或いは組換えREMAPを、REMAPに特異的な抗体を用いるイムノアフィニティークロマトグラフィにより実質的に精製する。イムノアフィニティークラムは、CNBr-活性化S 50

EPHAROSE (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REMAP抗体とを共有結合させることにより形成する。結合後に、製造者の使用説明書に従って樹脂をブロックし、洗浄する。

【0357】

REMAPを含む培養液をイムノアフィニティーカラムに通し、REMAP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REMAP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、REMAPを回収する。

【0358】

1.7 REMAP と相互作用する分子の同定

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

【0359】

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

【0360】

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

【0361】

1.8 REMAP 活性の実証

REMAP 活性に対する或るアッセイは、細胞表面におけるREMAPの発現を測定する。REMAPをコードするcDNAを、好適な哺乳動物細胞株に形質移入する。細胞表面タンパク質は、記載されているようにビオチンで標識する (de la Fuente, M.A. 他 (1997) *Blood* 90:2398-2405)。REMAP特異抗体を用いて免疫沈降を実行し、ドデシル硫酸ナトリウムポリアクリルアミドゲル電気泳動 (SDS-PAGE) 及び免疫プロット技術を用いて免疫沈降サンプルを分析する。標識された免疫沈降と未標識免疫沈降の比は、細胞表面に発現したREMAPの量に比例する。

【0362】

或いは、REMAP活性のためのアッセイは、リガンド/受容体が仲介する細胞増殖の調節のための原型アッセイに基づく。このアッセイでは、スイスマウス3T3細胞におけるDNA合成の速度を測定する。当分野で公知の形質移入方法を用いて、REMAPをコードするポリヌクレオチドを含むプラスミドを静止状態の3T3培養細胞に加える。一過性に形質移入された細胞は次に、放射性DNA前駆分子である ^3H チミジンの存在下でインキュベートする。次に、培養した細胞にREMAPリガンドの変化する量を加える。 ^3H チミジンの酸沈殿可能DNAへの取り込みは、ラジオアイソトープカウンターを用いて適切な時間間隔で測定する。取り込み量は、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REMAPリガンド濃度範囲に対する用量反応曲線がリニアであることは、受容体活性を示唆するものである。ミリリットル当りの活性の1単位は、50%の反応レベルを産出するREMAPの密度として画定され、100%であれば ^3H チミジンを酸沈殿可能DNAに最大限取り込むことを表す (McKay, I. and I. Leigh, eds. (1993) *Growth Factors: A Practical Approach*, Oxford University Press, New York NY, 73ページ)。

10

20

30

40

50

【0363】

或いは、REMAP 活性のためのアッセイは、REMAP ファミリータンパク質がGタンパク質活性化セカンドメッセンジャーシグナル伝達経路を調整する能力に基づく（例えばcAMP; Ga uclin, P. 他 (1998) J. Biol. Chem. 273:4990-4996）。当分野で既知の方法を用いて、完全長REMAP をコードするプラスミドを哺乳動物細胞系（例えばチャイニーズハムスター卵巣(CHO)またはヒト胎児腎臓(HEK-293)細胞系）に形質移入する。培地中の12穴トレイで形質移入された細胞を成長させ、培養液を廃棄して、接着した細胞をPBSで軽く洗浄する。培地においてリガンド有りまたはリガンドなしで細胞を30分間インキュベートし、次に培地を除去し、細胞を1Mの過塩素酸で処理して溶解する。溶解産物中のcAMPレベルは、当分野で既知の方法を用いて放射免疫アッセイにより測定する。リガンドに曝された細胞から得た溶解産物中のcAMPレベルをリガンドなしのものと比較したときの変化は、形質移入された細胞に存在するREMAP の量に比例する。

10

【0364】

イノシトールリン酸レベルの変化を測定するためには、 1×10^5 細胞/穴を含む24穴プレートで細胞を成長させ、イノシトールを含まない培地及び [^3H]ミオイノシトールを用いて2 μCi /穴で48時間インキュベートする。培養液を除去し、10 mMのLiClを含む緩衝液で細胞を洗浄し、その後リガンドを添加する。反応は、過塩素酸の添加により停止する。イノシトールリン酸は、Dowex AG1-X8 (Bio-Rad) 陰イオン交換樹脂及び液体シンチレーションにより計数された全ての標識されたイノシトールリン酸で抽出及び分離される。リガンドに曝された細胞から得た標識されたイノシトールリン酸のレベルをリガンドなしのものと比較したときの変化は、形質移入された細胞に存在するREMAP の量に比例する。

20

【0365】

更なる実施例において、REMAPのイオン伝導性機能は電気生理学的試験を用いて実証される。REMAPは、COS7、HeLa または CHOのような哺乳類細胞系をREMAPをコードする真核生物発現ベクターで形質転換することによって発現される。真核生物発現ベクターは市販されており、それらを細胞内に導入する技術は当業者には周知である。 - ガラクトシダーゼのような多数の標識遺伝子のいずれか一つを発現させる第二プラスミドの少量は、細胞へと同時形質転換され、異質なDNAを吸収し発現させるそれら細胞の迅速な同定を可能とする。細胞は形質転換の後、株化細胞がREMAP 及び13 - ガラクトシダーゼを発現し蓄積するのに適した条件下で、48 - 72時間に渡ってインキュベートされる。 - ガラクトシダーゼを発現する形質転換細胞は、本技術分野で既知の条件下で適切な比色基質を培地へ添加すると青く染色される。染色された細胞は、本技術分野で既知の電気生理学技術を用いて膜電気伝導度の違いを試験する。形質転換されていない細胞、及び/又はベクター配列だけ、若しくは - ガラクトシダーゼ配列だけのいずれかで形質転換された細胞は、対照として用いられ、並行して試験する。どちらかのREMAPに特異的な抗体を用いて細胞をインキュベートすることによってREMAP のカチオンまたはアニオン伝導性に寄与することが示される。それぞれの抗体は、REMAPの細胞外側に結合することにより、イオンチャンネル内のポアと、それに伴う伝導性とをブロックする。

30

更なる実施例において、REMAP の輸送活性は、アフリカツメガエル卵母細胞への標識された基質の取り込みを測定することによって検査される。

40

ステージ5及び6の卵母細胞は、REMAP mRNA (卵母細胞あたり10 ng)で注入され、OR2培地(82,5. 5mMのNaCl、2. 5 mMのKCl、1mMのCaCl₂、1mMのMgCl₂、1mM Na₂HPO₄、5 mMのHepes、3. 8 mMのNaOH、50 $\mu\text{g}/\text{ml}$ ゲンタマイシン、pH 7. 8) 内で18 で3日間インキュベートして、REMAP の発現を可能とする。卵母細胞は、次に標準取り込み培地(100mMのNaCl、2 mMのKCl、1mMのCaCl₂、1mMのMgCl₂、10 mMのHepes/Tris、pH 7. 5)へと移される。種々の基質(例えば、アミノ酸、糖、薬物および神経伝達物質)の取り込みは、卵母細胞にH³ 基質を添加することによって開始される。30分間インキュベートした後、卵母細胞をNa⁺遊離培養液中で卵母細胞を3回洗浄して取り込みを中止させ、取り込まれた ³Hを測定し、対照と比較する。REMAP活性は吸収された³H基質の濃度に比例する。

50

【0366】

さらなる実施例において、REMAP プロテインキナーゼ(PK)活性は、標識の $[^{32}\text{P}]\text{ATP}$ を用いたタンパク質基質のリン酸化と組み込まれた放射活性を定量する。REMAPはタンパク質基質、 $^{32}\text{P}\text{-ATP}$ 、及び適切なキナーゼバッファと共にインキュベートする。生成産物に組み込まれた ^{32}P は、電気泳動法で遊離 $^{32}\text{P}\text{-ATP}$ より分離され、組み込まれた ^{32}P をラジオアイソトープカウンターでカウントする。回収した ^{32}P の量はアッセイ中のREMAPのPK活性に比例する。リン酸化した特異的アミノ酸残基の判定を、加水分解したタンパク質のホスホアミノ酸解析で行う。

【0367】

19 REMAP リガンドの同定

REMAP は、CHO(チャイニーズハムスター卵巣)またはHEK(ヒト胎児腎臓)293などの真核細胞株において発現する。これらの細胞株は、GPCR発現過程が良好であり、発現したREMAPが下流エフェクターに機能的に共役できるような広範囲のGタンパク質を含む。候補リガンドの存在下で発現した受容体の活性化に対し、形質転換した細胞をアッセイする。活性は、cAMPまたは Ca^{2+} などの細胞内セカンドメッセンジャーの変化により測定する。当分野で公知の標準的な方法を用いるか、活性化受容体によるタンパク質キナーゼCの刺激に反応して発光タンパク質(例えばホタルルシフェラーゼまたは緑色蛍光タンパク質)がプロモーターの転写調節下にあるようなレポーター遺伝子アッセイを用いて直接測定する(Milligan, C.ら(1996) Trends Pharmacol. Sci. 17:235-237)。アッセイ技術は、これらセカンドメッセンジャーシステムの両方に利用可能であり、アデニリルシクラーゼ活性化FlashPlate Assay(NEN Life Sciences Products)などのマルチウェルプレートフォーマットまたはFluo-4 AM(Molecular Probes)などの蛍光 Ca^{2+} 指示薬でFLIPR蛍光定量的プレート読出しシステム(Molecular Devices)を併用して高処理読出しを可能にする。生理的関連性のあるセカンドメッセンジャー経路が知られていない場合、REMAPは、ホスホリパーゼC及び Ca^{2+} 可動化に関与する経路を介してREMAPのシグナル伝達を通すために、広範囲のGタンパク質と共役することが実証されているようなGタンパク質 $\text{G}_{15/16}$ と同時発現し得る(Offerrnanns, S. and M.I. Simon(1995) J. Biol. Chem. 270:15175-15180)。或いは、REMAPは内因性GPCRが不足しているような操作された酵母系に発現し、それによってREMAP活性化スクリーニングに対してバックグラウンドが存在しない利点を提供し得る。これらの酵母系は、ヒトGPCR及びGタンパク質を内因性酵母フェロモン受容体経路の対応する構成要素に対して置換する。シグナルに対する正常な酵母反応を、選択的培地の正の成長またはレポーター遺伝子発現に変換するように、下流シグナル伝達経路も変更する(Broach, J.R. and J. Thorner(1996) Nature 384(supp.):14-16)。既知のGPCRリガンド及びその他の天然の生理活性分子を含む推定上のリガンドに対して受容体をスクリーニングする。組織、生物学的液体及び細胞上澄みから抽出した生物学的抽出物もスクリーニングする。

【0368】

当業者は、本発明の範囲及び精神から逸脱することなく本発明に記載した方法及びシステムの種々の改変を行い得る。本発明について説明するにあたり特定の好適実施例に関連して説明を行ったが、本発明の範囲が、そのような特定の実施例に不当に制限されるべきではないことを理解されたい。実際に、分子生物学または関連分野の専門家には明らかな、本明細書に記載されている本発明の実施方法の様々な改変は、特許請求の範囲内にあるものとする。

【0369】

(表の簡単な説明)

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

【0370】

表2は、GenBank識別番号及び本発明のポリペプチドに最も近いGenBank相同体の注釈を示す。各ポリペプチドとそのGenBank相同体が一致する確率スコアも併せて示す。

10

20

30

40

50

【0371】

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

【0372】

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

【0373】

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。

【0374】

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

【0375】

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

【0376】

【表1】

表 1

Incyte プロジェクトID	Incyte ポリアブチド SEQ ID NO:	Incyte ポリアブチド ID	Incyte ポリスクレオチド SEQ ID NO:	Incyte ポリスクレオチド ID
71924779	1	71924779CD1	16	71924779CB1
2319430	2	2319430CD1	17	2319430CB1
7291877	3	7291877CD1	18	7291877CB1
1218126	4	1218126CD1	19	1218126CB1
7479161	5	7479161CD1	20	7479161CB1
7722591	6	7722591CD1	21	7722591CB1
2173285	7	2173285CD1	22	2173285CB1
7487619	8	7487619CD1	23	7487619CB1
7487607	9	7487607CD1	24	7487607CB1
7487616	10	7487616CD1	25	7487616CB1
7483204	11	7483204CD1	26	7483204CB1
7472099	12	7472099CD1	27	7472099CB1
7485443	13	7485443CD1	28	7485443CB1
3090414	14	3090414CD1	29	3090414CB1
7503710	15	7503710CD1	30	7503710CB1

10

20

30

【 0 3 7 7 】
【 表 2 - 1 】

表 2 - 1

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	基準 スコア	GenBank 相同体
1	71924779CD1	93213188	8.5E-245	[ラット]チロシン酪氨酸受容体α2イリ型 (Akmal, K.M. 他 (1996) Biol. Reprod. 54:1111-1119; Akmal, K.M. 他 (1997) Biol. Reprod. 56:549-556; Akmal, K.M. 他 (1998) Endocrinology 139:1239-1248)
2	2319430CD1	93395787	9.7E-61	[ヒト]視網膜通受容体 TR08 (Gennill, R.M. 他 (1998) Proc. Natl. Acad. Sci. USA 95:9572-9577)
3	7291877CD1	93834380	2.1E-26	[ラット]内因性因子 B12 受容体前駆体(Moestrup, S.K. 他(1998) J. Biol. Chem. 273:9289-9292)
4	1218126CD1	96650766	1.7E-30	[ヒト]PDZドメイン含有チロシンキナーゼ交換因子 I
5	7479161CD1	957734	6.3E-187	[ラット]潜在性リポタン結合タンパク質 (Dear, T.N. 他 (1991) EMBO J. 10:2813-2819)
6	7722591CD1	93449308	0.0	[ヒト]MEGF8 (Nakayama, M. 他 (1998) Genomics 51:27-34)
7	2173285CD1	913539682	0.0	[ヒト]コリチン結合タンパク質 HOOK3 (Valente, J.H. 他 (2001) J. Cell Biol. 152:923-934)
8	7487619CD1	91256393	7.1E-94	[ラット]味覚受容体タンパク質 TB 641 (Thomas, M.B 他(1996) Gene 178:1-5)
9	7487607CD1	95869218	4.5E-115	[マウス]味覚受容体(Strommann, J. 他(1999) Gene 236:281-291)
10	7487616CD1	91256393	1.6E-94	[ラット]味覚受容体タンパク質 TB 641 (Thomas, M.B 他(1996) Gene 178:1-5)
11	7483204CD1	91314663	9.7E-129	[家犬]犬嗅覚受容体 CEOLF2 (Issel-Tarver, L. and J. Rine (1996) Proc. Natl. Acad. Sci. USA 93:10897-10902)

10

20

30

【 0 3 7 8 】

【 表 2 - 2 】

表 2 - 2

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	確率 スコア	GenBank 相同体
12	7472099CD1	94680260	7.6E-81	[マウス] 臭気物質受容体 S18 (Malnic, B. 他 (1999) 細胞 96:713-723)
13	7485443CD1	95869918	2.7E-117	[マウス] 嗅覚受容体 (Strotmann, J. 他 (1999) Gene 236:281-291)
14	3090414CD1	94761597	3.4E-94	[マウス] MOR 3 Betal (Bulger, M. 他 (2000) Proc. Natl. Acad. Sci. USA 97:14560-14565)
15	7503710CD1	911877275	3.3E-207	[ヒト] hJ726C3.4 (オアルシログの依存性 ligand binding タンパク質 RYR3 (Rat))

【 0 3 7 9 】

【 表 3 - 1 】

10

20

30

表 3 - 1

SEQ ID NO:	Incyte 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	シグネチャ配列、ドメインおよびモチーフ	分析方法およびデータベース
1	71924779CD1	457	S61 S62 S98 S152 S227 S364 S383 S447 T248 T280 T392	N50 N53 N206 N207 N288 N448	核ホルモン受容体のリガンド結合ドメイン E225-S383 亜鉛フィンガー、C4タイプ(2つのドメイン): K81-D156, A326-R334 膜貫通ドメイン: D251-I268, N294-Q310 N末端は細胞質内 核ホルモン受容体 BL00031: C83-M115, Y117-M148 核内ホルモン受容体 DNA 結合領域シグネチャ S63-V130 核内ホルモン受容体 DNA 結合領域シグネチャ: C83-R109 C4-タイプスライド受容体 亜鉛フィンガーシグネチャ PR00047: A99-N114, R132-L140, L140-M148, C83-A99 ピタミンド受容体シグネチャ PR00350: C83-A99, C100-C119, G322-D341, M368-I391 スライドホルモン受容体シグネチャ PR00398: C83-C100, K104-H120, C125-C143, C230-I249, Q352-Y272, M279-H293 シチニン酸受容体、転写調節、DNA 結合、亜鉛フィンガー PD002627: K150-S224 PD005136: M1-P82 ホルモンファミリー受容体、転写調節、DNA 結合、亜鉛フィンガー PD000035: C83-S149 PD000112: F223-L329, N294-S383	HMMER-PFAM HMMER-PFAM TMAP BLIMPS-BLOCKS ProfileScan MOTIFS BLIMPS-PRINTS BLIMPS-PRINTS BLAST-PRODOM BLAST-PRODOM

10

20

30

【 0 3 8 0 】

【 表 3 - 2 】

表 3 - 2

SEQ ID NO:	Incyte 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	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
1					核内ホルモン受容体、DNA 結合領域: DM00047 P10276 77-345: S72-D341 DM00047 P10826 70-338: S72-P340 DM00047 P22605 104-372: S72-P340 DM00047 P22448 77-345: S72-P340.	BLAST-DMO
2	2319430CD1	663	S103 S248 S266 S356 S652 T179 T457 Y109	N455 N580	Zn フィンガー、 C3HC4 タイプ (RING フィンガー): C537-C574 膜貫通ドメイン A9-W29, Y53-P73, L80-Y100, I128-L148, E166-F186, G218-Y246, S272-Q300, T316-V336, I344-I364, A382-Q402, I413-L433, E462-I490 N 末端は細胞質内 ニューロレグリン (GAP-43) シグネチャ: Q606-H658 光化学系 II タンパク質 PF00421: T143-L197, G213-S248, L280-G314, Y206-D251, C271-A304 複数回膜貫通受容体 TRC8, patched PD1B5068: A3-I536	HMMER-PFAM TMAP
3	7291877CD1	504	S120 S146 S152 S318 S353 S400 S404 S424 S475 T68 T192 T244 T373 T410 T445 T445	N73 N90 N361 N409	シグナルペプチド: M1-G60 シグナルペプチド: M38-G60 CUB ドメイン: C65-Y170, C241-Y342 スリドメイン (SCR リポート): C178-C235 膜貫通ドメイン: M35-Y63 N 末端は細胞質内 アラジン A2B 受容体シグネチャ PRO0554: S353-T364, N263-D277 スリドメイン/リッチ PF00084. G197-Y208, A236-C235	BLAST- PRODOM SPScan HMMER HMMER-PFAM HMMER-PFAM TMAP BLIMPS- PRINTS BLIMPS-PFAM

10

20

30

【 0 3 8 1 】

【 表 3 - 3 】

表 3 - 3

SEQ ID NO:	Incyte 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	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース	
3					糖タンパク質ドメイン、EGF 様受容体、B-12 因子 PD00165: C241-Y342, C65-Y170 CIR/CIS ヲドド DM00162: A57190 826-947: Y63-E172, C241-Y342 I49540 438-552: C65-L173, C241-Y342 P98063 438-549: C65-Y170, C241-Y342 P98069 303-417: A240-Q343, T64-E171	BLAST- PRODOM BLAST-DMO	
4	1218126CD1	1114	S11 S62 S98 S106 S140 S254 S287 S316 S325 S332 S574 S584 S599 S645 S649 S656 S711 S775 S781 S843 S859 S894 S1012 T145 T164 T177 T191 T243 T519 T620 T666 T748 T796 T804 T877 T899 T1017 Y531	N189 N196 N222 N260 N393 N539 N572 N583 N773 N884			
5	7479161CD1	479	S152 S165 S377 S408 S422 S461 T144 T316 T386	N142 N288 N457	シグナル切断: M1-A18 シグナルペプチド: M1-P20 膜貫通ドメイン: A7-I34, V63-V89, L161-P181, Y186-L206, E211-S231; N-末端は細胞質内	SPSCAN HMMER TMAP	

【 0 3 8 2 】

【 表 3 - 4 】

表 3-4

SEQ ID NO:	Incyte 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	シグネチャ配列、ドメインおよびモチーフ	分析手法およびデータベース
5					潜在的リント結合タンパク質 RYA3: PD177882: F86-F261 PD053120: V4-R65 タンパク質前駆体 シガナル糖タンパク質 脂質輸送抗 生物質 膜貫通リポ多糖結合 LBP PD006440: G155-N468, L80-L119 リガンド: RV2G5, RYA3: DM05385 S17448 1-473: V4-A479 DM05385 S17447 1-470: G66-N468 リポ多糖結合タンパク質 DM02253 P18428 5-474: A7-I465 DM02253 P17213 11-486: L95-N468	BLAST- PRODOM BLAST- PRODOM BLAST-DOMO BLAST-DOMO
6	772259UCD1	1774	T102 T171 S264 S292 S343 S389 T467 T478 T499 S528 S530 T535 S549 T611 S733 S810 S850 T859 T880 S903 T951 S998 S1160 S1134 S1209 T1227 S1244 S1251 S1259 S1274 S1287 T1317 S1341 S1343 T1414 T1435	N100 N186 N211 N256 N376 N384 N448 N585 N816 N898 N919 N995 N1119 N1143 N1158 N1177 N1204 N1285 N1330 N1557 N1768	EGF 様ドメイン C63-C99, C434-C465, C1111-C1148, C393-C429, C103-P131, C1152-C1178 Kelch モチーフ: P550-Q598, P713-T758, E769-8830, G656-Q707 ラミニンEGF 様(ドメイン III および V): C136-C244, C148-C193, C1182-C1231 膜貫通ドメイン: A772-E799, Q1570-L1598 N 末端はサイトソル内 カルシウム結合 EGF 様ドメイン タンパク質/タンパク質 BL01187: C57-A68, C1124-Y1139 ラミニンタイプ EGF 様(LE)ドメインタンパク質 BL01248: C1202-C1214 タイプ III EGF 様シグネチャ PR00011: C209-C227, C447-C465, D165-C193, C447-C465	HMMER-PFAM HMMER-PFAM HMMER-PFAM HMMER-PFAM TMAP BLIMPS- BLOCKS BLIMPS- BLOCKS BLIMPS- PRINTS

10

20

30

【 0 3 8 3 】

【 表 3 - 5 】

表 3-5

SEQ ID NO:	Incyte 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	シグネチャ配列、ドメインおよびモチーフ	分析方法およびデータベース
6			T1508 S1538 T1734		EGF様ドメインシグネチャ 1: C216-C227, C418-C429, C454-C465, C1167-C1178 EGF様ドメインシグネチャ 2: C85-C99, C418-C429, C454-C465, C1133-C1148, C1167-C1178 アスパラギン酸とアスパラギン酸をドロキシル化部位: C76-C87, C1124-C1135 カルシウム結合 EGF様ドメイン シグネチャ: D59-C85 ラミニタンイブ EGF様(LB)ドメインシグネチャ: C164-C193, C216-C244, C1199-C2028, C1280-C1306	MOTIFS MOTIFS MOTIFS MOTIFS MOTIFS
7	2173285CD1	393	S190 S232 S296 S310 S314 S322 S353 S366 T31 T39 T181 T182 T249 T258 T291 T370	N228	タンパク質HOOK HOOK1 HOOK2 PD016676: E11-Y171 HOOK2 タンパク質 PD100271: L165-S322	BLAST- PRODOM BLAST- PRODOM BLAST- PRODOM
8	7487619CD1	311	S65 S87 S192 S263 T136 T288	N3 N63	シグナルペプチド, M1-G39 膜の回貫通受容体(ロトシアン7ファミリー): G39-Y287 膜貫通予測区分 L21-V49, M57-M81, A90-T115, S192-I220, A236-F256, V267-Y287 N-terminus cytosolic Gタンパク質共役受容体タンパク質 BL00237: E231-I257, T279-K295, R89-P128	SPSCAN HMMER-PFAM TMAP BLIMES- BLOCKS

【 0 3 8 4 】
【 表 3 - 6 】

表 3 - 6

SEQ ID NO:	Incyte 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	シグネチャ配列、ドメインおよびモチーフ	分析方法およびデータベース
8					Gタンパク質共役受容体シグネチャ g protein receptor.prf: Y101-T147 嗅覚受容体シグネチャ PR00245: M57-K78, Y176-D190, F237-G252, V271-L282, T288-K302 ヌラコルチン受容体ファミリーシグネチャ PR00534: V49-I61, I125-T136 受容体 嗅覚シグネチャ受容体様 Gタンパク質 共役膜貫通 精蛋白 多重遺伝子ファミリー PD000921: L165-I245 ED149621: V246-L301 Gタンパク質共役受容体 DM00013: P23274 18-306: F26-L301 S51356 18-307: L15-L298 F23266 17-306: L15-L301 P23275 17-306: P19-L301	PROFILESSCAN BLIMPS-PRINTS BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO
9	7487607CD1	318	S49 S67 S87 S193 S229 S297 T276	N5 N42 N65	膜7回貫通受容体(ロドプシンファミリー): G41-Y296 膜貫通予測区分: P21-G41, I48-F68, S95-Y123, I151-N171, T178-I198, L205-I225, G233-K261; N-terminus cytosolic Gタンパク質共役受容体ファミリー BL00237: K90-P129, I207-Y218, E232-M258, T288-K304 g protein receptor.prf: F102-A146 ロドプシン様 GPCR スーパーファミリーシグネチャ PR00237: L26-I50, M59-S80, G104-I126, A140-V161, M199-I222, A237-K261, K278-K304	HMMER-PFAM TWAP BLIMPS-BLOCKS PROFILESSCAN BLIMPS-PRINTS

【 0 3 8 5 】
【 表 3 - 7 】

表 3-7

SEQ ID NO:	Incyte 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	シグネチャ配列、ドメインおよびモチーフ	分析 方法および データベース
9					嗅覚受容体シグネチャ PR00245: M59-S80, F177-D191, S238-G253, I280-M291, S297-L311 受容体 嗅覚タンパク質 受容体様 Gタンパク質 共役膜 貫通 糖蛋白 多重遺伝子ファミリー: PD000921: L166-L245 PD149621: Y246-F315 Gタンパク質 共役受容体 DM00013: P23275 17-306: L17-L311 A57069 15-304: S18-L311 P37067 17-306: L17-L310 S51356 18-307: L17-V307 Gタンパク質 共役受容体シグネチャ: T110-I126	BLIMPS- PRINTS BLAST- PRODOM BLAST-DOMO
10	7487616CD1	311	S65 S87 S192 T136 T288	R3 N63 R303	膜/回貫通受容体(ロトシンファミリー): G39-Y287 膜貫通予測区分: L21-V49, M57-M81, I91-T115, N194- C222, A236-F256, L267-Y287; N-terminus non-cytosolic Gタンパク質 共役受容体タンパク質 BL00237: R89-P128, E231-I257, T279-K295 Gタンパク質 共役受容体シグネチャ g protein receptor.prf: Y101-T147 オプシナル結合部位: opsin.prf: Y258-E311 嗅覚受容体シグネチャ PR00245: M57-K78, Y176-D190, F237-G252, V271-F282, T288-K302	MOTIFS HMMER- PFAM TMAP BLIMPS- BLOCKS PROFILESCAN PROFILESCAN BLIMPS- PRINTS

10

20

30

【 0 3 8 6 】

【 表 3 - 8 】

表 3-8

SEQ ID NO:	Incyte 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	シグネチン配列、ドメインおよびモチーフ	分析手法およびデータベース
10					メラノリチン受容体ファミリーシグネチン PR00534; V49-L61, I125-T136 受容体 嗅覚タンパク質 受容体様 Gタンパク質 共役膜貫通 糖蛋白 多重遺伝子ファミリー PD000921: L165-L245 PD149621: V246-L301 Gタンパク質共役受容体 DMO0013: S51356 18-307: L15-L298 P23274 18-306: F26-L301 P23266 17-306: L15-L301 P23270 18-311: A18-L298	BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO
11	7483204CD1	310	S8 S67 S137 S188 S224 S291	N5 N38 N89 N135	シグナルペプチド: M23-N42 膜7回貫通受容体(ロドシノファミリ): A41-Y290 膜貫通予測区分 P21-I49, Y95-Y123, S138-R165, L197-V225, G233-R261; N末端は細胞質内 Gタンパク質共役受容体タンパク質 BL00237: K90-F129, E232-M258, V282-K298 Gタンパク質共役受容体シグネチン g_protein_receptor.prf: L102-G152 嗅覚受容体シグネチン PR00245: M59-K80, F177-D191, L238-G253, T274-L285, S291-K305 受容体 嗅覚タンパク質 受容体様 Gタンパク質 共役膜貫通 糖蛋白 多重遺伝子ファミリー PD000921: L166-L245 PD149621: S246-L304	HMMER HMMER-PFAM TMAP BLIMPS-BLOCKS PROFILESSCAN BLIMPS-PRINTS BLAST-PRODOM

10

20

30

【 0 3 8 7 】

【 表 3 - 9 】

表 3-9

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基	潜在的リン酸化部位	潜在的カリコシル化部位	シグナルペプチド、ドメインおよびモチーフ	分析手法およびデータベース
11					Gタンパク質共役受容体 DM00013: S51356 18-307: I17-L301 P37067 17-306: I17-L304 S29709 11-299: T18-K303 P30955 18-305: N20-L304 Gタンパク質共役受容体シグネチャ: S110-I126	BLAST-DOMO
12	7472099CD1	316	S109 S230 S306	N43	膜7回貫通受容体(ロドシノプアミン): G42-Y293 膜貫通予測区分: I19-L47, I60-F88, F95-R123, T143-R171, F195-I222, K237-V257, A266-P286; N-terminus CYCLOSOLIC Gタンパク質共役受容体タンパク質 BL00237: R91-F130, F252-H263, E233-S259, P285-R301 Gタンパク質共役受容体シグネチャ g_protein_receptor.pri: F103-R152 膜貫通受容体シグネチャ PR00245: F239-T254, L277-V288, I60-K81, C178-T192 推定上のGタンパク質共役受容体 RA1C PD170483: I246-R301 Gタンパク質共役受容体 DM00013: P23275 17-306: H24-I307 G45774 18-309: P20-L304 P45774 19-309: Y36-L304 P34982 17-305: P20-I307 Gタンパク質共役受容体シグネチャ: M111-I127	MOTIFS HMVER-PFAM TMAP BLIMPS-BLOCKS PROFILESSCAN BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO

10

20

30

【 0 3 8 8 】

【 表 3 - 1 0 】

表 3 - 1 0

SEQ ID NO:	Inocyte 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	シグネチャ配列、ドメインおよびモチーフ	分析手法およびデータベース
13	7485443CD1	318	S49 S67 S87 S193 S297 T276	N5 N42 N65	膜7回貫通受容体(ロトシンファミリー): G41-Y296 膜貫通予測区分: P21-G41, I48-F68, S95-Y123, Y141-V164, C169-C189, G194-I222, G233-K261; N-terminus non-cytosolic Gタンパク質共役受容体タンパク質 BL00237: K90-F129, T207-Y218, E232-M258, T288-K304 Gタンパク質共役受容体シグネチャ g_protein_receptor.prf: F102-G147 嗅覚受容体シグネチャ PRO0245: M59-S80, F177-D191, F238-G253, I280-M291, S297-P311 受容体 嗅覚タンパク質 受容体様 Gタンパク質 共役膜貫通 糖蛋白 多量遺伝子ファミリー PD00921: L166-L245 PDI49621: T246-F315 Gタンパク質共役受容体 DM00013: P23275 17-306: L17-L310 A57069 15-304: S18-L310 S51356 18-307: L17-V107 P37067 17-306: L17-L310	HMER-PFAM TMAP BLIMPS-BLOCKS PROFILESKAN BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO
14	3090414CD1	321	S69 S231 S263 T110 T165 T179	N5	Gタンパク質共役受容体シグネチャ: T110-I126 膜7回貫通受容体(ロトシンファミリー): G43-Y295 膜貫通予測区分: L27-H52, Y62-H90, I94-L119, I143-R167, N197-R225, A239-G267, I270-P287; N-末端はサイトソル内にない	MOTIFS HMER-PFAM TMAP

10

20

30

【 0 3 8 9 】

【 表 3 - 1 1 】

表 3-11

SEQ ID NO:	Incyte 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	シグネチャ配列、ドメインおよびモチーフ	分析手法およびデータベース
14					Gタンパク質共役受容体タンパク質 BL00237: D234-S260, P287-R303, K92-P131 嗅覚受容体シグネチャ PR00245: M61-K82, T179-N193, F240-T255 受容体 嗅覚タンパク質 受容体様 Gタンパク質 共役膜貫通糖蛋白 多重遺伝子ファミリー PD000921: L168-I247 Gタンパク質共役受容体 DM00013: G45774 18-309: P20-S311 P23273 18-306: F33-L310 P23274 18-306: F33-L310 S29708 18-306: S35-L310	BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO
15	7503710CD1	422	S152 S165 S351 S365 S404 T144 T316 T329	N142 N288 N400	Gタンパク質共役受容体シグネチャ: M112-I128 シグナル切断: M1-A18 シグナルペプチド M5-E20, M5-E23, M1-E20, M1-L24, M1-T27, M1-E23 リホリ-ムタンパク質 L1タンパク質 BL01199: L168-P181 潜在的リガンド結合タンパク質 RYA3 PD177882: F86-F261 PD053120: V4-R65 タンパク質前駆体 シグナル糖タンパク質 脂質輸送抗生物質 膜貫通 リポ多糖結合 LBP PD006440: A332-N411, G155-S357, L80-L119 do LIGAND; RYZG5; RYA3 DM05385: S17448 1-473: V4-L347, L237-A422 S17447 1-470: G66-I343, D246-N411, G58-G223, G58-V85	MOTIFS SPSCAN HMMER BLIMPS-BLOCKS BLAST-PRODOM BLAST-PRODOM BLAST-DOMO

【 0 3 9 0 】
【 表 4 - 1 】

10

20

30

表 4 - 1

ポリマレオキド SEQ ID NO./ Ineyte ID/ 配列番号	配列断片
16/71924779CBI/ 2192	1-210, 3-651, 8-539, 17-120, 24-606, 26-235, 31-485, 63-687, 132-555, 132-709, 201-754, 232-464, 244-496, 244-745, 244-904, 266-517, 268-792, 352-648, 352-951, 376-530, 385-937, 387-937, 458-1113, 506-702, 683-813, 697-1289, 701-894, 710-859, 711-755, 711-770, 711-797, 711-943, 711-1010, 711-1095, 711-1096, 711-1097, 711-1099, 711-1101, 711-1104, 711-1189, 720-1366, 736-1146, 763-1525, 789-1323, 815-1131, 830-1207, 835-1016, 854-1375, 858-1769, 891-1125, 891-1389, 891-1410, 891-1479, 915-1393, 927-1338, 947-1197, 960-1338, 982-1246, 984-1258, 984-1631, 1004-1600, 1006-1473, 1018-1282, 1039-1634, 1040-1294, 1041-1344, 1041-1591, 1067-1362, 1067-1603, 1071-1687, 1074-1300, 1076-1738, 1083-1356, 1104-1356, 1111-1554, 1111-1604, 1127-1648, 1144-1768, 1147-1694, 1161-1287, 1167-1498, 1254-1476, 1257-1548, 1257-1773, 1263-1607, 1269-1831, 1294-1558, 1338-1571, 1367-1870, 1396-2192, 1405-1656, 1406-1656, 1411-1599, 1422-1649, 1423-1662, 1444-1748, 1446-1679, 1448-1669, 1501-1768, 1520-1788, 1520-2015, 1580-1783, 1592-1864, 1602-1773, 1607-1834, 1626-1932
17/2319430CBI/ 3614	1-679, 331-627, 331-751, 331-869, 385-703, 444-883, 524-887, 538-764, 544-1198, 554-881, 597-804, 631-837, 631-848, 631-882, 649-888, 749-1042, 749-1271, 781-998, 878-1505, 891-1572, 985-1270, 1046-1508, 1141-1415, 1156-1587, 1238-1516, 1238-1688, 1238-1726, 1238-1731, 1277-1925, 1319-1541, 1368-1872, 1419-1678, 1419-1686, 1535-1998, 1616-1877, 1616-2246, 1643-2151, 1708-2305, 1738-2007, 1765-2020, 1775-2221, 1781-1975, 1803-2335, 1803-2474, 1826-2007, 1838-2359, 1844-2112, 1846-2434, 1853-2333, 1869-2434, 1879-2454, 1925-2176, 1926-2373, 1926-2459, 1927-2262, 1927-2423, 1943-2210, 1961-2603, 1967-2244, 1975-2160, 1975-2188, 2004-2662, 2012-2489, 2019-2399, 2043-2289, 2049-2677, 2059-2694, 2077-2640, 2107-2678, 2113-2385, 2118-2710, 2123-2808, 2137-2818, 2144-2548, 2163-2769, 2178-2808, 2180-2421, 2180-2708, 2218-2363, 2226-2888, 2260-2517, 2260-2533, 2279-2573, 2283-2895, 2303-2867, 2327-2833, 2335-2878, 2374-2756, 2374-2894, 2378-2570, 2380-2757, 2406-2690, 2406-2703, 2416-2876, 2442-2879, 2460-2835, 2464-3080, 2520-2764, 2520-3072, 2535-3005, 2585-3034, 2594-2835, 2595-3253, 2617-2879, 2642-2871, 2658-3206, 2675-2932, 2679-3255, 2695-3106, 2700-2963, 2712-2989, 2726-2954, 2726-3342, 2729-3044, 2734-3018, 2777-3188, 2798-3328, 2804-3035, 2825-3034, 2840-3379, 2853-3066, 2861-3559, 2869-3062, 2874-3134, 2929-3591, 2939-3579, 2946-3569, 2975-3243, 2975-3283, 2975-3512, 2977-3237, 2977-3592, 2978-3458, 2981-3264, 2986-3548, 2988-3241, 2989-3242, 2989-3245, 2989-3577, 2990-3452, 2994-3324, 2995-3262, 2996-3239, 2996-3252, 2998-3263, 3002-3258, 3002-3258, 3002-3258, 3005-3267, 3008-3591, 3016-3324, 3017-3269, 3017-3583, 3017-3591, 3368, 3005-3262, 3005-3267, 3008-3591, 3016-3324, 3017-3269, 3017-3583, 3017-3591,

【 0 3 9 1 】

【 表 4 - 2 】

表 4 - 2

ポリヌクレオチド SEQ ID NO: / Incyte ID / 配列長	配列断片
17	3027-3296, 3027-3552, 3027-3574, 3061-3313, 3063-3308, 3069-3331, 3069-3574, 3202-3464, 3409-3614
18/7291877CB1/ 1585	1-322, 1-583, 1-665, 2-233, 2-480, 71-248, 110-588, 110-748, 270-1027, 461-1027, 681-987, 767-989, 768-1411, 780-987, 988-1100, 1101-1585
19/1218126CB1/ 5618	1-452, 10-268, 21-417, 27-368, 27-591, 28-329, 32-301, 32-302, 34-286, 42-189, 49-661, 49-662, 101-716, 115-195, 141-385, 143-622, 162-695, 169-645, 178-583, 186-719, 273-719, 298-719, 335-719, 349-719, 356-719, 417-719, 450-552, 465-719, 479-2028, 502-719, 504-719, 513-1029, 620-719, 1144-1248, 1144-1289, 1144-1290, 1144-3363, 1417-1658, 1417-1929, 1462-1987, 1492-1659, 1499-2123, 1575-2236, 1710-2098, 1755-2249, 1848-2439, 2115-2495, 2115-2571, 2122-2454, 2181-2784, 2257-2492, 2257-2691, 2300-2547, 2586-3070, 2786-3026, 2786-3044, 2793-3001, 2793-3154, 2796-3154, 2834-3125, 2857-3324, 2880-3049, 3078-3672, 3103-3393, 3216-3755, 3251-3756, 3405-3865, 3554-4155, 3569-4063, 3598-3807, 3625-4034, 3739-4219, 3798-4098, 3798-4278, 3851-4418, 3859-4129, 3886-4145, 3948-4148, 4003-4508, 4293-4562, 4293-4767, 4393-4793, 4357-4781, 4488-4742, 4492-4763, 4493-4776, 4530-5079, 4539-5142, 4543-4775, 4543-4917, 4543-5012, 4572-4747, 4575-4839, 4593-4881, 4611-4836, 4623-4906, 4787-5029, 4792-5062, 4812-5055, 4832-5098, 4840-5090, 4840-5091, 4847-5085, 4936-5235, 4938-5156, 4954-5210, 5010-5275, 5012-5269, 5012-5554, 5012-5616, 5012-5618, 5015-5457
20/7479161CB1/ 1641	1-161, 1-190, 1-319, 1-372, 1-408, 1-419, 1-429, 1-433, 1-446, 1-459, 1-481, 1-485, 1-497, 1-499, 1-520, 1-524, 1-547, 1-550, 1-551, 1-553, 1-567, 1-582, 1-583, 1-602, 1-603, 1-608, 1-610, 1-612, 1-630, 1-660, 1-663, 1-667, 1-678, 1-783, 1-830, 3-634, 3-686, 15-722, 85-380, 135-515, 145-815, 162-780, 178-801, 183-625, 188-733, 195-674, 200-816, 203-907, 207-687, 208-740, 213-470, 224-607, 228-346, 228-457, 260-850, 273-378, 275-836, 280-748, 286-778, 287-918, 317-866, 323-831, 327-822, 338-879, 343-1023, 346-879, 347-886, 367-872, 368-873, 370-891, 403-830, 403-861, 405-834, 411-1016, 421-794, 421-914, 427-831, 427-908, 429-1052, 435-1141, 438-1120, 440-1035, 464-871, 471-830, 479-1035, 483-1048, 493-847, 495-996, 495-1018, 496-960, 497-1035, 501-1179, 504-1035, 509-1035, 522-923, 531-1209, 534-1037, 535-1147, 537-1035, 545-1035, 547-1064, 549-1003, 551-1245, 559-929, 566-1005, 569-786, 571-1035, 576-1035, 580-1121, 586-1273, 597-1035, 606-1094, 607-912, 612-1035, 613-1167, 617-880, 619-1238, 624-990, 626-1035, 628-990, 637-1030, 637-1168, 645-1035, 647-1035, 678-1035, 679-850, 680-1035, 681-1392, 681-1035, 682-935, 685-1233, 690-1349, 695-872, 695-939, 696-1106, 698-1035, 701-1035, 708-799, 708-800,

【 0 3 9 2 】
【 表 4 - 3 】

表 4 - 3

ボイシケイオナド SEQ ID NO./ Incyta ID/配列長	配列断片
20	711-1181, 714-1147, 724-1149, 745-1135, 746-1238, 751-883, 758-1160, 763-1327, 765-1465, 766-1467, 786-1038, 790-985, 791-1234, 791-1465, 799-1280, 801-982, 805-993, 809-1115, 809-1211, 809-1284, 817-1406, 820-1395, 830-1314, 832-1301, 833-1408, 838-1223, 863-1379, 863-1422, 869-1244, 906-1465, 908-1035, 913-1109, 918-1150, 918-1312, 963-1289, 971-1641, 997-1371, 1000-1404, 1001-1460, 1031-1465, 1206-1235, 1206-1262, 1206-1265, 1206-1372, 1206-1274, 1206-1279, 1206-1287, 1206-1299, 1206-1316, 1206-1333, 1206-1336, 1206-1342, 1206-1354, 1206-1355, 1206-1366, 1206-1423, 1206-1452, 1206-1457, 1206-1465, 1206-1466, 1208-1284
21/7722591CE1/ 6056	1-118, 1-130, 1-1697, 4-25, 215-324, 217-275, 777-1047, 779-857, 1068-1335, 1607-2232, 1658-2173, 1753-2341, 1826-2070, 2099-2412, 2099-2699, 2183-2888, 2184-2304, 2295-2855, 2301-2856, 2516-2777, 2609-3040, 2780-2856, 2854-3160, 2854-3240, 2856-3240, 2866-3127, 2971-3454, 3122-3233, 3196-3357, 3221-3616, 3231-3291, 3238-3663, 3306-3366, 3399-3637, 3431-4058, 3464-3549, 3514-3881, 3514-4173, 3671-3977, 3747-4136, 3965-4595, 4096-4843, 4148-4643, 4182-4795, 4286-4784, 4294-4441, 4294-4532, 4294-4548, 4294-4648, 4298-4441, 4412-4640, 4501-5250, 4532-4864, 4553-4648, 4580-4648, 4777-4983, 4777-5122, 4777-5259, 4779-6056, 4784-5021, 4891-5432, 4952-5205, 5526-5637
22/2173285CE1/ 1699	1-441, 1-495, 26-314, 52-726, 58-303, 66-727, 74-608, 164-428, 478-633, 566-757, 566-1066, 572-887, 789-955, 820-1076, 820-1307, 830-1182, 1095-1699, 1097-1699, 1438-1490, 1530-1687
23/7487619CE1/ 1661	1-878, 76-867, 118-867, 128-918, 130-916, 132-867, 154-866, 174-918, 177-912, 219-917, 281-918, 726-1661, 729-953, 736-953, 885-971, 885-1084, 888-1090, 894-1084
24/7487607CE1/ 2033	1-715, 7-715, 20-715, 36-715, 588-1445, 663-850, 1246-1569, 1246-1904, 1246-1958, 1246-1962, 1246-1964, 1246-1965, 1246-1968, 1246-2033
25/7487616CE1/ 1659	1-936, 572-774, 581-768, 581-777, 691-777, 709-921, 709-926, 709-933, 744-1381, 744-1486, 744-1532, 745-1441, 746-1530, 750-1483, 784-1659, 795-1528, 795-1542, 795-1584, 796-1495
26/7483204CE1/ 1175	1-1175, 216-758, 456-665
27/7472099CE1/ 1737	1-210, 39-641, 398-1061, 425-1025, 442-1095, 481-1117, 484-1102, 532-900, 656-1312, 925-982, 1104-1595, 1104-1614, 1106-1737, 1258-1646

【 0 3 9 3 】
【 表 4 - 4 】

表 4 - 4

ホリ双列オキド SEQ ID NO./ Incyte ID/配列長	配列断片
28/7485443CB1/ 972	1-215, 2-215, 16-972, 115-972, 187-377, 334-972, 773-972
29/3090414CB1/ 1592	1-722, 25-722, 40-723, 43-718, 59-719, 60-722, 66-722, 93-722, 153-710, 159-709, 214-463, 402-722, 403-722, 416-778, 416-928, 416-978, 427-1085, 429-715, 441-589, 441-634, 441-711, 441-713, 441-856, 441-877, 441-983, 441-1035, 441-1039, 464-960, 468-1083, 591-1128, 594-907, 594-1286, 594-1318, 607-1094, 619-829, 640-1247, 663-1285, 699-982, 706-919, 739-1002, 755-1295, 758-1318, 774-1060, 774-1309, 783-1041, 789-1205, 789-1395, 806-1365, 807-1338, 842-1099, 847-1374, 901-1365, 917-1380, 941-1380, 954-1592, 983-1376, 987-1348, 1059-1413
30/7503710CB1/ 1480	1-436, 1-439, 1-469, 1-507, 1-557, 1-561, 1-593, 1-620, 1-640, 2-613, 7-171, 7-491, 7-559, 7-577, 7-622, 11-200, 11-329, 11-382, 11-495, 11-534, 11-582, 11-611, 11-618, 11-673, 11-677, 11-1296, 11-1480, 13-644, 13-696, 14-612, 14-699, 155-825, 205-684, 210-826, 217-697, 227-750, 240-356, 270-860, 283-388, 297-928, 348-889, 357-895, 415-843, 459-597, 459-658, 507-1193, 514-1100, 544-1123, 547-1172, 556-1111, 556-1113, 556-1181, 579-795, 581-1138, 586-1118, 608-1068, 622-1205, 622-1262, 627-890, 655-1104, 657-1175, 688-1296, 689-856, 692-1126, 774-1303, 796-1045, 804-1304, 815-1002, 843-1154, 884-1123, 884-1125, 899-1180, 921-1114, 937-1101, 1087-1181

【 0 3 9 4 】
【 表 5 】

10

20

30

表 5

ポリヌクレオチド SEQ ID NO.	Ineyte プロジェクト ID	代表的 ライブラリ
16	71924779CB1	SINITMR01
17	2319430CB1	LUNCAST01
18	7291877CB1	BEAIPER05
19	1218126CB1	PHOSDNV34
20	7479161CB1	NOSEIC02
21	7722591CB1	SEMVIDE01
22	2173285CB1	BRANDIT04
23	7487619CB1	GPCRDFV02
24	7487607CB1	GPCRDFV02
25	7487616CB1	GPCRDFV02
27	7472099CB1	EMARNO02
28	7485443CB1	GPCRDFV02
29	3090414CB1	BRSTNOT19
30	7503710CB1	NOSEIC02

10

20

30

【 0 3 9 5 】

【 表 6 - 1 】

表 6-1

ライブラリ	ハフター	ライブラリの説明
EMARNO02	PBLUESCRIPT	ライブラリは 24 人の男性白人および女性白人の供与者(16 才~70 才)の骨髄から単離された RNA を用いて作成された。
BRAIFER05	PINCY	ライブラリは妊娠 23 週目で死産の左心低形成症胎児(白人男子)から採取した脳組織から単離した RNA を用いて作製した。
BRANDIT04	PINCY	ライブラリはうつ血性心不全で死亡した 88 才の白人女性から採取した松果体組織から単離した RNA を用いて作製した。神経病理は、軽度から中等度のアルツハイマー病、アテローム性動脈硬化症および多発性梗塞を示した。顕微鏡的には、大脳皮質全体にわたってびまん性かつ神経突起状アミロイド斑があった。側頭葉特に嗅内皮質に神経原線維変化があった。前頭葉には散在性、膨張した神経細胞が含まれていた。扁桃体には著しい神経膠症、神経突起状斑と細胞内神経原線維濃縮体が見られた。海馬には、神経突起および散在性斑、また神経原線維濃縮体が見られた。現病にはびまん性かつ限局性神経突起状アミロイド斑および散在性神経原線維濃縮体が見られた。左脳蒼球にグリアーノスで囲まれた囊胞性空洞化部分があった。淡蒼球には散在性細胞内神経原線維濃縮体が含まれていた。尾状核、被殻および中隔側生核にはびまん性斑が含まれていた。右小脳半球に脂質含有マクロファージを有する囊胞性空洞化部分があった。患者の病歴には、真性糖尿病、慢性膵炎、副甲状腺機能亢進、アミロイド心疾患および頻尿が含まれる。
BRSTNGT19	PINCY	ライブラリは 67 才の白人女性の片側性拡大単純乳房切除時に採取した乳房組織から単離した RNA を用いて作製した。関連する腫瘍組織の病理学検査では、残留浸潤性小葉癌が見られた。患者の病歴には抑鬱性疾患、良性大腸新生物、および痔瘻がある。家族歴には脳血管疾患、心臓血管疾患、および肺病がある。
GFCRDPV02	PCR2 - TOPOTA	ライブラリは異なる供与者から得た、プーパされた cDNA を用いて作製された。cDNA は以下から単離された mRNA を用いて作製した。大動脈、小脳、リンパ節、筋肉、扁桃(リンパ肥厚)、膀胱腫瘍(癌度 3 の移行上皮癌)、乳房(増殖性線維囊腫)変化、上皮腺管過形成、精巣腫瘍(胎生期癌)、脾臓、卵巢、副甲状腺、回腸、乳房皮膚、S 状結腸、陰茎。

【 0 3 9 6 】
【 表 6 - 2 】

表 6 - 2

ライブラリ	ペグナー	ライブラリの説明
GPCRDPV02 (cont.)		<p>腫瘍(菌株発育侵襲度4の扁平上皮癌)、胎児肝臓、胎児小腸、胎児肺、胎児皮膚、胎児陰茎、胎児骨、胎児膀胱腫瘍(グレード4の大円形細胞性芽状細胞腫)、卵巣(間質性卵胞莖腫増殖症)、膀胱、膀胱腫瘍(侵襲度3の移行細胞癌)、胃、リンパ節腫瘍(転移性基底様扁平上皮癌)、扁桃腺(反応性リンパ球過形成)、脛骨の骨髄、胎児膵臓、胎児脾臓、子宮ガン、子宮内腔(グレード3の腺癌扁平上皮癌)、精囊、肝臓、大動脈、副腎、リンパ節(転移性グレード3の扁平上皮癌)、舌筋肉、食道、食道癌(侵襲度3の腺癌)、回腸、膀胱、頭蓋骨の軟部組織腫瘍(グレード3の大腸腺癌)、肋骨腫瘍(転移性グレード3の骨肉腫)、肺、心臓、胎盤、胃、脾臓(うつ血性巨脾腫)、子宮、頸部(限局性扁平[上皮]化生を伴う軽度の慢性子宮頸[管]炎)、脾臓腫瘍(悪性リンパ腫)、びまん性大細胞型、B細胞表型(豊富な反応性T細胞と顕著な肉芽腫性炎症を伴う)、膈帯血液単核球、上葉節癌、(グレード3)扁平上皮癌)、子宮内膜(分泌相)、肝臓、肝臓腫瘍(転移性グレード2の神経内分泌癌)、大腸、膈帯血液、Th1細胞、非活性化、膈帯血液、Th2細胞、非活性化、冠狀動脈内皮細胞(未処理)、冠狀動脈平滑筋細胞(各々TNFおよびIL-1、10ng/mlで20時間処理)、膀胱(軽度の慢性膀胱炎)、喉頭癌、乳房皮膚、小腸、胎児前立腺文責線維芽細胞、前立腺上皮細胞(PyEC細胞)、胎児副腎、胎児肝臓、腎臓胎児性形質転換細胞系(5Aza-2デオキシシチジンで72時間処理)、哺乳上皮細胞(HMEC細胞)、末梢血液単球(0時間にIL-10、10ng/mlで処理、1時間後にLPSを5ng/ml添加して、24時間インキュベーション)、末梢血液単球(ランチントン舞踏病)、大腿と腓骨筋肉(ALS)、乳房皮膚線維芽細胞(未処理)(9CIS レチノイン酸 1μMで20時間処理)、乳房皮膚線維芽細胞(TNF-αとIL-1β、各10ng/mlで20時間処理)、胎児肝臓肥満細胞、</p>

【 0 3 9 7 】
【 表 6 - 3 】

表 6 - 3

ライブラリ	ペクター	ライブラリの説明
GPCRDBV02 (cont.)		<p>造血(hIL-6 と hSCF の存在下 18 日間培養したヒト胎児肝臓造血前駆細胞(CD34+幹細胞)から作製された肥満細胞)、大腸上皮層、気管上皮細胞(20%の懸液懸液媒体で 20 時間処理)、リンパ節、プールされた末梢血液単核細胞(未処理)、プールされた脳部分: 線条、淡蒼球および後核(アルツハイマー病)、下垂体、臍帯血液、CD34+由来樹状細胞 (SCF, GM-CSF および TNFα で 13 日間処理)、臍帯血液、CD34+山梨糖状細胞(SCF, GM-CSF & TNFα で 13 日間処理後)、PMD/イオノマイシンで 5 日間処理)、小腸、直腸、骨髄、神経芽細胞腫細胞系(SH-SY5Y 細胞を 6-ヒドロキシドパミン 100 μM で 8 時間処理)、骨髄、神経芽細胞腫細胞系(SH-SY5Y 細胞、未処理) 一人の供与者の脳部分: 扁桃、嗅内皮質、淡蒼球、無名質、線条、背側尾状核、背側被殻、腹側中隔側坐核、舌皮質(海馬前側と後側)、視床、大脳線核、中脳水道周囲灰白質、中脳、黒質および黒状核、松果体(アルツハイマー病)、脂肪細胞(未処理)、前脂肪細胞(ベルオルキニウム増殖因子活性化受容体 γ 作用物質、1 ミクロ M で 4 時間処理)、プールされた前立腺(腺癌腫形成)、プールされた腎臓、脂肪細胞(未処理)、プールされた脂肪細胞(ヒトインリン)で処理)、プールされた腸間膜と腹部脂肪、プールされた副腎、プールされた甲状腺(正常臓器と癌腫性過形成)、プールされた脾臓(正常臓器と特発性血小板減少性紫斑病)、プールされた左乳房、プールされた肺、プールされた鼻(アレルギー)、プールされた食道、正常および腫瘍(侵襲度 3 の腺癌)、プールされた皮膚線維芽細胞(9CIS レチノイン酸で処理されたもの)と TNF-α 及び IL-1 β で処理されたもの)、プールされた胆嚢(胆石症を伴う急性胆嚢炎)、急性胆嚢炎(臨床的水腫)、急性出血性胆石症合併急性胆嚢炎、慢性胆嚢炎および胆石症)、プールされた胎児心臓、(パトリー症候群と胎児死)、プールされた神経原性腫瘍細胞系、SK-N-MC、(神経上皮腫、眼窩上部に転移、未処理)および</p>

【 0 3 9 8 】
【 表 6 - 4 】

表 6 - 4

ライブラリ	ベクター	ライブラリの説明
GPCRDEV02 (cont.)		神経細胞、NT-2 細胞系、(マウスレプチン 1 μg/ml と 9cis レチノイン酸 3.3 μM で 6 日間処理)、プー ルされた卵巣(正常卵巣および毛嚢腫性卵巣)、プー ルされた前立腺、(腺癌腫形成)、プー ルされた精 嚢、プー ルされた小腸、プー ルされた胎児小腸、プー ルされた胃と胎児胃、前立腺上皮細胞、プー ルされた精 巣(正常および卵巣腫)、プー ルされた子宮、プー ルされた子宮嚢(グレード 3 の腺癌上皮癌と平滑筋腫)、 プー ルされた子宮、子宮内膜および子宮筋層、(正常および扁平上皮癌形成と限局性非定型を伴う腺癌性過形 成)、プー ルされた脳 (側頭葉、基底核、小脳と海馬(アルツハイマー病)、およびプー ルされた皮膚。 ライブラリは、脳外傷で死亡した 17 才白人男子の肺組織から単離された RNA を用いて作製された。患者の病歴には、喘息が ある。
LUNGAST01	PSPORT1	ライブラリは、鼻ポリープ組織から単離した RNA を用いて作製した。 この大型分画のライブラリは、鼻ポリープ組織から単離した RNA を用いて作製した。 ライブラリの作製には、111 名のドナーからプー ルした cDNA を用いた。cDNA の生成には、下記の組織から単離した mRNA を 用いた。まず、プー ルした骨格筋組織は、21~57 才の突然死した 10 名の白人男女のドナーから切除された。また、プー ルし た甲状腺組織は、18~32 才の突然死した 9 名の白人男女のドナーから切除された。プー ルした胎児肝臓組織は、18-24 週目の自然流産した 32 体の白人男女胎児から切除された。プー ルした胎児腎臓組織は、20-33 週目の自然流産した 59 体の白人男女胎児から切除された。最後に胎児脳組織は、或る 23 週齢白人男児で胎児死亡のものから切除した。
NOSEDI002 PHOSDNV34	PSPORT1 PCR2-TOPOTA	

10

20

30

【 0 3 9 9 】
【 表 6 - 5 】

表 6 - 5

ライブラリ	ベクター	ライブラリの説明
SEMVIDE01	PCDNA2.1	この5'UTRの無作為ライブラリは、63才の白人男性から閉鎖型前立腺生検、前立腺全摘出術および局所リンパ前切除時に摘出した精囊組織から単離したRNAを用いて作製した。関連腫瘍組織の病理は前立腺右側に Gleason グレード 2+3 の腺癌を示した。腺癌腫瘍性過形成が見られた。患者は前立腺癌、前立腺特異抗原の上昇および前立腺肥大はアブライドロ蒸留機摘出術があった。患者歴には腎結石、外因性喘息、良性肺腫瘍、腎痛、膝軟、およびニコチン中毒経験がある。以前の手術としてはアブライドロ蒸留機摘出術があった。患者の服用薬剤には Ventolin と Vanceril がある。家族歴は、母親にアテローム硬化型冠動脈疾患および急性心筋梗塞、父親にアテローム硬化型冠動脈疾患および急性心筋梗塞、祖父母に胃癌と外因性喘息がある。
SINITWIR01	PCDNA2.1	このランダムライブラリは70才白人女性の右半結腸切除、開腹式肝生検、S 状結腸鏡検査、結腸鏡検査および永久的結腸造設術時に摘出された回腸組織から単離された mRNA を用いて作製した。一致する腫瘍の病理は侵襲度 2 の腺癌で、回盲弁の Zcm 遠位部に潰瘍性瘻の形成を示した。患者の病歴には、悪性乳房腫瘍、II 型糖尿病、高脂血症、ウイルス性肝炎、不特定甲状腺疾患、骨関節炎、悪性皮膚腫瘍、欠乏性貧血および正常分娩が含まれている。家族の病歴には、乳癌、アテローム硬化型冠動脈疾患、良性の高血圧、脳血管疾患、卵巣癌、および高脂質血症が見られた。

10

20

30

【 0 4 0 0 】
【 表 7 - 1 】

表 7 - 1

プログラム	説明	引用文献	パラメータ 閾値
ABI FACTURA	核酸配列においてベクター配列を除くして、特定の塩基をマスクするプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finderは、アミノ酸または核酸配列の比較、注釈付けに有用である。	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler	核酸配列を構築するプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	Basic Local Alignment Search Tool (BLAST) はアミノ酸および核酸配列の配列類似性検索に有用であり、BLAST は blastp、blastn、blastx、tblastn および tblastx の 5 つの機能がある。	Altschul, S.F. 他 (1990) J Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) Nucleic Acids Res 25:3389-3402	ESTs: 確率値=1.0E-8 以下 完全長配列: 確率値1.0E-10 以下
FASTA	問合せ配列と同種の配列群配との類似性を検索する Pearson および Lipman アルゴリズム。FASTA は、最小5つの機能(fasta、fasta、fastx、fastxおよび ssearch)がある。	Pearson, W. R. and D. J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W. R. (1990) Methods Enzymol. 183:63-98; Smith, T. F. 及び M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E 値=1.0E-6 構築されたESTs: fasta I 同一性=95%以上 一致した長さ=200塩基以上 fastx E 値=1.0E-8 以下 完全長配列: fastx スコア=100以上
BLIMPS	配列をBLOCKS、PRINTS、DOMO、PRODOMおよびPFAM データベースの配列と対応させて選別する。配列相関性および構造的フィンガープリント領域を検索するBLOCKS IMPROVED サーチャー。	Henikoff, S. 及び J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. 及び S. Henikoff (1996) Methods Enzymol. 266:88-105; 及び Attwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37:417-424	確率値=1.0E-3以下
HMMER	PFAM のようなタンパク質ファミリーコンセンサス配列の隠れマルコフモデル(HMM)に基づいたデータベースに対して問合せ配列を検索するアルゴリズム。	Krogh, A. 他(1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. 他 (1988) Nucleic Acids Res. 26:320-322; Durbin, R. 他 (1998) Our World View, in a Nutshell, Cambridge Univ. Press, 1-350頁	PFAM ヒット: 確率値=1.0E-3 以下 シグナルペプチド ヒット: スコア=0 以上

10

20

30

【 0 4 0 1 】
【 表 7 - 2 】

表 7 - 2

プログラム	説明	引用文献	パラメータ関値
ProfileScan	Prositeで定義された配列パターンと一致するタンパク質配列の構造的モチーフおよび配列モチーフを検索するアルゴリズム。	Gribstkov, M. 他 (1988) CABIOS 4:61-66; Gribstkov, M. 他 (1989) Methods Enzymol 183 146-159, Bairoch, A. 他 (1997) Nucleic Acids Res. 25: 217-221.	標準化された黄のスコアとその特定の Prosite モチーフに対する GCG 指定の "HIGH" 値 通常、スコア=1.4-2.1
Phred	高い感度と精度で自動配列決定機の出る塩基配列を調べる塩基配列出力アルゴリズム。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. 及び P. Green (1998) Genome Res 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的なインプライメンテーションに基づく SWAT や CrossMatch を含む Phils Revised Assembly プログラムで、配列相同性の検索や DNA 配列の構築に有用である。	Smith, T. F. 及び M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T. F. 及び M. S. Waterman (1981) J. Mol. Biol. 147: 195-197, and Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=56以上
Conseq	Phrap で構築したものの表示、編集用グラフィカルツール。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキキャンして、分泌シグナルペプチドの存在を調べる重み付けマトリックス解析プログラム。	Nielson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J. M. 及び S. Audic (1997) CABIOS 12:431-439.	スコア=3.5以上
TMAP	加重マトリックスを用いて蛋白質配列での膜貫通セグメントの描写および配向を決定するプログラム。	Persson, B. および P. Argos (1994) J. Mol. Biol. 237:182-192, Persson, B. 及び P. Argos (1996) Protein Sci 5 363-371.	
TMHMMER	隠れMarkov モデル(HMM)を使って蛋白質配列の膜貫通セグメントを描写し、配向を決定するプログラム。	Sonnhammer, E.L. 他 (1998) Proc Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow 他, eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, 175-182頁.	
Motifs	Prositeで定義された配列と一致したパターンのアミノ酸配列を検索するプログラム。	Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, 第9版, M51-59頁, Genetics Computer Group, Madison, WI	

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
25 July 2002 (25.07.2002)

PCT

(10) International Publication Number
WO 02/057454 A2(51) International Patent Classification: C12N 15/12,
C07K 14/705, C12N 5/10, A01K 67/027, C07K 16/28,
C12Q 1/68, A61K 38/17, G01N 33/68, A61K 39/395[US/US]; 900 Olive Street, Menlo Park, CA 94025 (US).
GANDHI, Ameena, R. [US/US]; 705 5th Avenue, San
Francisco, CA 94118 (US); RICHARDSON, Thomas,
W. [US/US]; 616 Canyon Road, #107, Redwood City, CA
94062 (US); KHAN, Farrah, A. [IN/US]; 9445 Harrison
Street, Des Plaines, IL 60016 (US); LU, Yan [CN/US];
3885 Corrina Way, Palo Alto, CA 94303 (US); SWAR-
NAKAR, Anita [CA/US]; 8 Locksley Avenue, #5D, San
Francisco, CA 94122 (US); RAMKUMAR, Jayalaxmi
[IN/US]; 34359 Maybird Circle, Fremont, CA 94555
(US); NGUYEN, Daniel, B. [US/US]; 1403 Ridgewood
Drive, San Jose, CA 95118 (US); GRAUL, Richard, C.
[US/US]; 682-29th Avenue, San Francisco, CA 94121
(US); LU, Dyung, Aina, M. [US/US]; 233 Coy Drive, San
Jose, CA 95123 (US).

(21) International Application Number: PCT/US02/01339

(22) International Filing Date: 16 January 2002 (16.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/262,838 19 January 2001 (19.01.2001) US
60/265,927 2 February 2001 (02.02.2001) US
60/271,196 23 February 2001 (23.02.2001) US
60/274,540 9 March 2001 (09.03.2001) US
60/334,179 28 November 2001 (28.11.2001) US(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(71) Applicant (for all designated States except US): INCYTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): LEE, Ernestine,
A. [US/US]; 624 Kaits Street, Albany, CA 94706 (US);
WALLA, Narinder, K. [US/US]; 890 Davis Street, #205,
San Leandro, CA 94577 (US); BAUGHN, Mariah, R.
[US/US]; 14244 Santiago Road, San Leandro, CA 94577
(US); AZIMZAI, Yalda [US/US]; 5518 Boulder Canyon
Drive, Castro Valley, CA 94552 (US); TANG, Y., Tom
[US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US);
YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA
94087 (US); THANGAVELU, Kavitha [IN/US]; 1950
Montecito Avenue, #23, Mountain View, CA 94043 (US);
XU, Yuming [US/US]; 1739 Walnut Drive, Mountain
View, CA 94040 (US); ARVIZU, Chandra [US/US];
490 Sherwood Way, #1, Menlo Park, CA 94025 (US);
WARREN, Bridget, A. [US/US]; 1030 Parkwood Drive,
#2, Cupertino, CA 95014 (US); YAO, Monique, G.
[US/US]; 1189 Woodgate Drive, Carmel, IN 46033 (US);
AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane,
Brisbane, CA 94005 (US); HAFALIA, April, J., A.
[US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054
(US); ELLIOTT, Vicki, S. [US/US]; 3770 Polton Place
Way, San Jose, CA 95121 (US); KALLICK, Deborah, A.(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, GR, GU, HT, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MY, NZ, NI, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SL, SM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KI, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CI, CG, CL, CM, GA, GN, GQ, GW, ML, MR,
NI, SN, TD, TG).

Published:

— without international search report and to be republished
upon receipt of that report
— with sequence listing part of description published sepa-
rately in electronic form and available upon request from
the International Bureau.For 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: RECEPTORS AND MEMBRANE-ASSOCIATED PROTEINS

(57) Abstract: The invention provides human receptors and membrane-associated proteins (REMAP) and polynucleotides which identify and encode REMAP. 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 REMAP.



WO 02/057454 A2

WO 02/057454

PCT/US02/01339

RECEPTORS AND MEMBRANE-ASSOCIATED PROTEINS**TECHNICAL FIELD**

5 This invention relates to nucleic acid and amino acid sequences of receptors and membrane-associated proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative, autoimmune/inflammatory, neurological, metabolic, developmental, and endocrine disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of receptors and membrane-associated proteins.

BACKGROUND OF THE INVENTION

10 Signal transduction is the general process by which cells respond to extracellular signals. Signal transduction across the plasma membrane begins with the binding of a signal molecule, e.g., a hormone, neurotransmitter, or growth factor, to a cell membrane receptor. The receptor, thus activated, triggers an intracellular biochemical cascade that ends with the activation of an intracellular target molecule, such as a transcription factor. This process of signal transduction regulates all types of cell functions including cell proliferation, differentiation, and gene transcription.

15 Biological membranes surround organelles, vesicles, and the cell itself. Membranes are highly selective permeability barriers made up of lipid bilayer sheets composed of phosphoglycerides, fatty acids, cholesterol, phospholipids, glycolipids, proteoglycans, and proteins. Membranes contain ion pumps, ion channels, and specific receptors for external stimuli which transmit biochemical signals across the membranes. These membranes also contain second messenger proteins which interact with these pumps, channels, and receptors to amplify and regulate transmission of these signals.

Plasma Membrane Proteins

25 Plasma membrane proteins (MPs) are divided into two groups based upon methods of protein extraction from the membrane. Extrinsic or peripheral membrane proteins can be released using extremes of ionic strength or pH, urea, or other disruptors of protein interactions. Intrinsic or integral membrane proteins are released only when the lipid bilayer of the membrane is dissolved by detergent.

30 The majority of known integral membrane proteins are transmembrane proteins (TM) which are characterized by an extracellular, a transmembrane, and an intracellular domain. TM domains are typically comprised of 15 to 25 hydrophobic amino acids which are predicted to adopt an α -helical conformation. TM proteins are classified as bitopic (Types I and II) and polytopic (Types III and IV) (Singer, S.J. (1990) *Annu. Rev. Cell Biol.* 6:247-96). Bitopic proteins span the membrane once while polytopic proteins contain multiple membrane-spanning segments. TM proteins carry out a variety of

35

WO 02/057454

PCT/US02/01339

important cellular functions, including acting as cell-surface receptor proteins involved in signal transduction. These functions are represented by growth and differentiation factor receptors, and receptor-interacting proteins such as *Drosophila* pecanex and frizzled proteins, LIV-1 protein, NF2 protein, and GNS1/SUR4 eukaryotic integral membrane proteins. TM proteins also act as transporters of ions or metabolites, such as gap junction channels (connexins), and ion channels, and as cell anchoring proteins, such as lectins, integrins, and fibronectins. TM proteins are found in vesicle organelle-forming molecules, such as caveolins; or cell recognition molecules, such as cluster of differentiation (CD) antigens, glycoproteins, and mucins.

Many MPs contain amino acid sequence motifs that serve to localize proteins to specific subcellular sites. Examples of these motifs include PDZ domains, KDEL, RGD, NGR, and GSL sequence motifs, von Willebrand factor A (vWFA) domains, and EGF-like domains. RGD, NGR, and GSL motif-containing peptides have been used as drug delivery agents in targeted cancer treatment of tumor vasculature (Arap, W. et al. (1998) *Science*, 279:377-380). Furthermore, MPs may also contain amino acid sequence motifs that serve to interact with extracellular or intracellular molecules, such as carbohydrate recognition domains (CRD).

Chemical modification of amino acid residue side chains alters the manner in which MPs interact with other molecules, for example, phospholipid membranes. Examples of such chemical modifications to amino acid residue side chains are covalent bond formation with glycosaminoglycans, oligosaccharides, phospholipids, acetyl and palmitoyl moieties, ADP-ribose, phosphate, and sulphate groups.

RNA encoding membrane proteins may have alternative splice sites which give rise to proteins encoded by the same gene but with different messenger RNA and amino acid sequences. Splice variant membrane proteins may interact with other ligand and protein isoforms.

Receptors

The term receptor describes proteins that specifically recognize other molecules. The category is broad and includes proteins with a variety of functions. The bulk of receptors are cell surface proteins which bind extracellular ligands and produce cellular responses in the areas of growth, differentiation, endocytosis, and immune response. Other receptors facilitate the selective transport of proteins out of the endoplasmic reticulum and localize enzymes to particular locations in the cell. The term may also be applied to proteins which act as receptors for ligands with known or unknown chemical composition and which interact with other cellular components. For example, the steroid hormone receptors bind to and regulate transcription of DNA.

Cell surface receptors are typically integral plasma membrane proteins. These receptors recognize hormones such as catecholamines; peptide hormones; growth and differentiation factors; small peptide factors such as thyrotropin-releasing hormone; galanin, somatostatin, and tachykinins;

WO 02/057454

PCT/US02/01339

and circulatory system-borne signaling molecules. Cell surface receptors on immune system cells recognize antigens, antibodies, and major histocompatibility complex (MHC)-bound peptides. Other cell surface receptors bind ligands to be internalized by the cell. This receptor-mediated endocytosis functions in the uptake of low density lipoproteins (LDL), transferrin, glucose- or mannose-terminal glycoproteins, galactose-terminal glycoproteins, immunoglobulins, phosphovitellogenins, fibrin, proteinase-inhibitor complexes, plasminogen activators, and thrombospondin (Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York NY, p. 723; Mikhailenko, I. et al. (1997) *J. Biol. Chem.* 272:6784-6791).

Receptor Protein Kinases

Many growth factor receptors, including receptors for epidermal growth factor, platelet-derived growth factor, fibroblast growth factor, as well as the growth modulator α -thrombin, contain intrinsic protein kinase activities. When growth factor binds to the receptor, it triggers the autophosphorylation of a serine, threonine, or tyrosine residue on the receptor. These phosphorylated sites are recognition sites for the binding of other cytoplasmic signaling proteins. These proteins participate in signaling pathways that eventually link the initial receptor activation at the cell surface to the activation of a specific intracellular target molecule. In the case of tyrosine residue autophosphorylation, these signaling proteins contain a common domain referred to as a Src homology (SH) domain. SH2 domains and SH3 domains are found in phospholipase C- γ , PI-3-K p85 regulatory subunit, Ras-GTPase activating protein, and pp60^{src} (Lowenstein, E.J. et al. (1992) *Cell* 70:431-442). The cytokine family of receptors share a different common binding domain and include transmembrane receptors for growth hormone (GH), interleukins, erythropoietin, and prolactin.

Other receptors and second messenger-binding proteins have intrinsic serine/threonine protein kinase activity. These include activin/TGF- β /BMP superfamily receptors, calcium- and diacylglycerol-activated/phospholipid-dependant protein kinase (PK-C), and RNA-dependant protein kinase (PK-R). In addition, other serine/threonine protein kinases, including nematode Twitchin, have fibronectin-like, immunoglobulin C2-like domains.

G-protein coupled receptors

The G-protein coupled receptors (GPCRs), encoded by one of the largest families of genes yet identified, play a central role in the transduction of extracellular signals across the plasma membrane. GPCRs have a proven history of being successful therapeutic targets.

GPCRs are integral membrane proteins characterized by the presence of seven hydrophobic transmembrane domains which together form a bundle of antiparallel alpha (α) helices. GPCRs range in size from under 400 to over 1000 amino acids (Strosberg, A.D. (1991) *Eur. J. Biochem.* 196:1-10; Coughlin, S.R. (1994) *Curr. Opin. Cell Biol.* 6:191-197). The amino-terminus of a GPCR is extracellular, is of variable length, and is often glycosylated. The carboxy-terminus is cytoplasmic

WO 02/057454

PCT/US02/01339

and generally phosphorylated. Extracellular loops alternate with intracellular loops and link the transmembrane domains. Cysteine disulfide bridges linking the second and third extracellular loops may interact with agonists and antagonists. The most conserved domains of GPCRs are the transmembrane domains and the first two cytoplasmic loops. The transmembrane domains account, in part, for structural and functional features of the receptor. In most cases, the bundle of α helices forms a ligand-binding pocket. The extracellular N-terminal segment, or one or more of the three extracellular loops, may also participate in ligand binding. Ligand binding activates the receptor by inducing a conformational change in intracellular portions of the receptor. In turn, the large, third intracellular loop of the activated receptor interacts with a heterotrimeric guanine nucleotide binding (G) protein complex which mediates further intracellular signaling activities, including the activation of second messengers such as cyclic AMP (cAMP), phospholipase C, and inositol triphosphate, and the interaction of the activated GPCR with ion channel proteins. (See, e.g., Watson, S. and S. Arkininstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 2-6; Bolander, F.F. (1994) Molecular Endocrinology, Academic Press, San Diego CA, pp. 162-176; Baldwin, J.M. (1994) Curr. Opin. Cell Biol. 6:180-190.)

GPCRs include receptors for sensory signal mediators (e.g., light and olfactory stimulatory molecules); adenosine, γ -aminobutyric acid (GABA), hepatocyte growth factor, melanocortins, neuropeptide Y, opioid peptides, opsins, somatostatin, tachykinins, vasoactive intestinal polypeptide family, and vasopressin; biogenic amines (e.g., dopamine, epinephrine and norepinephrine, histamine, glutamate (metabotropic effect), acetylcholine (muscarinic effect), and serotonin); chemokines; lipid mediators of inflammation (e.g., prostaglandins and prostanoids, platelet activating factor, and leukotrienes); and peptide hormones (e.g., bombesin, bradykinin, calcitonin, C5a anaphylatoxin, endothelin, follicle-stimulating hormone (FSH), gonadotropic-releasing hormone (GnRH), neurokinin, and thyrotropin-releasing hormone (TRH), and oxytocin). GPCRs which act as receptors for stimuli that have yet to be identified are known as orphan receptors.

The largest family of GPCRs consists of the rhodopsin-like GPCRs, which transmit diverse extracellular signals including hormones, neurotransmitters, and light. Rhodopsin is a photosensitive GPCR found in animal retinas. In vertebrates, rhodopsin molecules are embedded in membranous stacks found in photoreceptor (rod) cells. Each rhodopsin molecule responds to a photon of light by triggering a decrease in cGMP levels which leads to the closure of plasma membrane sodium channels. In this manner, a visual signal is converted to a neural impulse. Other rhodopsin-like GPCRs are directly involved in responding to neurotransmitters. These GPCRs include the receptors for adrenaline (adrenergic receptors), acetylcholine (muscarinic receptors), adenosine, galanin, and glutamate (N-methyl-D-aspartate/NMDA receptors). (Reviewed in Watson, S. and S. Arkininstall (1994) The G-Protein Linked Receptor Facts Book, Academic Press, San Diego CA, pp. 7-9, 19-22,

WO 02/057454

PCT/US02/01339

32-35, 130-131, 214-216, 221-222; Habert-Ortoli, E. et al. (1994) Proc. Natl. Acad. Sci. USA 91:9780-9783.)

The largest subfamily of GPCRs, the olfactory receptors, are also members of the rhodopsin-like GPCR family. These receptors function by transducing odorant signals. Numerous distinct olfactory receptors are required to distinguish different odors. Each olfactory sensory neuron expresses only one type of olfactory receptor, and distinct spatial zones of neurons expressing distinct receptors are found in nasal passages. For example, the RA1c receptor which was isolated from a rat brain library, has been shown to be limited in expression to very distinct regions of the brain and a defined zone of the olfactory epithelium (Raming, K. et al. (1998) Receptors Channels 6:141-151).

The olfactory mucosa also appears to possess an additional group of odorant-binding proteins which recognize and bind separate classes of odorants. For example, cDNA clones from rat have been isolated which correspond to mRNAs highly expressed in olfactory mucosa but not detected in other tissues. The proteins encoded by these clones are homologous to proteins that bind lipopolysaccharides or polychlorinated biphenyls, and the different proteins appear to be expressed in specific areas of the mucosal tissue. These proteins are believed to interact with odorants before or after specific recognition by odorant receptors, perhaps acting as selective signal filters (Dear, T.N. et al. (1991) EMBO J. 10:2813-2819; Vogt, R.G. et al. (1991) J. Neurobiol. 22:74-84).

Members of the secretin-like GPCR subfamily have as their ligands peptide hormones such as secretin, calcitonin, glucagon, growth hormone-releasing hormone, parathyroid hormone, and vasoactive intestinal peptide. For example, the secretin receptor responds to secretin, a peptide hormone that stimulates the secretion of enzymes and ions in the pancreas and small intestine (Watson, supra, pp. 278-283). Secretin receptors are about 450 amino acids in length and are found in the plasma membrane of gastrointestinal cells. Binding of secretin to its receptor stimulates the production of cAMP.

Examples of secretin-like GPCRs implicated in inflammation and the immune response include the EGF module-containing, mucin-like hormone receptor (Emr1) and CD97 receptor proteins. CD97 is predominantly expressed in leukocytes and is markedly upregulated on activated B and T cells (McKnight, A.J. and S. Gordon (1998) J. Leukoc. Biol. 63:271-280). These GPCRs are members of the recently characterized EGF-TM7 receptors subfamily. These seven transmembrane hormone receptors exist as heterodimers *in vivo* and contain between three and seven potential calcium-binding EGF-like motifs. The EGF motif is about forty amino acid residues in length and includes six conserved cysteine residues, and a calcium-binding site near the N-terminus of the signature sequence. Post-translational hydroxylation of aspartic acid or asparagine residues has been associated with EGF-like domains in several proteins (Prosite PDOC00010 Aspartic acid and asparagine hydroxylation site).

WO 02/057454

PCT/US02/01339

A number of proteins that contain calcium-binding EGF-like domain signature sequences are involved in growth and differentiation. Examples include bone morphogenic protein 1, which induces the formation of cartilage and bone; crumbs, which is a *Drosophila* epithelial development protein; Notch and a number of its homologs, which are involved in neural growth and differentiation, and transforming growth factor beta-1 binding protein (Expasy PROSITE document PDOC00913; 5 Soler, C. and Carpenter, G., in Nicola, N.A. (1994) *The Cytokine Facts Book*, Oxford University Press, Oxford, UK, pp 193-197). EGF-like domains mediate protein-protein interactions for a variety of proteins. For example, EGF-like domains in the ECM glycoprotein fibulin-1 have been shown to mediate both self-association and binding to fibronectin (Tran, H. et al. (1997) *J. Biol. Chem.* 10 272:22600-22606). Point mutations in the EGF-like domains of ECM proteins have been identified as the cause of human disorders such as Marfan syndrome and pseudoachondroplasia (Maurer, P. et al. (1996) *Curr. Opin. Cell Biol.* 8:609-617).

GPCR mutations, which may cause loss of function or constitutive activation, have been associated with numerous human diseases (Coughlin, *supra*). For instance, retinitis pigmentosa may 15 arise from mutations in the rhodopsin gene. Furthermore, somatic activating mutations in the thyrotropin receptor have been reported to cause hyperfunctioning thyroid adenomas, suggesting that certain GPCRs susceptible to constitutive activation may behave as protooncogenes (Parma, J. et al. (1993) *Nature* 365:649-651). GPCR receptors for the following ligands also contain mutations associated with human disease: luteinizing hormone (precocious puberty); vasopressin V₂ (X-linked 20 nephrogenic diabetes); glucagon (diabetes and hypertension); calcium (hyperparathyroidism, hypocalcemia, hypercalcemia); parathyroid hormone (short limbed dwarfism); β_3 -adrenoceptor (obesity, non-insulin-dependent diabetes mellitus); growth hormone releasing hormone (dwarfism); and adrenocorticotropin (glucocorticoid deficiency) (Wilson, S. et al. (1998) *Br. J. Pharmacol.* 125:1387-1392; Stadel, J.M. et al. (1997) *Trends Pharmacol. Sci.* 18:430-437). GPCRs are also 25 involved in depression, schizophrenia, sleeplessness, hypertension, anxiety, stress, renal failure, and several cardiovascular disorders (Horn, F. and G. Vriend (1998) *J. Mol. Med.* 76:464-468).

In addition, within the past 20 years several hundred new drugs have been recognized that are directed towards activating or inhibiting GPCRs. The therapeutic targets of these drugs span a wide range of diseases and disorders, including cardiovascular, gastrointestinal, and central nervous system 30 disorders as well as cancer, osteoporosis and endometriosis (Wilson, *supra*; Stadel, *supra*). For example, the dopamine agonist L-dopa is used to treat Parkinson's disease, while a dopamine antagonist is used to treat schizophrenia and the early stages of Huntington's disease. Agonists and antagonists of adrenoceptors have been used for the treatment of asthma, high blood pressure, other cardiovascular disorders, and anxiety; muscarinic agonists are used in the treatment of glaucoma and 35 tachycardia; serotonin 5HT_{1D} antagonists are used against migraine; and histamine H₁ antagonists

WO 02/057454

PCT/US02/01339

are used against allergic and anaphylactic reactions, hay fever, itching, and motion sickness (Horn, supra).

Nuclear Hormone Receptors

The nuclear hormone receptors, also known as the nuclear receptors or the intracellular
5 receptors, constitute a protein superfamily whose members are both receptors and transcriptional
regulators. Nuclear hormone receptors rely on both their receptor function and their transcriptional
regulatory function to affect a broad array of biological processes, including development,
homeostasis, cell proliferation, and cell differentiation. (Reviewed in Mangelsdorf, D.J. et al. (1995)
Cell 83:835-840; Wen, D.X. and D.P. McDonnell (1995) Curr. Opin. Biotechnol. 6:582-589;
10 Perimann, T. and R.M. Evans (1997) Cell 90:391-397; Tenbaum, S. and A. Baniahmad (1997) Int. J.
Biochem. Cell Biol. 29:1325-1341; Moras, D. and H. Gronemeyer (1998) Curr. Opin. Cell Biol.
10:384-391; Willy, P.J. and D.J. Mangelsdorf (1998) in: Hormones and Signaling (ed: B.W.
O'Malley) vol. 1, Academic Press, San Diego CA, pp. 307-358; Weatherman, R.V. et al. (1999)
Annu. Rev. Biochem. 68:559-581.)

15 Nuclear hormone receptors as receptors

Generally, the term receptor describes a protein that specifically recognizes other molecules.
As receptors, nuclear hormone receptors specifically recognize and bind to their cognate ligands.
Although nuclear hormone receptors are located intracellularly, many receptors are extracellular cell
surface proteins which bind extracellular ligands. Such extracellular receptors produce cellular
20 responses affecting growth, differentiation, endocytosis, and the immune response. Other receptors
facilitate the selective transport of proteins out of the endoplasmic reticulum and localize enzymes to
particular regions of the cell. Transcriptional regulation by nuclear hormone receptors, propagation
of cellular signals by extracellular receptors, and transport and localization of proteins by other
receptors, all rely upon specific interactions between the receptors and a variety of cellular
25 components. In many cases, the identity of the cognate ligand to which a receptor binds is unknown.
Such receptors are termed orphan receptors. This term also applies to those nuclear hormone
receptors which carry out their transcriptional regulatory functions without binding any ligands.

Nuclear hormone receptors as transcriptional regulators

Multicellular organisms are comprised of diverse cell types that differ dramatically both in
30 structure and function. The identity of a cell is determined by its characteristic pattern of gene
expression, and different cell types express overlapping but distinctive sets of genes throughout
development. Spatial and temporal regulation of gene expression is critical for the control of cell
proliferation, cell differentiation, apoptosis, and other processes that contribute to organismal
development. As transcriptional regulators, nuclear hormone receptors play key roles in controlling
35 these fundamental biological processes. Other transcriptional regulators affect gene expression in

WO 02/057454

PCT/US02/01339

response to extracellular signals that mediate cell-cell communication and that coordinate the activities of different cell types.

In general, transcriptional regulators such as nuclear hormone receptors initiate, activate, repress, or terminate gene transcription by binding to the promoter, enhancer, and upstream regulatory regions of a gene in a sequence-specific manner. However, some transcriptional regulators bind regulatory elements within or downstream of a gene's coding region. Transcriptional regulatory proteins may bind to a specific region of DNA singly, or in a complex with other accessory factors. (Reviewed in Lewin, B. (1990) in: Genes IV, Oxford University Press, New York NY, and Cell Press, Cambridge MA, pp. 554-570.)

10 Mechanism of nuclear hormone receptor function

In the unliganded state, a nuclear hormone receptor exists in association with a multiprotein complex of chaperones, including heat shock proteins such as hsp90 and immunophilins such as hsp56. These chaperones maintain the ligand-free receptor in an inactive state which is amenable to binding of free ligand, and prevent the ligand-free receptor from translocating to the nucleus. Upon activation by its cognate ligand, the receptor may form a homodimer or heterodimer which translocates to the nucleus, binds to specific DNA sequences, and exerts its transcriptional regulatory function. In order to effectively carry out its regulatory roles, an activated nuclear hormone receptor dissociates from a histone deacetylase-containing corepressor complex and associates with a histone acetyltransferase-containing coactivator complex (Xu, L. et al. (1999) *Curr. Opin. Genet. Dev.* 9:140-147). The association of the activated receptor with coactivator proteins results in remodeling of chromatin so that it adopts an open transcriptionally active state, providing access to the transcriptional regulatory elements of the activated nuclear receptor (Lemon, B.D. and L.P. Freedman (1999) *Curr. Opin. Genet. Dev.* 9:499-504).

20 Structure of nuclear hormone receptors

Nuclear hormone receptors function as signal transducers by converting hormonal signals into transcriptional responses. In general, nuclear hormone receptors consist of a variable amino-terminal domain, a highly conserved DNA-binding domain, and a conserved C-terminal ligand-binding domain. In the steroid-binding nuclear hormone receptors, the amino-terminal domain harbors a trans-activation element termed AF-1. Some nuclear hormone receptors also contain a trans-activation element in the ligand-binding domain termed AF-2. The DNA-binding and ligand-binding domains of nuclear hormone receptors may contain dimerization elements, and the DNA-binding domain may contain a nuclear localization signal (Weatherman, R.V. et al. (1999) *Annu. Rev. Biochem.* 68:559-581).

The DNA-binding domain of nuclear hormone receptors is composed of two zinc finger motifs which mediate recognition of specific DNA sequences. A zinc finger motif contains

WO 02/057454

PCT/US02/01339

periodically spaced cysteine and histidine residues which coordinate Zn²⁺. Examples of this sequence pattern include the C2H2-type, C4-type, and C3HC4-type ("RING" finger) zinc fingers, and the PHD domain (Lewin, supra; Aasland, R. et al. (1995) Trends Biochem. Sci. 20:56-59). A zinc finger motif contains an α helix and an antiparallel β sheet whose proximity and conformation are maintained by the zinc ion. Contact with DNA is made by the arginine preceding the α helix and by the second, third, and sixth residues of the α helix. Zinc finger motifs may be repeated in a tandem array within a protein such that the α helix of each zinc finger in the protein makes contact with the major groove of the DNA double helix. This repeated contact between the protein and the DNA produces a strong and specific DNA-protein interaction. The strength and specificity of the interaction can be regulated by the number of zinc finger motifs within the protein. Although zinc fingers were originally identified in DNA-binding proteins as regions that interact directly with DNA, they have since been found in proteins that do not bind to DNA. (See, e.g., Lodish, H. et al. (1995) Molecular Cell Biology, Scientific American Books, New York NY, pp. 447-451.)

The ligand-binding domain of nuclear hormone receptors is responsible for binding to ligands, coactivator proteins, and corepressor proteins. This domain is composed of three layers of α helices, with the central layer consisting of two helices containing many hydrophobic side chains (Moras, D. and H. Gronemeyer (1998) Curr. Opin. Cell Biol. 10:384-391). These two central α helices thus create a hydrophobic pocket which is the site of ligand binding. A ligand bound in this hydrophobic ligand-binding site is completely buried inside the receptor protein and is not exposed to solvent. This suggests that large conformational changes in the ligand-binding domain would accompany binding of a ligand. One of the α helices of the ligand-binding domain provides many of the inter-subunit contacts in dimers of nuclear receptors. This α helix contacts the ligand when it is bound in the ligand-binding pocket, suggesting that ligand binding can affect formation of receptor dimers (Weatherman, R.V. et al. (1999) Annu. Rev. Biochem. 68:559-581).

25 Classes of nuclear hormone receptors

Nuclear hormone receptors can be grouped into three broad classes: the steroid receptors, the RXR-heterodimeric receptors, and the orphan nuclear hormone receptors. The steroid receptors bind to steroid hormones, and this class includes the androgen receptor, mineralocorticoid receptor, estrogen receptor, glucocorticoid receptor, and progesterone receptor. The RXR-heterodimeric receptors bind to nonsteroid ligands, and this class includes the thyroid hormone receptor, retinoic acid receptor, vitamin D receptor, ecdysone receptor, and peroxisome proliferator activated receptor. The orphan nuclear hormone receptors include steroidogenic factor 1, nerve growth factor-induced receptor, and X-linked orphan receptor DAX-1.

The steroid hormone receptors are activated upon binding to specific steroid hormones. The conformational change induced by ligand binding leads to dissociation of the receptor from heat

WO 02/057454

PCT/US02/01339

shock proteins and formation of receptor homodimers which recognize specific palindromic DNA sequences called hormone response elements (HREs). Upon binding to an HRE, a steroid hormone receptor homodimer can regulate the transcription of target genes.

For example, the progesterone receptor (PR) is a steroid hormone receptor which is activated
5 by progesterone, a 4-pregnene-3,20-dione derived from cholesterol which is a critical oscillating component of the female reproductive cycle. These oscillations correlate with anatomical and morphological changes including menstruation and pregnancy. The activities of progesterone are mediated through PR. In the cytoplasm, PR associates with several other proteins and factors known as the PR heterocomplex. This heterocomplex includes heat shock proteins and immunophilins such
10 as hsp70, hsp90, hsp27, p59 (hsp56), p48, and p23 (Johnson, J.L. et al. (1994) Mol. Cell. Biol. 14:1956-1963). Upon binding progesterone, activated PR translocates to the nucleus, binds to canonical DNA transcriptional elements, and regulates progesterone-regulated genes implicated in differentiation and the cell cycle (Moutsatsou, P and C.E. Sekeris (1997) Ann. N.Y. Acad. Sci. 816:99-115). The PR antagonist RU 486, which can be used to terminate a pregnancy, is an example
15 of a commercial therapeutic targeted toward a steroid hormone receptor.

The RXR-heterodimeric nuclear receptors are distinguished from the steroid hormone receptors in that members of the former group bind to their target DNA sequences upon formation of heterodimers with retinoid X receptors (RXRs) (Mangelsdorf, D.J. and R.M. Evans (1995) Cell 83:841-850). Three different isoforms of RXR have been identified (Minucci, S. and K. Ozato
20 (1996) Curr. Opin. Genet. Dev. 6:567-574). The retinoic acid receptors (RARs) are examples of RXR-heterodimeric nuclear receptors. Retinoic acid (RA) is a biologically active metabolite of vitamin A (retinol), a fat-soluble vitamin found mainly in fish liver oils, liver, egg yolk, butter, and cream. While 9-cis-RA binds to RARs and RXRs, all-trans-RA binds only to RXRs. RAR/RXR heterodimers bind with high affinity to specific DNA sequences known as retinoic acid response
25 elements (RAREs), thus acting as regulators of RA-dependent transcription.

Peroxisome proliferator activated receptors (PPARs) are therapeutically important RXR-heterodimeric nuclear receptors which are induced by fatty acids and eicosanoids. There are three known isotypes of PPAR, each with specific expression patterns, and these PPARs are involved in the regulation of genes involved in systemic homeostasis of glucose and lipids (Kliever, S.A. and T.M.
30 Willson (1998) Curr. Opin. Genet. Dev. 8:576-581; Michalik, L. and W. Wahli (1999) Curr. Opin. Biotechnol. 10:564-570). As such, PPARs are therapeutic targets for disorders such as diabetes, dyslipidemia, and obesity (Smith, S.A. (1996) Pharmacol. Rev. Commun. 8:57-64; Willson, T.M. and W. Wahli (1997) Curr. Opin. Chem. Biol. 1:235-241; Barroso, I. et al. (1999) Nature 402:880-883).

The orphan nuclear receptors either have no known activating ligand, or can exert their
35 transcriptional regulatory activities without benefit of ligand binding. For example, in *Caenorhabditis*

WO 02/057454

PCT/US02/01339

elegans, the X-chromosome encoded nuclear hormone receptor homologue SEX-1 regulates transcription of the sex determination gene *xol-1* (Carmi, I. et al. (1998) Nature 396:168-173). Rather than relying on ligand binding, SEX-1 acts as a transcriptional regulator in a dose-dependent manner, in effect controlling sexual differentiation through an X-chromosome-counting mechanism.

5 Some nuclear hormone receptors lack the conventional DNA-binding domain typically associated with the nuclear hormone receptor family. DAX-1 is one such nuclear hormone receptor lacking the conventional DNA-binding domain, and mutations in DAX-1 have been shown to cause X-linked adrenal hypoplasia congenita (Zanaria, E.F. et al. (1994) Nature 372:635-641). DAX-1 is an orphan nuclear receptor which interacts directly with steroidogenic factor 1 (SF-1) (Ito, M. et al. (1997) Mol. Cell. Biol. 17:1476-1483), and DAX-1 is capable of modulating the action of SF-1 in sex-specific gene expression (Nachtigal, M.W. et al. (1998) Cell 445-454). SF-1 is an orphan nuclear receptor which acts as a transcription factor for several steroidogenic enzyme genes in the adrenal gland and gonads (Lala, D.S. et al. (1992) Mol. Endocrinol. 6:1249-1258; Lynch, J.P. et al. (1993) Mol. Endocrinol. 7:776-786; Clemens, J.W. et al. (1994) Endocrinology 134:1499-1508), and can also regulate several genes expressed in pituitary gonadotrope cells (Barnhart, K.M. and P.L. Mellon (1994) Mol. Endocrinol. 8:878-885; Ingraham, H.A. et al. (1994) Genes Dev. 8:2302-2312; Halvorson, L.M. et al. (1996) J. Biol. Chem. 271:6645-6650; Keri, R.A. and J.H. Nilson (1996) J. Biol. Chem. 271:10782-10785).

SF-1 also acts as a potent transactivator of small heterodimer partner (SHP; short heterodimer partner) (Lee, Y.K. et al. (1999) J. Biol. Chem. 274:20869-20873). SHP is another example of a nuclear hormone receptor lacking the conventional DNA-binding domain (Seol, W. et al. (1996) Science 272:1336-1339; Lee, H.-K. et al. (1998) J. Biol. Chem. 273:14398-14402). SHP interacts with many members of the nuclear hormone receptor family, including retinoid receptors, estrogen receptor, thyroid hormone receptor, and the orphan receptor CAR. SHP acts as an inhibitor of estrogen receptor-mediated transcriptional activation by competing with coactivators for binding to estrogen receptor (Johansson, L. et al. (1999) J. Biol. Chem. 274:345-353). SHP also inhibits transactivation by the orphan receptor hepatocyte nuclear factor 4, and by retinoid X receptor (Lee, Y.K. et al. (2000) Mol. Cell. Biol. 20:187-195).

Consequences of defective transcription regulation

30 Many neoplastic disorders in humans can be attributed to inappropriate gene expression. Malignant cell growth may result from either excessive expression of tumor promoting genes or insufficient expression of tumor suppressor genes (Cleary, M.L. (1992) Cancer Surv. 15:89-104). Chromosomal translocations may also produce chimeric loci which fuse the coding sequence of one gene with the regulatory regions of a second unrelated gene. Such an arrangement likely results in inappropriate gene transcription, potentially contributing to malignancy.

WO 02/057454

PCT/US02/01339

In addition, the immune system responds to infection or trauma by activating a cascade of events that coordinate the progressive selection, amplification, and mobilization of cellular defense mechanisms. A complex and balanced program of gene activation and repression is involved in this process. However, hyperactivity of the immune system as a result of improper or insufficient regulation of gene expression may result in considerable tissue or organ damage. This damage is well documented in immunological responses associated with arthritis, allergens, heart attack, stroke, and infections. (See, e.g., Isselbacher et al. (1996) *Harrison's Principles of Internal Medicine*, 13/e, McGraw Hill, Inc. and Teton Data Systems Software.)

Furthermore, the growth of multicellular organisms is based upon the induction and coordination of cell differentiation at the appropriate stages of development. Central to this process is differential gene expression, which confers the distinct identities of cells and tissues throughout the body. Failure to regulate gene expression during development could result in developmental disorders.

Ligand-Gated Receptor Ion Channels

Ligand-gated receptor ion channels fall into two categories. The first category, extracellular ligand-gated receptor ion channels (ELGs), rapidly transduce neurotransmitter-binding events into electrical signals, such as fast synaptic neurotransmission. ELG function is regulated by post-translational modification. The second category, intracellular ligand-gated receptor ion channels (ILGs), are activated by many intracellular second messengers and do not require post-translational modification(s) to effect a channel-opening response.

ELGs depolarize excitable cells to the threshold of action potential generation. In non-excitabile cells, ELGs permit a limited calcium ion-influx during the presence of agonist. ELGs include channels directly gated by neurotransmitters such as acetylcholine, L-glutamate, glycine, ATP, serotonin, GABA, and histamine. ELG genes encode proteins having strong structural and functional similarities. ILGs are encoded by distinct and unrelated gene families and include receptors for cAMP, cGMP, calcium ions, ATP, and metabolites of arachidonic acid.

Macrophage Scavenger Receptors

Macrophage scavenger receptors with broad ligand specificity may participate in the binding of low density lipoproteins (LDL) and foreign antigens. Scavenger receptors types I and II are trimeric membrane proteins with each subunit containing a small N-terminal intracellular domain, a transmembrane domain, a large extracellular domain, and a C-terminal cysteine-rich domain. The extracellular domain contains a short spacer domain, an α -helical coiled-coil domain, and a triple helical collagenous domain. These receptors have been shown to bind a spectrum of ligands, including chemically modified lipoproteins and albumin, polyribonucleotides, polysaccharides, phospholipids, and asbestos (Matsumoto, A. et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:9133-9137;

WO 02/057454

PCT/US02/01339

Elomaa, O. et al. (1995) *Cell* 80:603-609). The scavenger receptors are thought to play a key role in atherogenesis by mediating uptake of modified LDL in arterial walls, and in host defense by binding bacterial endotoxins, bacteria, and protozoa.

T-Cell Receptors

5 T cells play a dual role in the immune system as effectors and regulators, coupling antigen recognition with the transmission of signals that induce cell death in infected cells and stimulate proliferation of other immune cells. Although a population of T cells can recognize a wide range of different antigens, an individual T cell can only recognize a single antigen and only when it is presented to the T cell receptor (TCR) as a peptide complexed with a major histocompatibility molecule (MHC) on the surface of an antigen presenting cell. The TCR on most T cells consists of immunoglobulin-like integral membrane glycoproteins containing two polypeptide subunits, α and β , of similar molecular weight. Both TCR subunits have an extracellular domain containing both variable and constant regions, a transmembrane domain that traverses the membrane once, and a short intracellular domain (Saito, H. et al. (1984) *Nature* 309:757-762). The genes for the TCR subunits are constructed through somatic rearrangement of different gene segments. Interaction of antigen in the proper MHC context with the TCR initiates signaling cascades that induce the proliferation, maturation, and function of cellular components of the immune system (Weiss, A. (1991) *Annu. Rev. Genet.* 25: 487-510). Rearrangements in TCR genes and alterations in TCR expression have been noted in lymphomas, leukemias, autoimmune disorders, and immunodeficiency disorders (Aisenberg, A.C. et al. (1985) *N. Engl. J. Med.* 313:529-533; Weiss, *supra*).

Selectins

Selectins, or LEC-CAMs, comprise a specialized lectin subfamily involved primarily in inflammation and leukocyte adhesion (reviewed in Lasky, L.A. (1991) *J. Cell. Biochem.* 45:139-146). Selectins mediate the recruitment of leukocytes from the circulation to sites of acute inflammation and are expressed on the surface of vascular endothelial cells in response to cytokine signaling. Selectins bind to specific ligands on the leukocyte cell membrane and enable the leukocyte to adhere to and migrate along the endothelial surface. Binding of selectin to its ligand leads to polarized rearrangement of the actin cytoskeleton and stimulates signal transduction within the leukocyte (Brenner, B. et al. (1997) *Biochem. Biophys. Res. Commun.* 231:802-807; Hidari, K.I. et al. (1997) *J. Biol. Chem.* 272:28750-28756). Members of the selectin family possess three characteristic motifs: a lectin or carbohydrate recognition domain; an epidermal growth factor-like domain; and a variable number of short consensus repeats (scr or "sushi" repeats). Sushi domains, also known as complement control protein (CCP) modules, or short consensus repeats (SCR), occur in a wide variety of complement and adhesion proteins (Norman, D.G. et al. (1991) *J. Mol. Biol.* 219:717-725).

Netrin Receptors

WO 02/057454

PCT/US02/01339

The netrins are a family of molecules that function as diffusible attractants and repellants to guide migrating cells and axons to their targets within the developing nervous system. The netrin receptors include the *C. elegans* protein UNC-5, as well as homologues recently identified in vertebrates (Leonardo, E.D. et al. (1997) Nature 386:833-838). These receptors are members of the immunoglobulin superfamily, and also contain a characteristic domain called the ZU5 domain. Mutations in the mouse member of the netrin receptor family, Rcm (rostral cerebellar malformation) result in cerebellar and midbrain defects as an apparent result of abnormal neuronal migration (Ackerman, S.L. et al. (1997) Nature 386:838-842).

VPS10 Domain Containing Receptors

The members of the VPS10 domain containing receptor family all contain a domain with homology to the yeast vacuolar sorting protein 10 (VPS10) receptor. This family includes the mosaic receptor SorLA, the neurotensin receptor sortilin, and SorCS, which is expressed during mouse embryonal and early postnatal nervous system development (Hermey, G. et al. (1999) Biochem. Biophys. Res. Commun. 266:347-351; Hermey, G. et al. (2001) Neuroreport 12:29-32). A recently identified member of this family, SorCS2, is highly expressed in the developing and mature mouse central nervous system. Its main site of expression is the floor plate, and high levels are also detected transiently in brain regions including the dopaminergic brain nuclei and the dorsal thalamus (Rezgaoui, M. (2001) Mech. Dev. 100:335-338).

Membrane-Associated Proteins

Tetraspan Family Proteins

The transmembrane 4 superfamily (TM4SF) or tetraspan family is a multigene family encoding type III integral membrane proteins (Wright, M.D. and Tomlinson, M.G. (1994) Immunol. Today 15:588). The TM4SF is comprised of membrane proteins which traverse the cell membrane four times. Members of the TM4SF include platelet and endothelial cell membrane proteins, melanoma-associated antigens, leukocyte surface glycoproteins, colonic carcinoma antigens, tumor-associated antigens, and surface proteins of the schistosome parasites (Jankowski, S.A. (1994) Oncogene 9:1205-1211). Members of the TM4SF share about 25-30% amino acid sequence identity with one another. A number of TM4SF members have been implicated in signal transduction, control of cell adhesion, regulation of cell growth and proliferation, including development and oncogenesis, and cell motility, including tumor cell metastasis. Expression of TM4SF proteins is associated with a variety of tumors and the level of expression may be altered when cells are growing or activated.

Tumor Antigens

Tumor antigens are surface molecules that are differentially expressed in tumor cells relative to normal cells. Tumor antigens distinguish tumor cells immunologically from normal cells and provide diagnostic and therapeutic targets for human cancers (Takagi, S. et al. (1995) Int. J. Cancer

WO 02/057454

PCT/US02/01339

61:706-715; Liu, E. et al. (1992) *Oncogene* 7:1027-1032).

Ion Channels

Ion channels are found in the plasma membranes of virtually every cell in the body. For example, chloride channels mediate a variety of cellular functions including regulation of membrane potentials and absorption and secretion of ions across epithelial membranes. When present in intracellular membranes of the Golgi apparatus and endocytic vesicles, chloride channels also regulate organelle pH. (See, e.g., Greger, R. (1988) *Annu. Rev. Physiol.* 50:111-122.) Electrophysiological and pharmacological properties of chloride channels, including ion conductance, current-voltage relationships, and sensitivity to modulators, suggest that different chloride channels exist in muscles, neurons, fibroblasts, epithelial cells, and lymphocytes. Many channels have sites for phosphorylation by one or more protein kinases including protein kinase A, protein kinase C, tyrosine kinase, and casein kinase II, all of which regulate ion channel activity in cells. Inappropriate phosphorylation of proteins in cells has been linked to changes in cell cycle progression and cell differentiation. Changes in the cell cycle have been linked to induction of apoptosis or cancer. Changes in cell differentiation have been linked to diseases and disorders of the reproductive system, immune system, and skeletal muscle.

Cerebellar granule neurons possess a non-inactivating potassium current which modulates firing frequency upon receptor stimulation by neurotransmitters and controls the resting membrane potential. Potassium channels that exhibit non-inactivating currents include the *ether a go-go* (EAG) channel. A membrane protein designated KCR1 specifically binds to rat EAG by means of its C-terminal region and regulates the cerebellar non-inactivating potassium current. KCR1 is predicted to contain 12 transmembrane domains, with intracellular amino and carboxyl termini. Structural characteristics of these transmembrane regions appear to be similar to those of the transporter superfamily, but no homology between KCR1 and known transporters was found, suggesting that KCR1 belongs to a novel class of transporters. KCR1 appears to be the regulatory component of non-inactivating potassium channels (Hoshi, N. et al. (1998) *J. Biol. Chem.* 273:23080-23085).

ABC Transporters

ATP-binding cassette (ABC) transporters, also called the "traffic ATPases", are a superfamily of membrane proteins that mediate transport and channel functions in prokaryotes and eukaryotes (Higgins, C.F. (1992) *Annu. Rev. Cell Biol.* 8:67-113). ABC proteins share a similar overall structure and significant sequence homology. All ABC proteins contain a conserved domain of approximately two hundred amino acid residues which includes one or more nucleotide binding domains. Mutations in ABC transporter genes are associated with various disorders, such as hyperbilirubinemia II/Dubin-Johnson syndrome, recessive Stargardt's disease, X-linked adrenoleukodystrophy, multidrug resistance, celiac disease, and cystic fibrosis.

WO 02/057454

PCT/US02/01339

Semaphorins and Neuropilins

Semaphorins are a large group of axonal guidance molecules consisting of at least 30 different members and are found in vertebrates, invertebrates, and even certain viruses. All semaphorins contain the sema domain which is approximately 500 amino acids in length. Neuropilin, a semaphorin receptor, has been shown to promote neurite outgrowth *in vitro*. The extracellular region of neuropilins consists of three different domains: CUB, discoidin, and MAM domains. The CUB and the MAM motifs of neuropilin have been suggested to have roles in protein-protein interactions and are thought to be involved in the binding of semaphorins through the sema and the C-terminal domains (reviewed in Raper, J.A. (2000) *Curr. Opin. Neurobiol.* 10:88-94).

Membrane Proteins Associated with Intercellular Communication

Intercellular communication is essential for the development and survival of multicellular organisms. Cells communicate with one another through the secretion and uptake of protein signaling molecules. The uptake of proteins into the cell is achieved by endocytosis, in which the interaction of signaling molecules with the plasma membrane surface, often via binding to specific receptors, results in the formation of plasma membrane-derived vesicles that enclose and transport the molecules into the cytosol. The secretion of proteins from the cell is achieved by exocytosis, in which molecules inside of the cell are packaged into membrane-bound transport vesicles derived from the *trans* Golgi network. These vesicles fuse with the plasma membrane and release their contents into the surrounding extracellular space. Endocytosis and exocytosis result in the removal and addition of plasma membrane components, and the recycling of these components is essential to maintain the integrity, identity, and functionality of both the plasma membrane and internal membrane-bound compartments.

Nogo has been identified as a component of the central nervous system myelin that prevents axonal regeneration in adult vertebrates. Cleavage of the Nogo-66 receptor and other glycosphatidylinositol-linked proteins from axonal surfaces renders neurons insensitive to Nogo-66, facilitating potential recovery from CNS damage (Fournier, A.E. et al. (2001) *Nature* 409:341-346).

The slit proteins are extracellular matrix proteins expressed by cells at the ventral midline of the nervous system. Slit proteins are ligands for the repulsive guidance receptor Roundabout (Robo) and thus play a role in repulsive axon guidance (Brose, K. et al. (1999) *Cell* 96:795-806).

Lysosomes are the site of degradation of intracellular material during autophagy and of extracellular molecules following endocytosis. Lysosomal enzymes are packaged into vesicles which bud from the *trans*-Golgi network. These vesicles fuse with endosomes to form the mature lysosome in which hydrolytic digestion of endocytosed material occurs. Lysosomes can fuse with autophagosomes to form a unique compartment in which the degradation of organelles and other

WO 02/057454

PCT/US02/01339

intracellular components occurs.

Protein sorting by transport vesicles, such as the endosome, has important consequences for a variety of physiological processes including cell surface growth, the biogenesis of distinct intracellular organelles, endocytosis, and the controlled secretion of hormones and neurotransmitters (Rothman, J.E. and Wieland, F.T. (1996) *Science* 272:227-234). In particular, neurodegenerative disorders and other neuronal pathologies are associated with biochemical flaws during endosomal protein sorting or endosomal biogenesis (Mayer R.J. et al. (1996) *Adv. Exp. Med. Biol.* 389:261-269).

Peroxisomes are organelles independent from the secretory pathway. They are the site of many peroxide-generating oxidative reactions in the cell. Peroxisomes are unique among eukaryotic organelles in that their size, number, and enzyme content vary depending upon organism, cell type, and metabolic needs (Waterham, H.R. and Cregg, J.M. (1997) *BioEssays* 19:57-66). Genetic defects in peroxisome proteins which result in peroxisomal deficiencies have been linked to a number of human pathologies, including Zellweger syndrome, rhizomelic chondrodysplasia punctata, X-linked adrenoleukodystrophy, acyl-CoA oxidase deficiency, bifunctional enzyme deficiency, classical Refsum's disease, DHAP alkyl transferase deficiency, and acatalasemia (Moser, H.W. and Moser, A.B. (1996) *Ann. NY Acad. Sci.* 804:427-441). In addition, Gartner, J. et al. (1991; *Pediatr. Res.* 29:141-146) found a 22 kDa integral membrane protein associated with lower density peroxisome-like subcellular fractions in patients with Zellweger syndrome.

Normal embryonic development and control of germ cell maturation is modulated by a number of secretory proteins which interact with their respective membrane-bound receptors. Cell fate during embryonic development is determined by members of the activin/TGF- β superfamily, cadherins, IGF-2, and other morphogens. In addition, proliferation, maturation, and redifferentiation of germ cell and reproductive tissues are regulated, for example, by IGF-2, inhibins, activins, and follistatins (Petraglia, F. (1997) *Placenta* 18:3-8; Mather, J.P. et al. (1997) *Proc. Soc. Exp. Biol. Med.* 215:209-222). Transforming growth factor beta (TGF β) signal transduction is mediated by two receptor Ser/Thr kinases acting in series, type II TGF β receptor and (TbetaR-II) phosphorylating type I TGF β receptor (TbetaR-I). TbetaR-I-associated protein-1 (TRECAP-1), which distinguishes between quiescent and activated forms of the type I transforming growth factor beta receptor, has been associated with TGF β signaling (Chang, M.J. et al. (1998) *J. Biol. Chem.* 273:9365-9368).

Retinoic acid receptor alpha (RAR alpha) mediates retinoic acid induced maturation and has been implicated in myeloid development. Genes induced by retinoic acid during granulocytic differentiation include E3, a hematopoietic-specific gene that is an immediate target for the activated RAR alpha during myelopoiesis (Scott, L.M. et al. (1996) *Blood* 88:2517-2530).

The μ -opioid receptor (MOR) mediates the actions of analgesic agents including morphine,

WO 02/057454

PCT/US02/01339

codeine, methadone, and fentanyl as well as heroin. MOR is functionally coupled to a G-protein-activated potassium channel (Mestek A. et al. (1995) J. Neurosci. 15:2396-2406). A variety of MOR subtypes exist. Alternative splicing has been observed with MOR-1 as with a number of G protein-coupled receptors including somatostatin 2, dopamine D2, prostaglandin EP3, and serotonin receptor subtypes 5-hydroxytryptamine4 and 5-hydroxytryptamine7 (Pan, Y.X. et al. (1999) Mol. Pharm. 56:396-403).

Peripheral and Anchored Membrane Proteins

Some membrane proteins are not membrane-spanning but are attached to the plasma membrane via membrane anchors or interactions with integral membrane proteins. Membrane anchors are covalently joined to a protein post-translationally and include such moieties as prenyl, myristyl, and glycosylphosphatidyl inositol groups. Membrane localization of peripheral and anchored proteins is important for their function in processes such as receptor-mediated signal transduction. For example, prenylation of Ras is required for its localization to the plasma membrane and for its normal and oncogenic functions in signal transduction.

15 **Extracellular Messengers**

Intercellular communication is essential for the growth and survival of multicellular organisms, and in particular, for the function of the endocrine, nervous, and immune systems. In addition, intercellular communication is critical for developmental processes such as tissue construction and organogenesis, in which cell proliferation, cell differentiation, and morphogenesis must be spatially and temporally regulated in a precise and coordinated manner. Cells communicate with one another through the secretion and uptake of diverse types of signaling molecules such as hormones, growth factors, neuropeptides, and cytokines.

Hormones

Hormones are signaling molecules that coordinately regulate basic physiological processes from embryogenesis throughout adulthood. These processes include metabolism, respiration, reproduction, excretion, fetal tissue differentiation and organogenesis, growth and development, homeostasis, and the stress response. Hormonal secretions and the nervous system are tightly integrated and interdependent. Hormones are secreted by endocrine glands, primarily the hypothalamus and pituitary, the thyroid and parathyroid, the pancreas, the adrenal glands, and the ovaries and testes.

The secretion of hormones into the circulation is tightly controlled. Hormones are often secreted in diurnal, pulsatile, and cyclic patterns. Hormone secretion is regulated by perturbations in blood biochemistry, by other upstream-acting hormones, by neural impulses, and by negative feedback loops. Blood hormone concentrations are constantly monitored and adjusted to maintain optimal, steady-state levels. Once secreted, hormones act only on those target cells that express

WO 02/057454

PCT/US02/01339

specific receptors.

Most disorders of the endocrine system are caused by either hyposecretion or hypersecretion of hormones. Hyposecretion often occurs when a hormone's gland of origin is damaged or otherwise impaired. Hypersecretion often results from the proliferation of tumors derived from hormone-secreting cells. Inappropriate hormone levels may also be caused by defects in regulatory feedback loops or in the processing of hormone precursors. Endocrine malfunction may also occur when the target cell fails to respond to the hormone.

Hormones can be classified biochemically as polypeptides, steroids, eicosanoids, or amines. Polypeptide hormones, which include diverse hormones such as insulin and growth hormone, vary in size and function and are often synthesized as inactive precursors that are processed intracellularly into mature, active forms. Amine hormones, which include epinephrine and dopamine, are amino acid derivatives that function in neuroendocrine signaling. Steroid hormones, which include the cholesterol-derived hormones estrogen and testosterone, function in sexual development and reproduction. Eicosanoid hormones, which include prostaglandins and prostacyclins, are fatty acid derivatives that function in a variety of processes. Most polypeptide hormones and some amine hormones are soluble in the circulation where they are highly susceptible to proteolytic degradation within seconds after their secretion. Steroid hormones and eicosanoid hormones are insoluble and must be transported in the circulation by carrier proteins. The following discussion will focus primarily on polypeptide hormones.

Hormones secreted by the hypothalamus and pituitary gland play a critical role in endocrine function by coordinately regulating hormonal secretions from other endocrine glands in response to neural signals. Hypothalamic hormones include thyrotropin-releasing hormone, gonadotropin-releasing hormone, somatostatin, growth-hormone releasing factor, corticotropin-releasing hormone, substance P, dopamine, and prolactin-releasing hormone. These hormones directly regulate the secretion of hormones from the anterior lobe of the pituitary. Hormones secreted by the anterior pituitary include adrenocorticotropic hormone (ACTH), melanocyte-stimulating hormone, somatotrophic hormones such as growth hormone and prolactin, glycoprotein hormones such as thyroid-stimulating hormone, luteinizing hormone (LH), and follicle-stimulating hormone (FSH), β -lipotropin, and β -endorphins. These hormones regulate hormonal secretions from the thyroid, pancreas, and adrenal glands, and act directly on the reproductive organs to stimulate ovulation and spermatogenesis. The posterior pituitary synthesizes and secretes antidiuretic hormone (ADH, vasopressin) and oxytocin.

Disorders of the hypothalamus and pituitary often result from lesions such as primary brain tumors, adenomas, infarction associated with pregnancy, hypophysectomy, aneurysms, vascular malformations, thrombosis, infections, immunological disorders, and complications due to head

WO 02/057454

PCT/US02/01339

trauma. Such disorders have profound effects on the function of other endocrine glands. Disorders associated with hypopituitarism include hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism. Disorders associated with hyperpituitarism include acromegaly, gigantism, and syndrome of inappropriate ADH secretion (SIADH), often caused by benign adenomas.

Hormones secreted by the thyroid and parathyroid primarily control metabolic rates and the regulation of serum calcium levels, respectively. Thyroid hormones include calcitonin, somatostatin, and thyroid hormone. The parathyroid secretes parathyroid hormone. Disorders associated with hypothyroidism include goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism. Disorders associated with hyperthyroidism include thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease. Disorders associated with hyperparathyroidism include Conn disease (chronic hypercalcemia) leading to bone resorption and parathyroid hyperplasia.

Hormones secreted by the pancreas regulate blood glucose levels by modulating the rates of carbohydrate, fat, and protein metabolism. Pancreatic hormones include insulin, glucagon, amylin, γ -aminobutyric acid, gastrin, somatostatin, and pancreatic polypeptide. The principal disorder associated with pancreatic dysfunction is diabetes mellitus caused by insufficient insulin activity. Diabetes mellitus is generally classified as either Type I (insulin-dependent, juvenile diabetes) or Type II (non-insulin-dependent, adult diabetes). The treatment of both forms by insulin replacement therapy is well known. Diabetes mellitus often leads to acute complications such as hypoglycemia (insulin shock), coma, diabetic ketoacidosis, lactic acidosis, and chronic complications leading to disorders of the eye, kidney, skin, bone, joint, cardiovascular system, nervous system, and to decreased resistance to infection.

The anatomy, physiology, and diseases related to hormonal function are reviewed in McCance, K.L. and Huether, S.E. (1994) Pathophysiology: The Biological Basis for Disease in Adults and Children, Mosby-Year Book, Inc., St. Louis, MO; Greenspan, F.S. and Baxter, J.D. (1994) Basic and Clinical Endocrinology, Appleton and Lange, East Norwalk, CT.

Growth Factors

Growth factors are secreted proteins that mediate intercellular communication. Unlike hormones, which travel great distances via the circulatory system, most growth factors are primarily local mediators that act on neighboring cells. Most growth factors contain a hydrophobic N-terminal signal peptide sequence which directs the growth factor into the secretory pathway. Most growth factors also undergo post-translational modifications within the secretory pathway. These modifications can include proteolysis, glycosylation, phosphorylation, and intramolecular disulfide

WO 02/057454

PCT/US02/01339

bond formation. Once secreted, growth factors bind to specific receptors on the surfaces of neighboring target cells, and the bound receptors trigger intracellular signal transduction pathways. These signal transduction pathways elicit specific cellular responses in the target cells. These responses can include the modulation of gene expression and the stimulation or inhibition of cell division, cell differentiation, and cell motility.

Growth factors fall into at least two broad and overlapping classes. The broadest class includes the large polypeptide growth factors, which are wide-ranging in their effects. These factors include epidermal growth factor (EGF), fibroblast growth factor (FGF), transforming growth factor- β (TGF- β), insulin-like growth factor (IGF), nerve growth factor (NGF), and platelet-derived growth factor (PDGF), each defining a family of numerous related factors. The large polypeptide growth factors, with the exception of NGF, act as mitogens on diverse cell types to stimulate wound healing, bone synthesis and remodeling, extracellular matrix synthesis, and proliferation of epithelial, epidermal, and connective tissues. Members of the TGF- β , EGF, and FGF families also function as inductive signals in the differentiation of embryonic tissue. NGF functions specifically as a neurotrophic factor, promoting neuronal growth and differentiation.

Another class of growth factors includes the hematopoietic growth factors, which are narrow in their target specificity. These 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. These factors include the colony-stimulating factors (G-CSF, M-CSF, GM-CSF, and CSF1-3), erythropoietin, and the cytokines. The cytokines are specialized hematopoietic factors secreted by cells of the immune system and are discussed in detail below.

Growth factors play critical roles in neoplastic transformation of cells *in vitro* and in tumor progression *in vivo*. Overexpression of the large polypeptide growth factors promotes the proliferation and transformation of cells in culture. Inappropriate expression of these growth factors by tumor cells *in vivo* may contribute to tumor vascularization and metastasis. Inappropriate activity of hematopoietic growth factors can result in anemias, leukemias, and lymphomas. Moreover, growth factors are both structurally and functionally related to oncoproteins, the potentially cancer-causing products of proto-oncogenes. Certain FGF and PDGF family members are themselves homologous to oncoproteins, whereas receptors for some members of the EGF, NGF, and FGF families are encoded by proto-oncogenes. Growth factors also affect the transcriptional regulation of both proto-oncogenes and oncosuppressor genes. (Pimentel, E. (1994) Handbook of Growth Factors, CRC Press, Ann Arbor, MI; McKay, I. and Leigh, I., eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York, NY; Habenicht, A., ed. (1990) Growth Factors, Differentiation Factors, and Cytokines, Springer-Verlag, New York, NY.)

WO 02/057454

PCT/US02/01339

In addition, some of the large polypeptide growth factors play crucial roles in the induction of the primordial germ layers in the developing embryo. This induction ultimately results in the formation of the embryonic mesoderm, ectoderm, and endoderm which in turn provide the framework for the entire adult body plan. Disruption of this inductive process would be catastrophic to embryonic development.

Small Peptide Factors - Neuropeptides and Vasomediators

Neuropeptides and vasomediators (NP/VM) comprise a family of small peptide factors, typically of 20 amino acids or less. These factors generally function in neuronal excitation and inhibition of vasoconstriction/vasodilation, muscle contraction, and hormonal secretions from the brain and other endocrine tissues. 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, gastrin, and many of the peptide hormones discussed above. NP/VMs can transduce signals directly, modulate the activity or release of other neurotransmitters and hormones, and act as catalytic enzymes in signaling 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.)

Cytokines

Cytokines comprise a family of signaling molecules that modulate the immune system and the inflammatory response. Cytokines are usually secreted by leukocytes, or white blood cells, in response to injury or infection. Cytokines function as growth and differentiation factors that act primarily on cells of the immune system such as B- and T-lymphocytes, monocytes, macrophages, and granulocytes. Like other signaling molecules, cytokines bind to specific plasma membrane receptors and trigger intracellular signal transduction pathways which alter gene expression patterns. There is considerable potential for the use of cytokines in the treatment of inflammation and immune system disorders.

Cytokine structure and function have been extensively characterized in vitro. Most cytokines are small polypeptides of about 30 kilodaltons or less. Over 50 cytokines have been identified from human and rodent sources. Examples of cytokine subfamilies include the interferons (IFN- α , - β , and - γ), the interleukins (IL1-IL13), the tumor necrosis factors (TNF- α and - β), and the chemokines. Many cytokines have been produced using recombinant DNA techniques, and the activities of individual cytokines have been determined in vitro. These activities include regulation of leukocyte proliferation, differentiation, and motility.

WO 02/057454

PCT/US02/01339

The activity of an individual cytokine *in vitro* may not reflect the full scope of that cytokine's activity *in vivo*. Cytokines are not expressed individually *in vivo* but are instead expressed in combination with a multitude of other cytokines when the organism is challenged with a stimulus. Together, these cytokines collectively modulate the immune response in a manner appropriate for that particular stimulus. Therefore, the physiological activity of a cytokine is determined by the stimulus itself and by complex interactive networks among co-expressed cytokines which may demonstrate both synergistic and antagonistic relationships.

Chemokines comprise a cytokine subfamily with over 30 members. (Reviewed in Wells, T.N.C. and Peitsch, M.C. (1997) *J. Leukoc. Biol.* 61:545-550.) Chemokines were initially identified as chemotactic proteins that recruit monocytes and macrophages to sites of inflammation. Recent evidence indicates that chemokines may also play key roles in hematopoiesis and HIV-1 infection. Chemokines are small proteins which range from about 6-15 kilodaltons in molecular weight. Chemokines are further classified as C, CC, CXC, or CX₂C based on the number and position of critical cysteine residues. The CC chemokines, for example, each contain a conserved motif consisting of two consecutive cysteines followed by two additional cysteines which occur downstream at 24- and 16-residue intervals, respectively (ExPASy PROSITE database, documents PS00472 and PDOC00434). The presence and spacing of these four cysteine residues are highly conserved, whereas the intervening residues diverge significantly. However, a conserved tyrosine located about 15 residues downstream of the cysteine doublet seems to be important for chemotactic activity. Most of the human genes encoding CC chemokines are clustered on chromosome 17, although there are a few examples of CC chemokine genes that map elsewhere. Other chemokines include lymphotactin (C chemokine); macrophage chemotactic and activating factor (MCAF/MCP-1; CC chemokine); platelet factor 4 and IL-8 (CXC chemokines); and fractalkine and neurotractin (CX₂C chemokines). (Reviewed in Luster, A.D. (1998) *N. Engl. J. Med.* 338:436-445.)

Chromogranins and secretogranins are acidic proteins present in the secretory granules of endocrine and neuro-endocrine cells (Huttner, W.B. et al. (1991) *Trends Biochem. Sci.* 16:27-30) (Simon, J.-P. et al. (1989) *Biochem. J.* 262:1-13). Granins may be precursors of biologically-active peptides, or they may be helper proteins in the packaging of peptide hormones and neuropeptides - their precise role is unclear.

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

WO 02/057454

PCT/US02/01339

SUMMARY OF THE INVENTION

The invention features purified polypeptides, receptors and membrane-associated proteins, referred to collectively as "REMAP" and individually as "REMAP-1," "REMAP-2," "REMAP-3," "REMAP-4," "REMAP-5," "REMAP-6," "REMAP-7," "REMAP-8," "REMAP-9," "REMAP-10," "REMAP-11," "REMAP-12," "REMAP-13," "REMAP-14," and "REMAP-15." 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:1-15, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-15.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-15. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:16-30.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting

WO 02/057454

PCT/US02/01339

of SEQ ID NO:1-15, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. The method comprises a) 5 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 10 polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an 15 immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.

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:16-30, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at 20 least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, 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 25 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:16-30, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, c) a polynucleotide complementary to the polynucleotide of a), d) a 30 polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments 35 thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if

WO 02/057454

PCT/US02/01339

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:16-30, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, 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:1-15, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional REMAP, 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:1-15, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. 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

WO 02/057454

PCT/US02/01339

treating a disease or condition associated with decreased expression of functional REMAP, 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:1-15, b) a polypeptide
5 comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the
10 group consisting of SEQ ID NO:1-15. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of
15 treating a disease or condition associated with overexpression of functional REMAP, 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:1-15, b) a polypeptide comprising a
20 naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group
25 consisting of SEQ ID NO:1-15. 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:1-15, b) a polypeptide comprising a
30 naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group
35 consisting of SEQ ID NO:1-15. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the

WO 02/057454

PCT/US02/01339

polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, b) detecting altered expression of the target polynucleotide, and c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

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

WO 02/057454

PCT/US02/01339

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

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

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

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

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

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold
15 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
20 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
25 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.
30 Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the
35 invention is not entitled to antedate such disclosure by virtue of prior invention.

WO 02/057454

PCT/US02/01339

DEFINITIONS

"REMAP" refers to the amino acid sequences of substantially purified REMAP 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.

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

10 An "allelic variant" is an alternative form of the gene encoding REMAP. 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
15 or more times in a given sequence.

"Altered" nucleic acid sequences encoding REMAP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as REMAP or a polypeptide with at least one functional characteristic of REMAP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide
20 probe of the polynucleotide encoding REMAP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding REMAP. 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 REMAP. Deliberate amino acid substitutions may be made on the basis of similarity in
25 polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of REMAP 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.
30 Amino acids with uncharged side chains having similar hydrophobicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

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

WO 02/057454

PCT/US02/01339

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.

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

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

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

WO 02/057454

PCT/US02/01339

cross-linker. (See, e.g., Brody, E.N. and L. Gold (2000) *J. Biotechnol.* 74:5-13.)

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

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

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; 10 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 15 molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

20 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 REMAP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

25 "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 30 or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding REMAP or fragments of REMAP 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 35 sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

WO 02/057454

PCT/US02/01339

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

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

	Original Residue	Conservative Substitution
	Ala	Gly, Ser
15	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
	Gln	Asn, Glu, His
20	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
	Leu	Ile, Val
25	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
	Thr	Ser, Val
30	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

WO 02/057454

PCT/US02/01339

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.

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

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

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

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

25 A fragment of SEQ ID NO:16-30 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:16-30, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:16-30 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:16-30 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:16-30 and the region of SEQ ID NO:16-30 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

30 A fragment of SEQ ID NO:1-15 is encoded by a fragment of SEQ ID NO:16-30. A fragment of SEQ ID NO:1-15 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-15. For example, a fragment of SEQ ID NO:1-15 is useful as an immunogenic peptide

WO 02/057454

PCT/US02/01339

for the development of antibodies that specifically recognize SEQ ID NO:1-15. The precise length of a fragment of SEQ ID NO:1-15 and the region of SEQ ID NO:1-15 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

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

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

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

15 Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of 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
20 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms
25 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
30 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
35 compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version

WO 02/057454

PCT/US02/01339

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

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

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

10 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
15 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
20 sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative
25 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
30 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
35 comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version

WO 02/057454

PCT/US02/01339

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

5 *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.

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

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

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

30 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

WO 02/057454

PCT/US02/01339

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

10 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 15 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 20 hybridization complex may be formed in solution (e.g., $C_{3,t}$ or $R_{3,t}$ analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

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

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

30 An "immunogenic fragment" is a polypeptide or oligopeptide fragment of REMAP 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 REMAP which is useful in any of the antibody production methods disclosed herein or known in the art.

35 The term "microarray" refers to an arrangement of a plurality of polynucleotides,

WO 02/057454

PCT/US02/01339

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.

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

10 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.

"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
15 necessary to join two protein coding regions, in the same reading frame.

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

"Post-translational modification" of an REMAP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of REMAP.

25 "Probe" refers to nucleic acid sequences encoding REMAP, 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.

30 "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
35 be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100,

WO 02/057454

PCT/US02/01339

or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

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

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

WO 02/057454

PCT/US02/01339

This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

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

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

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

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

The term "sample" is used in its broadest sense. A sample suspected of containing REMAP, nucleic acids encoding REMAP, 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,

WO 02/057454

PCT/US02/01339

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.

5 "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.

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

"Transformation" describes a process by which exogenous DNA is introduced into a recipient 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
15 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.

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

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of
35 the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-

WO 02/057454

PCT/US02/01339

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

THE INVENTION

25 The invention is based on the discovery of new human receptors and membrane-associated proteins (REMAP), the polynucleotides encoding REMAP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, neurological, metabolic, developmental, and endocrine disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide
30 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
35 an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

WO 02/057454

PCT/US02/01339

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 scores for the matches between each polypeptide and its homolog(s). Column 5 shows the annotation of the GenBank homolog(s) along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. 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 receptors and membrane-associated proteins. For example, SEQ ID NO:1 is 97% identical to rat retinoic acid receptor alpha 2 isoform (GenBank ID g3213188) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $8.5e-245$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:1 also contains a nuclear hormone receptor ligand-binding domain and a C4 type zinc finger as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:1 is a retinoic acid receptor. In an alternative example, SEQ ID NO:2 is 31% identical, from residue A3 to residue H575, to human multiple membrane spanning receptor TRC8 (GenBank ID g3395787) as determined by BLAST with a probability score of $9.7e-61$. (See Table 2.) Data from additional BLAST analyses provide further corroborative evidence that SEQ ID NO:2 is a multiple membrane spanning receptor. In an alternative example, SEQ ID NO:5 is 76% identical, from residue V4 to residue A479, to rat potential ligand-binding protein (GenBank ID g57734) as determined by BLAST with a probability score of $6.3e-187$. (See Table 2.) Data from additional BLAST analyses provide further corroborative evidence that SEQ ID NO:5 is an olfactory ligand binding protein. In an alternative example, SEQ ID

WO 02/057454

PCT/US02/01339

NO:11 is 77% identical, from residue M1 to residue F310, to Canis familiaris olfactory receptor CFOLF2 (GenBank ID g1314663) as determined by BLAST with a probability score of 9.7e-129. (See Table 2.) SEQ ID NO:11 also contains a 7-transmembrane receptor (rhodopsin family) active site domain as determined by searching for statistically significant matches in the HMM-based PFAM database. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:11 is a G-protein coupled receptor. In an alternative example, SEQ ID NO:15 is 99% identical, from residue M5 to residue M328, and is 89% identical, from residue V313 to residue E410, to a human protein which is an ortholog of the potential ligand-binding protein RYA3 (GenBank ID g11877275) as determined by BLAST with a probability score of 3.3e-207. (See Table 2.) Data from BLIMPS and additional BLAST analyses provide further corroborative evidence that SEQ ID NO:15 is a ligand-binding protein. SEQ ID NO:3-4, SEQ ID NO:6-10, and SEQ ID NO:12-14 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-15 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Column 1 lists the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:), the corresponding Incyte polynucleotide consensus sequence number (Incyte ID) for each polynucleotide of the invention, and the length of each polynucleotide sequence in basepairs. Column 2 shows the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences used to assemble the full length polynucleotide sequences of the invention, and of fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:16-30 or that distinguish between SEQ ID NO:16-30 and related polynucleotide sequences.

The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotide sequences. In addition, the polynucleotide fragments described in column 2 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the polynucleotide fragments described in column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, a polynucleotide sequence identified as

WO 02/057454

PCT/US02/01339

FL_XXXXXX_N₁_N₂_YYYYY_N₃_N₄ represents a "stitched" sequence in which XXXXXX is the identification number of the cluster of sequences to which the algorithm was applied, and YYYYY is the number of the prediction generated by the algorithm, and N_{1,2,3,...} if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, a polynucleotide sequence identified as FL_XXXXXX_gAAAAA_gBBBBB_1_N is a "stretched" sequence, with XXXXXX being the Incyte project identification number, gAAAAA being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, gBBBBB being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and N referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, gBBBBB).

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

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

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

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

WO 02/057454

PCT/US02/01339

The invention also encompasses REMAP variants. A preferred REMAP 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 REMAP amino acid sequence, and which contains at least one functional or structural characteristic of REMAP.

5 The invention also encompasses polynucleotides which encode REMAP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:16-30, which encodes REMAP. The polynucleotide sequences of SEQ ID NO:16-30, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the
10 sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding REMAP. 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 REMAP. A particular aspect of the invention encompasses a variant of a
15 polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:16-30 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:16-30. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of REMAP.

20 In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide sequence encoding REMAP. A splice variant may have portions which have significant sequence identity to the polynucleotide sequence encoding REMAP, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing of exons during mRNA processing. A splice variant may have less
25 than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to the polynucleotide sequence encoding REMAP over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide sequence encoding REMAP. For example, a polynucleotide
30 comprising a sequence of SEQ ID NO:30 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:20. Any one of the splice variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of REMAP.

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 REMAP, some bearing minimal
35 similarity to the polynucleotide sequences of any known and naturally occurring gene, may be

WO 02/057454

PCT/US02/01339

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

Although nucleotide sequences which encode REMAP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring REMAP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding REMAP 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 REMAP 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 REMAP and REMAP 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 REMAP 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:16-30 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

WO 02/057454

PCT/US02/01339

sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding REMAP 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

WO 02/057454

PCT/US02/01339

emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode REMAP may be cloned in recombinant DNA molecules that direct expression of REMAP, 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 REMAP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter REMAP-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 No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of REMAP, 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 REMAP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids

WO 02/057454

PCT/US02/01339

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

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

15 In order to express a biologically active REMAP, the nucleotide sequences encoding REMAP
or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which
contains the necessary elements for transcriptional and translational control of the inserted coding
sequence in a suitable host. These elements include regulatory sequences, such as enhancers,
constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in
polynucleotide sequences encoding REMAP. Such elements may vary in their strength and
20 specificity. Specific initiation signals may also be used to achieve more efficient translation of
sequences encoding REMAP. Such signals include the ATG initiation codon and adjacent sequences,
e.g. the Kozak sequence. In cases where sequences encoding REMAP and its initiation codon and
upstream regulatory sequences are inserted into the appropriate expression vector, no additional
transcriptional or translational control signals may be needed. However, in cases where only coding
25 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 enhancers appropriate for the particular host cell system used.
(See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

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

WO 02/057454

PCT/US02/01339

16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding REMAP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, *supra*; Ausubel, *supra*; Van 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 REMAP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding REMAP 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 REMAP 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 REMAP are needed, e.g. for the production of antibodies, vectors which direct high level expression of REMAP 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 REMAP. 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

WO 02/057454

PCT/US02/01339

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 REMAP. Transcription of sequences encoding REMAP 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.)

10 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 REMAP 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 REMAP 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.

20

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.)

25

For long term production of recombinant proteins in mammalian systems, stable expression of REMAP in cell lines is preferred. For example, sequences encoding REMAP 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.

30 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.

35 Any number of selection systems may be used to recover transformed cell lines. These

WO 02/057454

PCT/US02/01339

include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* 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 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), β glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) *Methods Mol. Biol.* 55:121-131.)

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 REMAP is inserted within a marker gene sequence, transformed cells containing sequences encoding REMAP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding REMAP 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 REMAP and that express REMAP 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 REMAP 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 REMAP 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

WO 02/057454

PCT/US02/01339

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 REMAP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. 5 Alternatively, the sequences encoding REMAP, 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 10 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 radioisotopes, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding REMAP may be cultured under 15 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 REMAP may be designed to contain signal sequences which direct secretion of REMAP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the 20 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.

25 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 30 sequences encoding REMAP 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 REMAP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of REMAP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available 35 affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST),

WO 02/057454

PCT/US02/01339

maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the REMAP encoding sequence and the heterologous protein sequence, so that REMAP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

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

REMAP of the present invention or fragments thereof may be used to screen for compounds that specifically bind to REMAP. At least one and up to a plurality of test compounds may be screened for specific binding to REMAP. 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 REMAP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) *Current Protocols in Immunology* 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which REMAP 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 REMAP, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing REMAP or cell membrane fractions which contain REMAP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either REMAP or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with REMAP, either in solution or affixed to a solid support, and detecting the binding of REMAP to the compound.

WO 02/057454

PCT/US02/01339

Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

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

In another embodiment, polynucleotides encoding REMAP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For
20 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
25 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
30 heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding REMAP 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
35 into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al.

WO 02/057454

PCT/US02/01339

(1998) Science 282:1145-1147).

Polynucleotides encoding REMAP 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 REMAP is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress REMAP, e.g., by secreting REMAP in its milk, may also serve as a convenient source of that protein (Jaane, 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 REMAP and receptors and membrane-associated proteins. In addition, examples of tissues expressing REMAP can be found in Table 6. Therefore, REMAP appears to play a role in cell proliferative, autoimmune/inflammatory, neurological, metabolic, developmental, and endocrine disorders. In the treatment of disorders associated with increased REMAP expression or activity, it is desirable to decrease the expression or activity of REMAP. In the treatment of disorders associated with decreased REMAP expression or activity, it is desirable to increase the expression or activity of REMAP.

Therefore, in one embodiment, REMAP 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 REMAP. 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,

WO 02/057454

PCT/US02/01339

glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hyper eosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic

5 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 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

10 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-

15 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, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders,

20 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

25 frontotemporal dementia; a metabolic disorder such as Addison's disease, cerebrotendinous xanthomatosis, congenital adrenal hyperplasia, coumarin resistance, cystic fibrosis, fatty hepatocirrhosis, fructose-1,6-diphosphatase deficiency, galactosemia, goiter, glucagonoma, glycogen storage diseases, hereditary fructose intolerance, hyperadrenalism, hypoadrenalism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia, hyperthyroidism, hypoglycemia,

30 hypothyroidism, hyperlipidemia, hyperlipemia, lipid myopathies, lipodystrophies, lysosomal storage diseases, mannosidosis, neuraminidase deficiency, obesity, osteoporosis, phenylketonuria, pseudovitamin D-deficiency rickets, disorders of carbohydrate metabolism such as congenital type II dyserythropoietic anemia, diabetes, insulin-dependent diabetes mellitus, non-insulin-dependent diabetes mellitus, galactose epimerase deficiency, glycogen storage diseases, lysosomal storage

35 diseases, fructosuria, pentosuria, and inherited abnormalities of pyruvate metabolism, disorders of

WO 02/057454

PCT/US02/01339

lipid metabolism such as fatty liver, cholestasis, primary biliary cirrhosis, carnitine deficiency, carnitine palmitoyltransferase deficiency, myoadenylate deaminase deficiency, hypertriglyceridemia, lipid storage disorders such Fabry's disease, Gaucher's disease, Niemann-Pick's disease, metachromatic leukodystrophy, adrenoleukodystrophy, GM₂ gangliosidosis, and ceroid

5 lipofuscinosis, abetalipoproteinemia, Tangier disease, hyperlipoproteinemia, lipodystrophy, lipomatosis, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipoid adrenal hyperplasia, minimal change disease, lipomas, atherosclerosis, hypercholesterolemia, hypercholesterolemia with hypertriglyceridemia, primary hypoalphalipoproteinemia, hypothyroidism, renal disease, liver disease, lecithin:cholesterol acyltransferase deficiency, cerebrotendinous

10 xanthomatosis, sitosterolemia, hypocholesterolemia, Tay-Sachs disease, Sandhoff's disease, hyperlipidemia, hyperlipemia, and lipid myopathies, and disorders of copper metabolism such as Menke's disease, Wilson's disease, and Ehlers-Danlos syndrome type IX diabetes; 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,

15 myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, a seizure disorder such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; and

20 an endocrine disorder such as a disorder of the hypothalamus and/or pituitary resulting from lesions such as a primary brain tumor, adenoma, infarction associated with pregnancy, hypophysectomy, aneurysm, vascular malformation, thrombosis, infection, immunological disorder, and complication due to head trauma, a disorder associated with hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism, a disorder associated with hyperpituitarism including acromegaly, gigantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma, a disorder associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism, a

30 disorder associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease, a disorder associated with hyperparathyroidism including Conn disease (chronic hypercalcemia), a pancreatic disorder such as Type I or Type II diabetes mellitus and associated complications, a disorder associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal

35 cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's disease,

WO 02/057454

PCT/US02/01339

Liddle's syndrome, and Arnold-Healy-Gordon syndrome, pheochromocytoma tumors, and Addison's disease, a disorder associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbation of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis, and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, a hypergonadal disorder associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α -reductase, and gynecomastia.

In another embodiment, a vector capable of expressing REMAP 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 REMAP including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified REMAP 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 REMAP including, but not limited to, those provided above.

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

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

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding REMAP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of REMAP 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.

WO 02/057454

PCT/US02/01339

An antagonist of REMAP may be produced using methods which are generally known in the art. In particular, purified REMAP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind REMAP. Antibodies to REMAP 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 REMAP 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 REMAP 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 REMAP 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 REMAP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030; and Cole, S.P. et al. (1984) *Mol. Cell Biol.* 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the 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 REMAP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) *Proc. Natl. Acad. Sci. USA* 88:10134-10137.)

WO 02/057454

PCT/US02/01339

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.)

5 Antibody fragments which contain specific binding sites for REMAP may also be generated. For example, such fragments include, but are not limited to, F(ab)₂ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab)₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. 10 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 REMAP and its 15 specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering REMAP 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 REMAP. Affinity is expressed as an 20 association constant, K_a, which is defined as the molar concentration of REMAP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple REMAP epitopes, represents the average affinity, or avidity, of the antibodies for REMAP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific 25 for a particular REMAP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10⁹ to 10¹² L/mole are preferred for use in immunoassays in which the REMAP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10⁶ to 10⁷ L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of REMAP, 30 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 35 example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml,

WO 02/057454

PCT/US02/01339

preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of REMAP-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*.)

5 In another embodiment of the invention, the polynucleotides encoding REMAP, 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 REMAP. Such technology is well known in the art, and antisense oligonucleotides or larger
10 fragments can be designed from various locations along the coding or control regions of sequences encoding REMAP. (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
15 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*
20 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.)

25 In another embodiment of the invention, polynucleotides encoding REMAP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency
30 (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), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii)
35 express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated

WO 02/057454

PCT/US02/01339

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

Expression vectors that may be effective for the expression of REMAP include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). REMAP may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, supra), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding REMAP 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

WO 02/057454

PCT/US02/01339

these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to REMAP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding REMAP 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 No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well 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 REMAP to cells which have one or more genetic abnormalities with respect to the expression of REMAP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding REMAP to target cells which have one or more genetic abnormalities with respect to the expression of REMAP. The use of herpes simplex virus (HSV)-based vectors may be

WO 02/057454

PCT/US02/01339

especially valuable for introducing REMAP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 5 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this 10 patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of 15 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 REMAP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based 20 on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for REMAP into the 25 alphavirus genome in place of the capsid-coding region results in the production of a large number of REMAP-coding RNAs and the synthesis of high levels of REMAP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy 30 application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of REMAP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the 35 art.

WO 02/057454

PCT/US02/01339

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

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

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

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis.

Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding REMAP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

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

WO 02/057454

PCT/US02/01339

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 REMAP.

5 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
10 promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased REMAP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding REMAP may be therapeutically useful, and in the treatment of disorders associated with decreased REMAP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding REMAP may be therapeutically useful.

15 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
20 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 REMAP 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
25 REMAP 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 REMAP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a
30 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; Amdt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa
35 cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular

WO 02/057454

PCT/US02/01339

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).

5 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.)

10 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.

15 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 REMAP, antibodies to REMAP, and mimetics, agonists, antagonists, or inhibitors of

20 REMAP.

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.

25 Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the

30 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.

35 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.

WO 02/057454

PCT/US02/01339

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising REMAP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, REMAP 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
5 transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs,
10 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.

A therapeutically effective dose refers to that amount of active ingredient, for example REMAP or fragments thereof, antibodies of REMAP, and agonists, antagonists or inhibitors of
15 REMAP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Compositions which exhibit large
20 therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the
25 subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and
30 response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

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

WO 02/057454

PCT/US02/01339

inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind REMAP may be used for the
5 diagnosis of disorders characterized by expression of REMAP, or in assays to monitor patients being
treated with REMAP or agonists, antagonists, or inhibitors of REMAP. Antibodies useful for
diagnostic purposes may be prepared in the same manner as described above for therapeutics.
Diagnostic assays for REMAP include methods which utilize the antibody and a label to detect
REMAP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or
10 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 REMAP, including ELISAs, RIAs, and FACS, are
known in the art and provide a basis for diagnosing altered or abnormal levels of REMAP expression.
15 Normal or standard values for REMAP expression are established by combining body fluids or cell
extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to
REMAP under conditions suitable for complex formation. The amount of standard complex
formation may be quantitated by various methods, such as photometric means. Quantities of REMAP
expressed in subject, control, and disease samples from biopsied tissues are compared with the
20 standard values. Deviation between standard and subject values establishes the parameters for
diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding REMAP 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
25 and quantify gene expression in biopsied tissues in which expression of REMAP may be correlated
with disease. The diagnostic assay may be used to determine absence, presence, and excess
expression of REMAP, and to monitor regulation of REMAP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide
sequences, including genomic sequences, encoding REMAP or closely related molecules may be used
30 to identify nucleic acid sequences which encode REMAP. The specificity of the probe, whether it is
made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a
conserved motif, and the stringency of the hybridization or amplification will determine whether the
probe identifies only naturally occurring sequences encoding REMAP, allelic variants, or related
sequences.

35 Probes may also be used for the detection of related sequences, and may have at least 50%

WO 02/057454

PCT/US02/01339

sequence identity to any of the REMAP 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:16-30 or from genomic sequences including promoters, enhancers, and introns of the REMAP gene.

Means for producing specific hybridization probes for DNAs encoding REMAP include the cloning of polynucleotide sequences encoding REMAP or REMAP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ³²P or ³⁵S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding REMAP may be used for the diagnosis of disorders associated with expression of REMAP. 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 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 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,

WO 02/057454

PCT/US02/01339

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

5 nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders,

10 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

15 frontotemporal dementia; a metabolic disorder such as Addison's disease, cerebrotendinous xanthomatosis, congenital adrenal hyperplasia, coumarin resistance, cystic fibrosis, fatty hepatocirrhosis, fructose-1,6-diphosphatase deficiency, galactosemia, goiter, glucagonoma, glycogen storage diseases, hereditary fructose intolerance, hyperadrenalism, hypoadrenalism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia, hyperthyroidism, hypoglycemia,

20 hypothyroidism, hyperlipidemia, hyperlipemia, lipid myopathies, lipodystrophies, lysosomal storage diseases, mannosidosis, neuraminidase deficiency, obesity, osteoporosis, phenylketonuria, pseudovitamin D-deficiency rickets, disorders of carbohydrate metabolism such as congenital type II dyserythropoietic anemia, diabetes, insulin-dependent diabetes mellitus, non-insulin-dependent diabetes mellitus, galactose epimerase deficiency, glycogen storage diseases, lysosomal storage

25 diseases, fructosuria, pentosuria, and inherited abnormalities of pyruvate metabolism, disorders of lipid metabolism such as fatty liver, cholestasis, primary biliary cirrhosis, carnitine deficiency, carnitine palmitoyltransferase deficiency, myoadenylate deaminase deficiency, hypertriglyceridemia, lipid storage disorders such as Fabry's disease, Gaucher's disease, Niemann-Pick's disease, metachromatic leukodystrophy, adrenoleukodystrophy, GM₂ gangliosidosis, and ceroid

30 lipofuscinosis, abetalipoproteinemia, Tangier disease, hyperlipoproteinemia, lipodystrophy, lipomatosis, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipid adrenal hyperplasia, minimal change disease, lipomas, atherosclerosis, hypercholesterolemia, hypercholesterolemia with hypertriglyceridemia, primary hypoalphalipoproteinemia, hypothyroidism, renal disease, liver disease, lecithin:cholesterol acyltransferase deficiency, cerebrotendinous

35 xanthomatosis, sitosterolemia, hypocholesterolemia, Tay-Sachs disease, Sandhoff's disease,

WO 02/057454

PCT/US02/01339

hyperlipidemia, hyperlipemia, and lipid myopathies, and disorders of copper metabolism such as Menke's disease, Wilson's disease, and Ehlers-Danlos syndrome type IX diabetes; 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, 5 myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, a seizure disorder such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; and 10 an endocrine disorder such as a disorder of the hypothalamus and/or pituitary resulting from lesions such as a primary brain tumor, adenoma, infarction associated with pregnancy, hypophysectomy, aneurysm, vascular malformation, thrombosis, infection, immunological disorder, and complication due to head trauma, a disorder associated with hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe 15 disease, sarcoidosis, empty sella syndrome, and dwarfism, a disorder associated with hyperpituitarism including acromegaly, gigantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma, a disorder associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism, a 20 disorder associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease, a disorder associated with hyperparathyroidism including Conn disease (chronic hypercalcemia), a pancreatic disorder such as Type I or Type II diabetes mellitus and associated complications, a disorder associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's disease, 25 Liddle's syndrome, and Arnold-Healy-Gordon syndrome, pheochromocytoma tumors, and Addison's disease, a disorder associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbation of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, 30 hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis, and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, a hypergonadal disorder associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α -reductase, and gynecomastia. The polynucleotide sequences encoding REMAP may be used in Southern or northern analysis, dot blot, or other membrane-based 35 technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in

WO 02/057454

PCT/US02/01339

microarrays utilizing fluids or tissues from patients to detect altered REMAP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding REMAP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding REMAP may be labeled by standard methods and added to a fluid or tissue
5 sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding REMAP in the
10 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 REMAP, a normal or standard profile for expression is established. This may be accomplished by
15 combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding REMAP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values
20 obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from
25 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
30 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 REMAP may involve the use of PCR. These oligomers may be chemically synthesized, generated
35 enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide

WO 02/057454

PCT/US02/01339

encoding REMAP, or a fragment of a polynucleotide complementary to the polynucleotide encoding REMAP, 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.

5 In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding REMAP 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, 10 oligonucleotide primers derived from the polynucleotide sequences encoding REMAP 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 15 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 20 of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA). SNPs may be used to study the genetic basis of human disease. For example, at least 16 25 common SNPs have been associated with non-insulin-dependent diabetes mellitus. SNPs are also useful for examining differences in disease outcomes in monogenic disorders, such as cystic fibrosis, sickle cell anemia, or chronic granulomatous disease. For example, variants in the mannose-binding lectin, MBL2, have been shown to be correlated with deleterious pulmonary outcomes in cystic fibrosis. SNPs also have utility in pharmacogenomics, the identification of genetic variants that influence a patient's response to a drug, such as life-threatening toxicity. For example, a variation in 30 N-acetyl transferase is associated with a high incidence of peripheral neuropathy in response to the anti-tuberculosis drug isoniazid, while a variation in the core promoter of the ALOX5 gene results in diminished clinical response to treatment with an anti-asthma drug that targets the 5-lipoxygenase pathway. Analysis of the distribution of SNPs in different populations is useful for investigating genetic drift, mutation, recombination, and selection, as well as for tracing the origins of populations 35 and their migrations. (Taylor, J.G. et al. (2001) Trends Mol. Med. 7:507-512; Kwok, P.-Y. and Z. Gu

WO 02/057454

PCT/US02/01339

(1999) *Mol. Med. Today* 5:538-543; Nowotny, P. et al. (2001) *Curr. Opin. Neurobiol.* 11:637-641.).

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

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

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

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seifhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

WO 02/057454

PCT/US02/01339

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

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

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

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under

WO 02/057454

PCT/US02/01339

given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for REMAP to quantify the levels of REMAP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; 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 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 useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of

WO 02/057454

PCT/US02/01339

each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

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

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

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

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

WO 02/057454

PCT/US02/01339

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 REMAP 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, REMAP, 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 REMAP 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 REMAP, or fragments thereof, and washed. Bound REMAP is then detected by methods well known in the art. Purified REMAP 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 REMAP specifically compete with a test compound for binding REMAP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with REMAP.

In additional embodiments, the nucleotide sequences which encode REMAP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on

WO 02/057454

PCT/US02/01339

properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

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

The disclosures of all patents, applications and publications, mentioned above and below, including U.S. Ser. No. 60/262,838, U.S. Ser. No. 60/265,927, U.S. Ser. No. 60/271,196, U.S. Ser. No. 60/274,549, and U.S. Ser. No. 60/334,179, are expressly incorporated by reference herein.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA). Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

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

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSOFT 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 02/057454

PCT/US02/01339

PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.B.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.

WO 02/057454

PCT/US02/01339

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and 5 BLOCKS, PRINTS, DOMO, PRODOM; PROTEOME databases with sequences from Homo sapiens, Rattus norvegicus, Mus musculus, Caenorhabditis elegans, Saccharomyces cerevisiae, Schizosaccharomyces pombe, and Candida albicans (Incyte Genomics, Palo Alto CA); 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. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV 15 and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues 20 of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, the PROTEOME databases, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San 25 Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of 30 Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the 35 strength of a match between two sequences (the higher the score or the lower the probability value,

WO 02/057454

PCT/US02/01339

the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:16-30. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative receptors and membrane-associated proteins were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin (1998) Curt. 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 receptors and membrane-associated proteins, the encoded polypeptides were analyzed by querying against PFAM models for receptors and membrane-associated proteins. Potential receptors and membrane-associated proteins were also identified by homology to Incyte cDNA sequences that had been annotated as receptors and membrane-associated proteins. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

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

WO 02/057454

PCT/US02/01339

based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants.

Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpi public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

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

VI. Chromosomal Mapping of REMAP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:16-30 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:16-30 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

WO 02/057454

PCT/US02/01339

Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

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

VII. Analysis of Polynucleotide Expression

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

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

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

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

WO 02/057454

PCT/US02/01339

BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79%
5 identity and 100% overlap.

Alternatively, polynucleotide sequences encoding REMAP 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
10 classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number
15 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 REMAP. cDNA sequences and cDNA library/tissue
20 information are found in the LIFSEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of REMAP 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
25 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.

30 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 Mg²⁺, (NH₄)₂SO₄,
35 and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme

WO 02/057454

PCT/US02/01339

(Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN 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 electrophoresis on a 1% agarose gel to determine which reactions were successful in extending the sequence.

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

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

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides

WO 02/057454

PCT/US02/01339

designed for such extension, and an appropriate genomic library.

IX. Identification of Single Nucleotide Polymorphisms in REMAP Encoding

Polynucleotides

Common DNA sequence variants known as single nucleotide polymorphisms (SNPs) were identified in SEQ ID NO:16-30 using the LIFESEQ database (Incyte Genomics). Sequences from the same gene were clustered together and assembled as described in Example III, allowing the identification of all sequence variants in the gene. An algorithm consisting of a series of filters was used to distinguish SNPs from other sequence variants. Preliminary filters removed the majority of basecall errors by requiring a minimum Phred quality score of 15, and removed sequence alignment errors and errors resulting from improper trimming of vector sequences, chimeras, and splice variants. An automated procedure of advanced chromosome analysis analysed the original chromatogram files in the vicinity of the putative SNP. Clone error filters used statistically generated algorithms to identify errors introduced during laboratory processing, such as those caused by reverse transcriptase, polymerase, or somatic mutation. Clustering error filters used statistically generated algorithms to identify errors resulting from clustering of close homologs or pseudogenes, or due to contamination by non-human sequences. A final set of filters removed duplicates and SNPs found in immunoglobulins or T-cell receptors.

Certain SNPs were selected for further characterization by mass spectrometry using the high throughput MASSARRAY system (Sequenom, Inc.) to analyze allele frequencies at the SNP sites in four different human populations. The Caucasian population comprised 92 individuals (46 male, 46 female), including 83 from Utah, four French, three Venezuelan, and two Amish individuals. The African population comprised 194 individuals (97 male, 97 female), all African Americans. The Hispanic population comprised 324 individuals (162 male, 162 female), all Mexican Hispanic. The Asian population comprised 126 individuals (64 male, 62 female) with a reported parental breakdown of 43% Chinese, 31% Japanese, 13% Korean, 5% Vietnamese, and 8% other Asian. Allele frequencies were first analyzed in the Caucasian population; in some cases those SNPs which showed no allelic variance in this population were not further tested in the other three populations.

X. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:16-30 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ - 32 P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a

WO 02/057454

PCT/US02/01339

SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10⁷ counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

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

XI. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the
15 aforementioned technologies should be uniform and solid with a non-porous surface (Schna (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to
20 those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schna, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)

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

Tissue or Cell Sample Preparation

35 Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and

WO 02/057454

PCT/US02/01339

poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μ l 5X SSC/0.2% SDS.

15 Microarray Preparation

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

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

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

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

WO 02/057454

PCT/US02/01339

0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μ l of sample mixture consisting of 0.2 μ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

Detection

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

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

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from 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

WO 02/057454

PCT/US02/01339

(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 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).

XII. Complementary Polynucleotides

Sequences complementary to the REMAP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring REMAP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of REMAP. 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 REMAP-encoding transcript.

XIII. Expression of REMAP

Expression and purification of REMAP is achieved using bacterial or virus-based expression systems. For expression of REMAP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac operator* regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express REMAP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of REMAP 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 REMAP 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.

WO 02/057454

PCT/US02/01339

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, REMAP 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 REMAP 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 REMAP obtained by these methods can be used directly in the assays shown in Examples XVII, XVIII, and XIX, where applicable.

XIV. Functional Assays

REMAP function is assessed by expressing the sequences encoding REMAP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as

WO 02/057454

PCT/US02/01339

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 REMAP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding REMAP 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 REMAP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XV. Production of REMAP Specific Antibodies

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

XVI. Purification of Naturally Occurring REMAP Using Specific Antibodies

Naturally occurring or recombinant REMAP is substantially purified by immunoaffinity chromatography using antibodies specific for REMAP. An immunoaffinity column is constructed by covalently coupling anti-REMAP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing REMAP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of REMAP (e.g., high ionic strength

WO 02/057454

PCT/US02/01339

buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/REMAP 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 REMAP is collected.

XVII. Identification of Molecules Which Interact with REMAP

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

Alternatively, molecules interacting with REMAP 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).

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

XVIII. Demonstration of REMAP Activity

An assay for REMAP activity measures the expression of REMAP on the cell surface. cDNA
20 encoding REMAP is transfected into an appropriate mammalian cell line. Cell surface proteins are labeled with biotin as described (de la Fuente, M.A. et al. (1997) *Blood* 90:2398-2405). Immunoprecipitations are performed using REMAP-specific antibodies, and immunoprecipitated samples are analyzed using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled
25 immunoprecipitant is proportional to the amount of REMAP expressed on the cell surface.

In the alternative, an assay for REMAP activity is based on a prototypical assay for ligand/receptor-mediated modulation of cell proliferation. This assay measures the rate of DNA synthesis in Swiss mouse 3T3 cells. A plasmid containing polynucleotides encoding REMAP is added to quiescent 3T3 cultured cells using transfection methods well known in the art. The
30 transiently transfected cells are then incubated in the presence of [³H]thymidine, a radioactive DNA precursor molecule. Varying amounts of REMAP ligand are then added to the cultured cells. Incorporation of [³H]thymidine into acid-precipitable DNA is measured over an appropriate time interval using a radioisotope counter, 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
35 REMAP ligand concentration range is indicative of receptor activity. One unit of activity per

WO 02/057454

PCT/US02/01339

milliliter is defined as the concentration of REMAP producing a 50% response level, where 100% represents maximal incorporation of [³H]thymidine into acid-precipitable DNA (McKay, I. and I. Leigh, eds. (1993) *Growth Factors: A Practical Approach*, Oxford University Press, New York NY, p. 73.)

5 In a further alternative, the assay for REMAP activity is based upon the ability of GPCR family proteins to modulate G protein-activated second messenger signal transduction pathways (e.g., cAMP; Gaudin, P. et al. (1998) *J. Biol. Chem.* 273:4990-4996). A plasmid encoding full length REMAP is transfected into a mammalian cell line (e.g., Chinese hamster ovary (CHO) or human embryonic kidney (HEK-293) cell lines) using methods well-known in the art. Transfected cells are
10 grown in 12-well trays in culture medium for 48 hours, then the culture medium is discarded, and the attached cells are gently washed with PBS. The cells are then incubated in culture medium with or without ligand for 30 minutes, then the medium is removed and cells lysed by treatment with 1 M perchloric acid. The cAMP levels in the lysate are measured by radioimmunoassay using methods well-known in the art. Changes in the levels of cAMP in the lysate from cells exposed to ligand
15 compared to those without ligand are proportional to the amount of REMAP present in the transfected cells.

To measure changes in inositol phosphate levels, the cells are grown in 24-well plates containing 1×10^6 cells/well and incubated with inositol-free media and [³H]myoinositol, 2 μ Ci/well, for 48 hr. The culture medium is removed, and the cells washed with buffer containing 10 mM LiCl
20 followed by addition of ligand. The reaction is stopped by addition of perchloric acid. Inositol phosphates are extracted and separated on Dowex AG1-X8 (Bio-Rad) anion exchange resin, and the total labeled inositol phosphates counted by liquid scintillation. Changes in the levels of labeled inositol phosphate from cells exposed to ligand compared to those without ligand are proportional to the amount of REMAP present in the transfected cells.

25 In a further alternative, the ion conductance capacity of REMAP is demonstrated using an electrophysiological assay. REMAP is expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector encoding REMAP. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. A small amount of a second plasmid, which expresses any one of a number of
30 marker genes such as β -galactosidase, is co-transformed into the cells in order to allow rapid identification of those cells which have taken up and expressed the foreign DNA. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression and accumulation of REMAP and β -galactosidase. Transformed cells expressing β -galactosidase are stained blue when a suitable colorimetric substrate is added to the culture media
35 under conditions that are well known in the art. Stained cells are tested for differences in membrane

WO 02/057454

PCT/US02/01339

conductance due to various ions by electrophysiological techniques that are well known in the art. Untransformed cells, and/or cells transformed with either vector sequences alone or β -galactosidase sequences alone, are used as controls and tested in parallel. The contribution of REMAP to cation or anion conductance can be shown by incubating the cells using antibodies specific for either REMAP.

5 The respective antibodies will bind to the extracellular side of REMAP, thereby blocking the pore in the ion channel, and the associated conductance.

In a further alternative, REMAP transport activity is assayed by measuring uptake of labeled substrates into *Xenopus laevis* oocytes. Oocytes at stages V and VI are injected with REMAP mRNA (10 ng per oocyte) and incubated for 3 days at 18 °C in OR2 medium (82.5 mM NaCl, 2.5 mM KCl, 1 mM CaCl₂, 1 mM MgCl₂, 1 mM Na₂HPO₄, 5 mM Hepes, 3.8 mM NaOH, 50 μ g/ml gentamycin, pH 7.8) to allow expression of REMAP protein. Oocytes are then transferred to standard uptake medium (100 mM NaCl, 2 mM KCl, 1 mM CaCl₂, 1 mM MgCl₂, 10 mM Hepes/Tris pH 7.5). Uptake of various substrates (e.g., amino acids, sugars, drugs, and neurotransmitters) is initiated by adding a ³H substrate to the oocytes. After incubating for 30 minutes, uptake is terminated by washing the

15 oocytes three times in Na⁺-free medium, measuring the incorporated ³H, and comparing with controls. REMAP activity is proportional to the level of internalized ³H substrate.

In a further alternative, REMAP protein kinase (PK) activity is measured by phosphorylation of a protein substrate using gamma-labeled [³²P]-ATP and quantitation of the incorporated radioactivity using a gamma radioisotope counter. REMAP is incubated with the protein substrate,

20 [³²P]-ATP, and an appropriate kinase buffer. The ³²P incorporated into the product is separated from free [³²P]-ATP by electrophoresis and the incorporated ³²P is counted. The amount of ³²P recovered is proportional to the PK activity of REMAP in the assay. A determination of the specific amino acid residue phosphorylated is made by phosphoamino acid analysis of the hydrolyzed protein.

XIX. Identification of REMAP Ligands

25 REMAP is expressed in a eukaryotic cell line such as CHO (Chinese Hamster Ovary) or HEK (Human Embryonic Kidney) 293 which have a good history of GPCR expression and which contain a wide range of G-proteins allowing for functional coupling of the expressed REMAP to downstream effectors. The transformed cells are assayed for activation of the expressed receptors in the presence of candidate ligands. Activity is measured by changes in intracellular second messengers, such as

30 cyclic AMP or Ca²⁺. These may be measured directly using standard methods well known in the art, or by the use of reporter gene assays in which a luminescent protein (e.g. firefly luciferase or green fluorescent protein) is under the transcriptional control of a promoter responsive to the stimulation of protein kinase C by the activated receptor (Milligan, G. et al. (1996) Trends Pharmacol. Sci. 17:235-237). Assay technologies are available for both of these second messenger systems to allow high

35 throughput readout in multi-well plate format, such as the adenylyl cyclase activation FlashPlate

WO 02/057454

PCT/US02/01339

Assay (NEN Life Sciences Products), or fluorescent Ca^{2+} indicators such as Fluo-4 AM (Molecular Probes) in combination with the FLIPR fluorimetric plate reading system (Molecular Devices). In cases where the physiologically relevant second messenger pathway is not known, REMAP may be coexpressed with the G-proteins $G_{\alpha 12/16}$ which have been demonstrated to couple to a wide range of G-proteins (Offermanns, S. and M.I. Simon (1995) *J. Biol. Chem.* 270:15175-15180), in order to funnel the signal transduction of the REMAP through a pathway involving phospholipase C and Ca^{2+} mobilization. Alternatively, REMAP may be expressed in engineered yeast systems which lack endogenous GPCRs, thus providing the advantage of a null background for REMAP activation screening. These yeast systems substitute a human GPCR and G_{α} protein for the corresponding components of the endogenous yeast pheromone receptor pathway. Downstream signaling pathways are also modified so that the normal yeast response to the signal is converted to positive growth on selective media or to reporter gene expression (Broach, J.R. and J. Thorner (1996) *Nature* 384 (suppl.):14-16). The receptors are screened against putative ligands including known GPCR ligands and other naturally occurring bioactive molecules. Biological extracts from tissues, biological fluids and cell supernatants are also screened.

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

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
71924779	1	71924779CD1	16	71924779CH1
2319430	2	2319430CD1	17	2319430CH1
7291877	3	7291877CD1	18	7291877CH1
1218126	4	1218126CD1	19	1218126CH1
7479161	5	7479161CD1	20	7479161CH1
7722591	6	7722591CD1	21	7722591CH1
2173285	7	2173285CD1	22	2173285CH1
7487619	8	7487619CD1	23	7487619CH1
7487607	9	7487607CD1	24	7487607CH1
7487616	10	7487616CD1	25	7487616CH1
7483204	11	7483204CD1	26	7483204CH1
7472099	12	7472099CD1	27	7472099CH1
7485443	13	7485443CD1	28	7485443CH1
3090414	14	3090414CD1	29	3090414CH1
7503710	15	7503710CD1	30	7503710CH1

WO 02/057454

PCT/US02/01339

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability Score	GenBank Homolog
1	7132479CD1	g3213188	8.5E-245	[Rattus norvegicus] retinoid acid receptor alpha 2 isoform 1 (Rma1, K.M. et al. (1996) Biol. Reprod. 54:111-119; Rma1, K.M. et al. (1997) Biol. Reprod. 56:542-556; Rma1, K.M. et al. (1998) Endocrinology 139:1233-1246)
2	2319430CD1	g3395787	9.7E-61	[Homo sapiens] multiple membrane spanning receptor TRC8 (Gemmill, R.M. et al. (1998) Proc. Natl. Acad. Sci. USA 95:9572-9577)
3	7291877CD1	g3834380	2.1E-26	[Rattus norvegicus] intrinsic factor-B12 receptor precursor (Moestrup, S.K. et al. (1996) J. Biol. Chem. 271:5235-5242)
4	1218126CD1	g6650766	1.7E-30	[Homo sapiens] PDZ domain-containing guanine nucleotide exchange factor 1
5	7479161CD1	g57734	6.3E-187	[Rattus rattus] potential ligand-binding protein (Dear, T.N. et al. (1991) EMBO J. 10:2813-2819)
6	7722591CD1	g3449308	0.0	[Homo sapiens] MEGF8 (Nakayama, M. et al. (1998) Genomics 51:27-34)
7	2173285CD1	g13539682	0.0	[Homo sapiens] Golgi-associated microtubule-binding protein HOOK3 (Walenta, J.H. et al. (2001) J. Cell Biol. 152:923-934)
8	7487619CD1	g1256393	7.1E-94	[Rattus norvegicus] taste bud receptor protein TB 641 (Thomas, M.B. et al. (1996) Gene 178:1-5)
9	7487607CD1	g5869918	4.5E-115	[Mus musculus] olfactory receptor (Strotmann, J. et al. (1999) Gene 236:281-291)
10	7487616CD1	g1256393	1.6E-94	[Rattus norvegicus] taste bud receptor protein TB 641 (Thomas, M.B. et al. (1996) Gene 178:1-5)
11	7483204CD1	g1314663	9.7E-129	[Canis familiaris] canine olfactory receptor CfOlf2 (Lassel-Parker, L. and J. Rine (1996) Proc. Natl. Acad. Sci. USA 93:10997-10992)

WO 02/057454

PCT/US02/01339

Table 2 (cont.)

Poly peptide Seq ID NO:	Incyte Poly peptide ID	GenBank ID NO:	Probability Score	GenBank Homolog
12	747209CD1	g4680260	7.5E-81	[Mus musculus] odorant receptor S18 (Mahnic, E. et al. (1999) Cell 96:713-723)
13	7485443CD1	g5809918	2.7E-117	[Mus musculus] olfactory receptor (Strotmann, J. et al. (1999) Gene 236:281-291)
14	3090414CD1	g4761597	3.4E-94	[Mus musculus] MOR 3 ^β (Bulger, M. et al. (2000) Proc. Natl. Acad. Sci. USA 97:14560-14565)
15	7503710CD1	g11877275	3.3E-207	[Homo sapiens] dJ726C3.4 (ortholog of potential ligand binding protein RVA3 (Rat))

WO 02/057454

PCT/US02/01339

Table 3

SEQ ID No.	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Software
1	71924779CD1	457	S61, S62, S98, S152, S227, S364, S383, S467, T248, T280, T332	N50, N63, N206, N247, N288, M448	Stand-binding domain of nuclear hormone receptor: E225-S383 Zinc Finger, C4 type (two domains): K31-D156, A326-R333 Transmembrane domains: E291-E298, N294-Q310 N-terminus is cytosolic Nuclear hormone receptor BL0031: C83-M115, Y117-R148 Nuclear hormone receptors DNA-binding region signature: S63-V130 Nuclear hormone receptors DNA-binding region signature: C83-R109 C4-type steroid receptor zinc finger signature PR00047: A99-N114, R132-L140, L140-M148, C83-A99 Vitamin D receptor signature PR00350: C83-A99, C100-C119, G322-D341, M368-F391 Steroid hormone receptor signature PR00398: C83-C100, K104-H120, C125-C143, C230-T249, Q252-Y272, M279-E293 Retinoic acid receptor, transcription regulation, DNA-binding, zinc finger PD002627: K150-S224 PD005136: M1-P82 Hormone family receptor, transcription regulation, DNA-binding, zinc finger PD000035: C83-S149 PD000112: F223-L329, N294-S383	HMMP-PTAM HMMP-PTAM FMRP BLIMPS-ELCKAS ProfileScan MOTIFS BLIMPS-PRINTS BLIMPS-PRINTS BLIMPS-PRINTS BLIMPS-PRINTS BLIMPS-PRINTS BLIMPS-PRINTS BLAST-PRODOM BLAST-PRODOM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1					Nuclear hormone receptor, DNA-binding DM00047 E10276 77-345: S72-D341 DM00047 E10028 70-338: S72-D340 DM00047 E22603 10-338: S72-D340 DM00047 E22603 17-372: S72-D340 DM00047 E22646 77-345: S72-D340 Zinc finger, C3HC4 type (RING finger) F199671: C337-C374 Transmembrane domains: A9-W29, Y3-F73, L60-V100, I128-L148, E166-E186, G218-F246, S272-Q300, T316-V336, I344-I364, A382-Q402, I413-L433, E462-I490 N-terminus is cytosolic Neurotrophin (GAP-43) signature: Q606-H658 Photosystem II proteins. PF00421: T143-L197, G213-S248, L280-G314, V206-D251, C271-A304 Multiple membrane spanning receptor FPC9, patched PF195068: A3-I536 Signal peptide: M1-G60 CUB domain: C65-V170, C241-Y342 Sushi domain (SCR repeat): C178-C235 Transmembrane domain: K35-Y63 N-terminus is cytosolic Adenosine A2B receptor signature PF00554: S353-I364, M263-D277 Sushi domain proteins PF00084: G197-Y208, A226-C235	BLAST-DBO HMMER-PFAM
2	2319430CD1	663	S103 S248 S266 S356 S632 E179 T457 Y109	N455 N580		HMMER-PFAM TMAP
3	7291877CD1	504	S120 S146 S152 S318 S353 S400 S404 S424 S475 T68 T192 T244 T373 T410 T423 T445	N73 N90 N361 N409		BLAST-PRODOM HMMER HMMER-PFAM TMAP BLIMPS-PFAM BLIMPS-PFAM BLIMPS-PFAM BLIMPS-PFAM BLIMPS-PFAM

WO 02/057454

PCT/US02/01339

Table 3 (cont.)

SEQ ID No.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Software	
3					Glycoprotein domain, EGF-like, Yeastar Factor B-12 PD000165; C241-Y342, C65-Y170, C18/C19 consensus DV00162, A571901896-947: Y63-E172, C241-Y342, I205201438-559: C65-E172, C241-Y342, P980631438-549: C65-Y170, C241-Y342, P980631393-417: A240-Q243, T64-E171	BLAST-DOMO	
4	121816CD1	1114	S11 S62 S98 S106 S140 S214 S267 S313 S375 S384 S509 S605 S649 S656 S711 S775 S781 S843 S869 S894 S1012 T145 T164 T177 T191 T243 T519 T620 T666 T748 T796 T804 T877 T899 T1017 Y531	N189 N196 N222 N250 N274 N287 N372 N383 N773 N884			
5	747916CD1	479	S352 S165 S377 S408 S422 S461 T144 T316 T386	N142 N288 N457	signal cleavage: M1-R18 signal peptide: M1-P20 Transmembrane domains: A7-I34, V63-V98, L161-P181, V186-L206, E211-S231, N-terminus is cytosolic	SPFGN HMMER TMHSP	

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residue#	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Software
5					<p>PORENTIAL LIGAND-BINDING PROTEIN NVA3: PD177882; F66-F261 PD053120; V4-R85</p> <p>PROTEIN PRECURSOR SIGNAL CYCLOPROTEIN LIPID TRANSPORT ANTIBIOTIC TRANSMEMBRANE LIPOPOLYSACCHARIDE-BINDING LBP: PD00646; C155-R468, L80-L119</p> <p>EGF-like domain: C39-C99, C103-P131, C1152-C1178</p> <p>Ca²⁺-binding RGF-like domain: C57-A58, C1124-Y1139</p> <p>Laminin-type EGF-like (LE) domain: C209-C227, C447-C465, D165-C193, C447-C465</p>	<p>BLAST-PRODOM</p> <p>BLAST-PRODOM</p> <p>BLAST-PRODOM</p> <p>BLAST-PRODOM</p>
6	772259ICD1	1774	<p>T102 T171</p> <p>S324 S391</p> <p>S343 S386</p> <p>T467 T478</p> <p>T499 S528</p> <p>S530 T535</p> <p>S549 T535</p> <p>S733 S811</p> <p>S830 T850</p> <p>T880 S908</p> <p>S1160 S1314</p> <p>S1203 S1327</p> <p>S1243 S1351</p> <p>S1259 S1374</p> <p>S1287 T1317</p> <p>S1341 S1343</p> <p>T1414 T1435</p>	<p>N100 N126</p> <p>N111 N226</p> <p>N176 N284</p> <p>N449 N585</p> <p>N819 N898</p> <p>N919 N995</p> <p>N1179</p> <p>N1148</p> <p>N1179</p> <p>N1704</p> <p>N1285</p> <p>N1357</p> <p>N1557</p> <p>N1768</p>	<p>EGF-like domain: C39-C99, C103-P131, C1152-C1178</p> <p>Ca²⁺-binding RGF-like domain: C57-A58, C1124-Y1139</p> <p>Laminin-type EGF-like (LE) domain: C209-C227, C447-C465, D165-C193, C447-C465</p>	<p>HMMER-PFAM</p> <p>HMMER-PFAM</p> <p>HMMER-PFAM</p> <p>HMMER-PFAM</p> <p>TMAP</p> <p>BLIMPS-BLOCKS</p> <p>BLIMPS-BLOCKS</p> <p>PRINTS</p>

WO 02/057454

PCT/US02/01339

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Software
6			TI408 S1538 TI1734		RCF-like domain signature 1: C216-C227, C418-C423, C454-C465, C1167-C1178 RCF-like domain signature 2: C85-C99, C418-C423, C454-C465, C1133-C1148, C1167-C1178 Aspartic acid and asparagine hydroxylation site: C16-C67, C1124-C1135 Calcium-binding SRF-like domain signature: D93-C93 Laminin-type B9F-like (LR) domain signature: C169-C193, C216-C244, C1199-C2028, C1280-C1306 PROTEIN HOOK HOOK2 PD016676: E11-E171 HOOK2 PROTEIN PD100271: L185-S322	MOTIFS MOTIFS MOTIFS MOTIFS MOTIFS MOTIFS BLAST- PRODOM BLAST- PRODOM BLAST- PRODOM SPPSCAN HMMER-PFAM
7	2173285CD1	393	S190 S332 S298 S310 S314 S322 S353 S366 T31 T99 T181 T182 T249 T258 T291 T370	N228	PROTEIN COILED COIL CHAIN MYOSIN REPEAT HEAVY APPENDING FILAMENT HEP7AD PD000002: L174-S377	BLAST- PRODOM BLAST- PRODOM BLAST- PRODOM
8	7487619CD1	311	S65 S87 S192 S263 T136 T288	N3 N63	Signal peptide: M1-G39 7 transmembrane receptor (rhodopsin family): G39-Y287 Predicted transmembrane segments: L21-V49, M57-M81, A90-T115, S192-T220, A236-F256, V267-Y287 N-terminus cytosolic C-protein coupled receptors proteins BL00237: E231-T257, T279-E295, R89-F128	SPPSCAN HMMER-PFAM TMAP BLIWP8- BLOCKS

WO 02/057454

PCT/US02/01339

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
9					Olfactory receptor signature PR00245: M68-S60, F177-L191, S338-G235, Y280-M291, E227-L311 RECEPTOR PROTEIN RECEPTOR-LIKE GPROTEIN COUPLED TRANSMEMBRANE GLYCOPROTEIN MULTIGENE FAMILY: F000921: L166-L245 F000921: L246-F313 G-PROTEIN COUPLED RECEPTORS DK00013: K23275 17-306: L47-L311 K57069 15-304: S18-L311 F37067 17-306: L47-L310 S51356 18-307: L47-V307 G-protein coupled receptors signature: T110-L126	BLIMP3-PRINTS BLAST-PRODOM BLAST-DOMO
10	7487616CD1	311	865-887 S192 T136 Y288	N3 N63 N303	7 transmembrane receptor (rhodopsin family): G39-Y287 Predicted transmembrane segments: L21-V49, M57-M81, I91-T115, M194-C222, A236-F236, L267-Y287; N-terminus non-cytosolic G-protein coupled receptors proteins SL000237: R89-F128, E231-I257, I279-K295 G-protein coupled receptors signature G-protein receptor.prf: Y101-M47 Opsins retinal binding site opsins.prf: Y258-E311 Olfactory receptor signature PR00245: M57-K78, Y176-D190, F237-G252, Y271-F282, T288-K302	HMMER-FFAM TRAP BLIMP3-BLOCKS PROFILESCAN PROFILESCAN BLIMP3-PRINTS

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
10					Metacortin receptor family signature PR00594; V49-L61, I125-V136 RECEPTORY PROTEIN RECEPTOR-LIKE PROTEIN COUPLED TRANSMEMBRANE GLYCOPROTEIN MULTIGENE FAMILY: PD149621: L66-L73 PD149621: V246-L249	BLAST-DMO BLAST-DMO PRODOM
11	7483204CD1	310	S6 S67 S137 S186 S224 S291	N5 R38 R89 N135	G-PROTEIN COUPLED RECEPTORS DMC0013: S135E118_306: L5-L26 P23274_18_306: F26-L201 P23266_17_306: L15-L201 P23270_18_311: A18-L298 Signal Peptide: R23-R42 7 transmembrane receptor (rhodopsin family): A41-Y290 predicted transmembrane segments: P21-L49, Y95-Y123, S138-R165, L197-V229, G233-R261, N-terminus cytosolic G-protein coupled receptors proteins R100237: K90-P129, E232-R238, Y282-R298 G-protein coupled receptors signature 5 protein receptor, PFI: L102-G152 Olfactory receptor signature PR00245: M59-R60, F177-D191, L238-G253, E274-L285, S291-K305 RECEPTOR COUPLED TRANSMEMBRANE LIKE G-PROTEIN COUPLED TRANSMEMBRANE GLYCOPROTEIN MULTIGENE FAMILY: PD000921: L166-L245 PD149621: S246-L304	HMMER HMMER-FFAM TWAP BLTMS-BLOCKS PROFILESCAN BLTMS-PRINTS BLAST-PRODOM

WO 02/057454

PCT/US02/01339

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
11					G-PROTEIN COUPLED RECEPTORS DM00013: S81356 18-307; I17-L301 F37067 17-306; I17-L304 S29709 11-299; T18-K303 P30955 18-305; N20-L304 G-protein coupled receptors signature: S110-I126	BLAST-DMO
12	7472099CD1	316	S109 S230 S306	N43	7 transmembrane receptor (rhodopsin family): G42-V293 Predicted transmembrane segments: I19-L47, I60-P88, F95-R123, M43-R171, P195-I222, K237-V257, A266-P286; N-terminus cytosolic G-protein coupled receptors proteins BL00237; R01-F130, F252-R263, E233-S259, F285-R301 G-protein coupled receptors signature 5_protein_receptor.pri: F103-R152 Olfactory receptor signature PR00245: P239-P254, L277-V288, I60-K81, C178-T192 PUTATIVE GPROTEIN COUPLED RECEPTOR RALC P2170483: I246-R301 PRODOM G-PROTEIN COUPLED RECEPTORS DM00013: E23275 17-306; H24-L307 G45774 18-309; P20-L304 F45774 19-309; Y36-L304 P34982 17-305; P20-I307 G-protein coupled receptors signature: M11-I127	HMMER-EPAM TWAP BLIMPS-BLOCKS PROFILESAM BLIMPS-PRINTS BLAST-PRODOM BLAST-DMO MOTIFS

WO 02/057454

PCT/US02/01339

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
13	7485443CD1	318	S49, S67, S67, S193, S297, I276	N5, N42, N65	7 transmembrane receptor (rhodopsin family); G41-Y236 Predicted transmembrane segments: F21-G41, I46-F68, S95-Y123, Y141-V164, G169-C189, G194-I222, G233-K261; N-terminus non-cytosolic G-protein coupled receptors proteins HM00237: K90-P129, I207-Y218, S232-W258, F288-K304 G-protein coupled receptors signature G-protein_receptor.prf: F102-G147 Olfactory receptor signature PR00245: M59-S80, F177-D191, F238-G253, I280-M291, S297-F311 RECEPTOR OLFACTORY PROTEIN RECEPTOR-LIKE GPROTEIN COUPLED TRANSMEMBRANE GLYCOPROTEIN MULTIGENE FAMILY PD149621: F246-F315 G-PROTEIN COUPLED RECEPTORS DM00013: P23275 I7-306; L17-L310 A57069 I5-304; S18-L310 S51356 I8-307; L17-V307 P37067 I7-306; L17-L310 G-protein coupled receptors signature: T110-I126	HMER-PFAM TMAP
14	3090414CD1	321	S69, S231, S263, T110, T165, T179	N5	7 transmembrane receptor (rhodopsin family); G43-Y295 Predicted transmembrane segments: I27-H52, Y62-H90, I94-L119, I143-R167, M197-R225, A239-G267, I270-P287; N-terminus non-cytosolic	HMER-PFAM TMAP

WO 02/057454

PCT/US02/01339

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
14					G-protein coupled receptors Proteins BL000237; D234-S260, E287-R303, R32-P131 Olfactory receptor signature PR00245; M51-R62, T179-N193, F240-T255 RECEPTOR OLFACTORY PROTEIN RECEPTOR-LIKE GPROTEIN COUPLED TRANSMEMBRANE GLYCOPROTEIN MULTIGENE FAMILY ED000921; L168-L247 G-PROTEIN COUPLED RECEPTORS DM00013; G45774 18-309; P20-S311 P23273 18-306; F33-L310 P23274 18-306; F33-L310 S29708 18-306; S35-L310 G-protein coupled receptors signature: M112-L128	BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO
15	7503710cd1	422	S152 S165 S351 S365 S404 T144 T316 T329	N142 N288 N400	signal cleavage: M1-A18 signal peptide: M5-P20, M5-E23, M1-P20, M1-T24, M1-T27, M1-E23 Ribosomal protein L1 proteins BL01199; L168-E161 POTENTIAL LIGAND-BINDING PROTEIN RY3 ED053120; V4-R65 ED177882; P86-P261 ED053120; V4-R65 PROTEIN PRECURSOR SIGNAL GLYCOPROTEIN LIPID TRANSPORT ANTIETHIOTIC TRANSMEMBRANE LIPOPOLYSACCHARIDE-BINDING LBP ED006440: A332-M41L, G155-S357, L80-L119 Go LIGAND; RX265; RY3 DM05385; S17448 1-473; V4-L347, L237-A422 S17447 1-470; G66-I343, D246-M41L, G58-G223, G58-V85	SPSCAN HMMER BLIMPS-BLOCKS BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM BLAST-DOMO

Table 4 (cont.)

Polynucleotide seq. ID NO./ nucleotide sequence accession number	Sequence Fragments
21/7722591CBI/ 6056	711-1181, 714-1147, 724-1149, 745-1135, 746-1238, 751-883, 758-1160, 763-1327, 765-1465, 766-1467, 786-1038, 790-985, 791-1234, 791-1465, 799-1280, 801-982, 805-993, 809-1115, 809-1211, 809-1284, 817-1406, 820-1395, 830-1314, 832-1301, 833-1408, 838-1223, 863-1379, 863-1422, 869-1244, 906-1465, 908-1035, 913-1109, 918-1150, 918-1312, 963-1289, 974-1641, 997-1371, 1000-1404, 1001-1460, 1031-1465, 1206-1235, 1206-1252, 1206-1365, 1206-1272, 1206-1274, 1206-1279, 1206-1287, 1206-1299, 1206-1316, 1206-1333, 1206-1336, 1206-1342, 1206-1354, 1206-1355, 1206-1366, 1206-1423, 1206-1452, 1206-1457, 1206-1465, 1206-1466, 1208-1284
21/7722591CBI/ 6056	1-118, 1-130, 1-1697, 4-25, 215-324, 217-275, 777-1047, 779-857, 1068-1135, 1607-2232, 1658-2173, 1753-2343, 1826-2070, 2049-2412, 2059-2599, 2183-2688, 2184-2304, 2285-2855, 2301-2856, 2516-2777, 2609-3040, 2780-2856, 2854-3160, 2856-3240, 2856-3127, 2971-3454, 3122-3233, 3196-3357, 3221-3616, 3231-3291, 3238-3663, 3306-3366, 3399-3637, 3431-4058, 3464-3549, 3514-3881, 3514-4173, 3671-3977, 3747-4136, 3985-4595, 4096-4643, 4138-4643, 4182-4795, 4286-4784, 4294-4441, 4294-4532, 4294-4548, 4294-4548, 4298-4441, 4412-4640, 4501-5250, 4532-4864, 4553-4648, 4580-4648, 4777-4983, 4777-5122, 4777-5259, 4479-6056, 4784-5021, 4891-5432, 4952-5205, 5526-5637
22/2173285CBI/ 1699	1-441, 1-485, 26-314, 52-726, 58-303, 66-727, 74-608, 154-428, 478-633, 566-757, 566-1065, 572-487, 789-955, 820-1076, 820-1307, 830-1182, 1095-1699, 1097-1699, 1438-1490, 1530-1587
23/7487616CBI/ 1661	1-978, 76-867, 118-867, 128-918, 130-916, 132-867, 154-866, 174-918, 177-912, 219-917, 281-918, 726-1661, 729-953, 736-953, 885-971, 885-1084, 888-1080, 894-1084
24/7487607CBI/ 2033	2-715, 7-715, 20-715, 36-715, 508-1445, 563-950, 1246-1505, 1246-1904, 1246-1958, 1246-1963, 1963, 1246-1564, 1246-1065, 1246-1868, 1246-2033
25/7487616CBI/ 1659	1-826, 572-724, 481-768, 581-777, 891-777, 709-921, 709-926, 709-933, 764-1301, 764-1486, 748-1532, 745-1441, 768-1330, 759-1330, 784-1659, 784-1659, 795-1528, 795-1542, 795-1384, 795-1495
26/7483204CBI/ 6125	1-1175, 215-758, 456-665
27/7472093CBI/ 1737	1-210, 39-641, 388-1061, 425-1025, 442-1065, 481-1117, 484-1102, 532-900, 656-1312, 925-982, 1104-1595, 1104-1614, 1106-1737, 1258-1646

WO 02/057454

PCT/US02/01339

Table 4 (cont.)

Polynucleotide SEQ ID NO./ Accession ID/ Gene ID/ Gene Name	Sequence Fragments
28/746543CBI/ 972	1-215, 2-215, 16-972, 115-972, 187-377, 334-972, 773-972
29/709041CBI/ 1592	1-722, 28-722, 40-723, 43-718, 59-719, 60-722, 66-722, 91-722, 163-710, 159-709, 214-463, 402-722, 403-722, 416-778, 416-928, 416-978, 427-1085, 429-715, 441-589, 441-634, 441- 711, 441-713, 441-856, 441-877, 441-983, 441-1035, 441-1039, 454-960, 468-1083, 591-1128, 594-907, 594-1286, 594-1318, 607-1084, 619-829, 640-1247, 663-1285, 699-982, 706-919, 739-1002, 755-1285, 758-1318, 774-1060, 774-1309, 783-1041, 789-1205, 789-1395, 806-1365, 807-1338, 842-1089, 847-1374, 901-1365, 917-1380, 941-1380, 954-1592, 983-1376, 987-1348, 1059-1413
30/750371CBI/ 1480	1-436, 1-439, 1-469, 1-507, 1-557, 1-561, 1-593, 1-620, 1-640, 2-613, 7-171, 7-491, 7- 559, 7-577, 7-622, 11-200, 11-329, 11-382, 11-495, 11-534, 11-562, 11-611, 11-618, 11- 673, 11-677, 11-1286, 11-1480, 13-644, 13-696, 14-612, 46-699, 155-895, 205-694, 210-826, 217-697, 227-750, 240-356, 270-360, 263-388, 297-928, 348-869, 357-885, 415-843, 459-597, 459-658, 507-1193, 514-1100, 544-1123, 547-1172, 556-1111, 556-1113, 556-1191, 579-795, 581-1138, 586-1118, 608-1068, 622-1205, 622-1262, 627-890, 655-1104, 657-1175, 688-1236, 689-656, 692-1126, 774-1303, 796-1045, 804-1304, 815-1002, 843-1154, 884-1123, 884-1126, 899-1180, 921-1114, 937-1101, 1067-1181

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
16	71924779CB1	SITIMR01
17	2319430CB1	LUNGAST01
18	7251877CB1	BRALFER05
19	1218126CB1	PHOS2NV34
20	7479161CB1	NOSBDC02
21	7722591CB1	SEAVIDE01
22	2173285CB1	BRANDTC4
23	7487619CB1	GPCRDPV04
24	7487607CB1	GPCRDPV02
25	7487616CB1	GPCRDPV02
27	7472099CB1	EMARNO102
28	7485443CB1	GPCRDPV02
29	3090414CB1	BRSTN0T19
30	7503710CB1	NOSBDC02

WO 02/057454

PCT/US02/01339

Table 6

Library	Vector	Library Description
BMAR0702	PELJESCRIFT	Library was constructed using RNA isolated from the bone marrow of 24 male and female Caucasian donors, 16 to 70 years old. (RNA came from Clontech.)
BRA1P805	PINCY	Library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks' gestation.
BRANDIT04	PINCY	Library was constructed using RNA isolated from pineal gland tissue removed from a 68-year-old Caucasian female who died from congestive heart failure. Neuropathology indicated mild to moderate Alzheimer disease, atherosclerosis, and multiple infarctions. Microscopically, there were diffuse and neuritic amyloid plaques throughout the cerebral cortex, there were neurofibrillary tangles in the temporal lobes particularly the entorhinal cortex. The frontal cortex contained scattered, ballooned neurons. The amygdala contained marked gliosis, neuritic plaques and intracellular neurofibrillary tangles. The hippocampus contained neuritic and diffuse plaques, and neurofibrillary tangles. The thalamus contained diffuse and focal neuritic amyloid plaques and scattered neurofibrillary tangles. There was area of cystic cavitation with surrounding gliosis in the left globus pallidus. The pallidum contained scattered intracellular neurofibrillary tangles. The caudate, putamen and nucleus accumbens contained diffuse plaques. There was an area of cystic cavitation with lipid-laden macrophages in the right cerebellar hemisphere. Patient history included diabetes mellitus, rheumatoid arthritis, hyperthyroidism, amyloid heart disease, and dementia.
BRSTW019	PINCY	Library was constructed using RNA isolated from breast tissue removed from a 67-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated residual invasive lobular carcinoma. Patient history included depressive disorder, benign large bowel neoplasm, and hemorrhoids. Family history included cerebrovascular and cardiovascular disease and lung cancer.
GFCDP002	PCR2-TOPOVA	Library was constructed using pooled cDNA from different donors. cDNA was generated using mRNA isolated from the following: aorta, cerebellum, lymph nodes, muscle, tonsil (lymphoid hyperplasia), bladder tumor (invasive grade 3 transitional cell carcinoma), breast (proliferative fibrocystic changes without atypia characterized by epithelial ductal hyperplasia, testicle tumor (embryonal carcinoma), spleen, ovary, parathyroid, ileum, breast skin, sigmoid colon, penis

Table 6 (cont.)

Library GPKDFV02 (cont.)	Vector	Library Description
		<p>tumor (fungating invasive grade 4 squamous cell carcinoma), fetal lung, breast, fetal small intestine, fetal liver, fetal pancreas, fetal lung, fetal skin, fetal penis, fetal bone, fetal ribs, frontal brain tumor (grade 4 gemistocytic astrocytoma), ovary (stromal hyperthecosis), bladder, bladder tumor (invasive grade 3 transitional cell carcinoma), stomach, lymph node tumor (metastatic basaloid squamous cell carcinoma), tonsil (reactive lymphoid hyperplasia), periosteum from the tibia, fetal brain, fetal spleen, uterus tumor, endometrial (grade 3 adenocarcinoma), seminal vesicle, liver, aorta, adrenal gland, lymph node (metastatic grade 3 squamous cell carcinoma), glomus muscle, esophagus, esophagus tumor (invasive grade 3 adenocarcinoma), ileum, pancreas, soft tissue tumor from the skull (grade 3 ependymoma), transverse colon, (benign familial polyposis), rectum tumor (grade 3 colonic adenocarcinoma), fib tumor (metastatic grade 3 osteosarcoma), lung, heart, placenta, thymus, stomach, spleen (splenomegaly with congestion), uterus, cervix (mild chronic cervicitis with focal squamous metaplasia), spleen tumor (malignant lymphoma, diffuse large cell type, B-cell phenotype with abundant reactive T-cells and marked granulomatous response), umbilical cord blood mononuclear cells, upper lobe lung tumor, (grade 3 squamous cell carcinoma), endometrium (secretory phase), liver, liver tumor (metastatic grade 2 neuroendocrine carcinoma), colon, umbilical cord blood, Th1 artery endothelial cells (untreated), coronary artery smooth muscle cells, (untreated), coronary artery smooth muscle cells (treated with TNF & IL-1 10ng/ml each for 20 hrs), bladder (mild chronic cystitis), epiglottis, breast skin, small intestine, fetal prostate stroma fibroblasts, prostate epithelial cells (PRSC cells), fetal adrenal glands, fetal liver, kidney transformed embryonal cell line (Z93-EMNA) (untreated), kidney transformed embryonal cell line (HMC cells) (treated with Shiga-2deoxycytidine for 72 hours), mammary epithelial cells, (HMC cells), peripheral blood monocytes (treated with IL-10 at time 0, 10ng/ml, 10ng/ml, 10ng/ml at 1 hour at 5ng/ml, incubation 24 hrs), peripheral blood monocytes (treated with anti-IL-10 at time 0, 10ng/ml, 10ng/ml, 10ng/ml at 1 hour at 5ng/ml, incubation 24 hrs), splenic cord, spleen of medulla (Huntington's chorea), thigh and arm muscle (GCS), breast skin fibroblasts (untreated), breast skin fibroblasts (treated with GCS), breast skin fibroblasts (treated with TNF-alpha & IL-1 beta, 10ng/ml each for 20 hrs), fetal liver mast cells, hematopoietic (mast</p>

WO 02/057454

PCT/US02/01339

Table 6 (cont.)

Library GRCRDP02 (cont.)	Vector	Library Description
		<p>cells prepared from human fetal liver hematopoietic progenitor cells (CD34+ stem cells) cultured in the presence of hr-5 and hSCR for 18 days), epithelial layer of colon, bronchial epithelial cells (treated for 20hrs with 20% smoke conditioned media), lymph node, pooled peripheral blood mononuclear cells (untreated), pooled brain segments: striatum, globus pallidus and posterior putamen (Alzheimer's Disease), pituitary gland, umbilical cord blood, CD34+ derived dendritic cells (treated with SCF, GM-CSF & TNF alpha, 13 days), umbilical cord blood, CD34+ derived dendritic cells (treated with SCF, GM-CSF & TNF alpha, 13 days followed by PMA/Ionomycin for 5 hours), small intestine, rectum, bone marrow neuroblastoma cell line (SH-SY5Y cells, untreated), brain segments from one donor: amygdala, entorhinal cortex, globus pallidus, substantia innominata, striatum, dorsal caudate nucleus, dorsal putamen, ventral nucleus accumbens, archaocortex (hippocampus anterior and posterior), thalamus, nucleus raphe magnus, periaqueductal gray, midbrain, substantia nigra, and dentate nucleus, pineal gland (Alzheimer's Disease), preadipocytes (untreated), preadipocytes (treated with a peroxisome proliferator-activated receptor gamma agonist, Imcrom, 4 hours), Pooled prostate (Adenofibromatous hyperplasia), pooled kidney, pooled adipocytes (untreated), pooled adipocytes (treated with human insulin), pooled mesenteric and abdominal fat, pooled adrenal glands, pooled thyroid (normal and adenomatous hyperplasia), pooled spleen (normal and with changes consistent with idiopathic thrombocytopenic purpura), pooled right and left breast, pooled lung, pooled nasal Polyps, pooled fat, pooled synovium (normal and rheumatoid arthritis), pooled brain (meningioma, gemistocytic astrocytoma, and Alzheimer's disease), pooled fetal colon, pooled colon: ascending, descending (chronic ulcerative colitis), and rectal tumor (adenocarcinoma), pooled esophagus, normal and tumor (invasive grade 3 adenocarcinoma), pooled breast skin fibroblast (one treated w/5α Retinoic Acid and the other with TNF-alpha & IL-1 beta), pooled gallbladder (acute necrotizing cholecystitis with cholelithiasis (clinically hydrops), acute hemorrhagic cholecystitis with cholelithiasis, chronic cholecystitis and cholelithiasis), pooled fetal heart, (Patau's and fetal demise), pooled neurogenic tumor cell line, SK-N-SIC, neuroepithelioma, metastasis to supra-orbital area, untreated) and neuron, NT-2 cell line, (treated with mouse leptin at 1 μg/ml and 5αs retinoic</p>

WO 02/057454

PCT/US02/01339

Table 6 (cont.)

Library	Vector	Library Description
GRCRNV02 (cont.)		acid at 3.3 μM for 6 days), pooled ovary (normal and polycystic ovarian disease), pooled prostate, (adenofibromatous hyperplasia), pooled seminal vesicle, pooled small intestine, pooled fetal small intestine, pooled stomach and fetal stomach, prostate epithelial cells, pooled testis (normal and embryonal carcinoma), pooled uterus, pooled uterus tumor (grade 3 adenosquamous carcinoma and leiomyoma), pooled uterus, endometrium, and myometrium, (normal and adenomatous hyperplasia with squamous metaplasia and focal atypia), pooled brain: (temporal lobe meningioma, cerebellum and hippocampus (Alzheimer's Disease), and pooled skin.
LUNGAST01	PSP0RV1	Library was constructed using RNA isolated from the lung tissue of a 17-year-old Caucasian male, who died from head trauma. Patient history included asthma.
NOERDICO2	PSP0RV1	This large size fractionated library was constructed using RNA isolated from nasal Polyp tissue.
PHOSDNV34	FCR2-TOFOVA	Library was constructed using pooled cDNA from 111 different donors. cDNA was generated using mRNA isolated from pooled skeletal muscle tissue removed from 10 Caucasian male and female donors, ages 21-57, who died from sudden death; from pooled thymus tissue removed from 9 Caucasian male and female donors, ages 18-32, who died from sudden death; from pooled fetal liver tissue removed from 32 Caucasian male and female fetuses, ages 18-24 weeks, who died from spontaneous abortions; from pooled fetal kidney tissue removed from 59 Caucasian male and female fetuses, ages 20-33 weeks, who died from spontaneous abortions; and from fetal brain tissue removed from a 23-week-old Caucasian male fetus who died from fetal demise.

WO 02/057454

PCT/US02/01339

Table 6 (cont.)

Library	Vector	Library Description
SEM7DR01	PCDM2.1	This 5' biased random primed library was constructed using RNA isolated from seminal vesicle tissue removed from a 63-year-old Caucasian male during closed prostatic biopsy, radical prostatectomy, and regional lymph node excision. Pathology for the associated tumor tissue indicated Gleason grade 2+3 adenocarcinoma in the right side of the prostate. Adenofibromatous hyperplasia was present. The patient presented with prostate cancer, elevated prostate specific antigen and prostatic hyperplasia. Patient history included kidney calculus, extrinsic asthma, benign bowel neoplasm, backache, tremor, and tobacco abuse in remission. Previous surgeries included adenotonsillectomy. Patient medications included Ventolin and Vancocin. Family history included atherosclerotic coronary artery disease and acute myocardial infarction in the mother; atherosclerotic coronary artery disease and acute myocardial infarction in the father; and stomach cancer and extrinsic asthma in the grandparent(s).
SINTTR01	PCDM2.1	This random primed library was constructed using RNA isolated from ileum tissue removed from a 70-year-old Caucasian female during right hemicolectomy, open liver biopsy, flexible sigmoidoscopy, colonoscopy, and permanent colostomy. Pathology for the matched tumor tissue indicated invasive grade 2 adenocarcinoma forming an ulcerated mass, situated 2 cm distal to the ileocecal valve. Patient history included a malignant breast neoplasm, type II diabetes, hyperlipidemia, viral hepatitis, an unspecified thyroid disorder, osteoarthritis, a malignant skin neoplasm, deficiency anemia, and normal delivery. Family history included breast cancer, atherosclerotic coronary artery disease, benign hypertension, cerebrovascular disease, ovarian cancer, and hyperlipidemia.

WO 02/057454

PCT/US02/01339

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.	
ABIPARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) <i>J. Mol. Biol.</i> 215:403-410; Altschul, S.F. et al. (1997) <i>Nucleic Acids Res.</i> 25:3389-3402.	<i>ESTs</i> : Probability values= 1.0E-8 or less <i>Full Length sequences</i> : 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, fastn, fastx, fasty, and search.	Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci. USA</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183:63-98; and Smith, T.F. and M.S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489.	<i>ESTs</i> : fasta E value=1.0E-6 <i>Assembled ESTs</i> : fasta Identity= 95% or greater and Match length=200 bases or greater fasta E value=1.0E-8 or less <i>Full Length sequences</i> : fastx scores=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) <i>Nucleic Acids Res.</i> 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) <i>Methods Enzymol.</i> 266:88-105; and Atwood, T.K. et al. (1997) <i>J. Chem. Inf. Comput. Sci.</i> 37:417-424.	Probability values= 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) <i>J. Mol. Biol.</i> 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) <i>Nucleic Acids Res.</i> 26:320-322; Durbin, R. et al. (1998) <i>Our World View</i> , in a Nutshell, Cambridge Univ. Press, pp. 1-350.	<i>PFAM hits</i> : Probability value= 1.0E-3 or less <i>Signal peptide hits</i> : Score= 0 or greater

WO 02/057454

PCT/US02/01339

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.	Gribisov, M. et al. (1988) CABIOS 4:61-66; Gribisov, M. et al. (1989) Methods Enzymol. 183:146-159; Baroch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores: GCG-specified "HIGP" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	Scores=120 or greater; Match length= 56 or greater
Phrap	A Plus 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. 147:105-197; and Green, P., University of Washington, Seattle, WA.	
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SFScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nelson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Scores=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) Proteins 24:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Somhammer, E.L. et al. (1998) Proc. Sixth Int. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Baroch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 02/057454

PCT/US02/01339

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15,
 - b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-14,
 - c) a polypeptide comprising a naturally occurring amino acid sequence at least 98% identical to the amino acid sequence of SEQ ID NO:15,
 - d) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and
 - e) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.
2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method of producing a polypeptide of claim 1, the method comprising:
 - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant

WO 02/057454

PCT/US02/01339

- polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
- b) recovering the polypeptide so expressed.
- 5 10. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.
11. An isolated antibody which specifically binds to a polypeptide of claim 1.
- 10 12. An isolated polynucleotide selected from the group consisting of:
- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30,
- b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of
- 15 c) a polynucleotide complementary to a polynucleotide of a),
- d) a polynucleotide complementary to a polynucleotide of b), and
- e) an RNA equivalent of a)-d).
- 20 13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.
14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:
- 25 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
- 30 b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.
- 35 16. A method of detecting a target polynucleotide in a sample, said target polynucleotide

WO 02/057454

PCT/US02/01339

having a sequence of a polynucleotide of claim 12, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
- b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

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

18. A composition of claim 17, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.

19. A method for treating a disease or condition associated with decreased expression of functional REMAP, comprising administering to a patient in need of such treatment the composition of claim 17.

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

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

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

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

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

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

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

WO 02/057454

PCT/US02/01339

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

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

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

10

27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1.
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

15

20

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

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

25

30

29. A method of assessing toxicity of a test compound, the method comprising:

- a) treating a biological sample containing nucleic acids with the test compound,
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions

35

WO 02/057454

PCT/US02/01339

- whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,
- 5 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.
- 10 30. A diagnostic test for a condition or disease associated with the expression of REMAP in a biological sample, the method comprising:
- a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and
15 b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.
31. The antibody of claim 11, wherein the antibody is:
- a) a chimeric antibody,
20 b) a single chain antibody,
c) a Fab fragment,
d) a F(ab')₂ fragment, or
e) a humanized antibody.
- 25 32. A composition comprising an antibody of claim 11 and an acceptable excipient.
33. A method of diagnosing a condition or disease associated with the expression of REMAP in a subject, comprising administering to said subject an effective amount of the composition of claim 32.
30
34. A composition of claim 32, wherein the antibody is labeled.
- 35 35. A method of diagnosing a condition or disease associated with the expression of REMAP in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

WO 02/057454

PCT/US02/01339

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:
- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
 - b) isolating antibodies from said animal, and
 - c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.
37. A polyclonal antibody produced by a method of claim 26.
38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.
39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:
- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
 - b) isolating antibody producing cells from the animal,
 - c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
 - d) culturing the hybridoma cells, and
 - e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.
40. A monoclonal antibody produced by a method of claim 39.
41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.
42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.
43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant

WO 02/057454

PCT/US02/01339

immunoglobulin library.

44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15 in a sample, the method comprising:

- 5 a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15 in the sample.

10

45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15 from a sample, the method comprising:

- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- 15 b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

- a) labeling the polynucleotides of the sample,
- 25 b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and
- c) quantifying the expression of the polynucleotides in the sample.

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

35

WO 02/057454

PCT/US02/01339

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.
50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.
51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.
52. An array of claim 48, which is a microarray.
53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.
54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.
55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.
56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.
57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.
58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.
59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.
60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.
61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

WO 02/057454

PCT/US02/01339

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.
63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.
- 5 64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.
65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.
66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.
- 10 67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
- 15 70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.
71. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
20 NO:16.
72. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:17.
- 25 73. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:18.
74. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:19.
- 30 75. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:20.
76. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
35 NO:21.

WO 02/057454

PCT/US02/01339

77. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:22.
78. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
5 NO:23.
79. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:24.
- 10 80. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:25.
81. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:26.
- 15 82. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:27.
83. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
20 NO:28.
84. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:29.
- 25 85. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:30.

WO 02/057454

PCT/US02/01339

<110> INCYTE GENOMICS, INC.
 LEE, Ernestine A.
 WALLA, Narinder K.
 BAUGHN, Mariah R.
 AZIMZAI, Yalda
 TANG, Y. Tom
 YUE, Henry
 THANGAVELU, Kavitha
 XU, Yuming
 ARVIZU, Chandra
 WARREN, Bridget A.
 YAO, Monique G.
 AU-YOUNG, Janice
 HAFALIA, April J.A.
 ELLIOTT, Vicki S.
 KALLICK, Deborah A.
 GANDHI, Ameena R.
 RICHARDSON, Thomas W.
 KHAN, Farrah A.
 LU, Yan
 SWARNAKAR, Anita
 RAMKUMAR, Jayalaxmi
 NGUYEN, Damiel B.
 GRAUL, Richard
 LU, Dyung Aina M.

<120> RECEPTORS AND MEMBRANE-ASSOCIATED PROTEINS

<130> EI-0346 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/262,838; 60/265,927; 60/271,196; 60/274,549; 60/334,179
 <151> 2001-01-19; 2001-02-02; 2001-02-23; 2001-03-09; 2001-11-28

<160> 30

<170> PERL Program

<210> 1
 <211> 457
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 71924779CD1

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

WO 02/057454

PCT/US02/01339

```

Ser Gly Tyr His Tyr Gly Val Ser Ala Cys Glu Gly Cys Lys Gly
95 100
Phe Phe Arg Arg Ser Ile Gln Lys Asn Met Val Tyr Thr Cys His
110 115
Arg Asp Lys Asn Cys Ile Ile Asn Lys Val Thr Arg Asn Arg Cys
125 130
Gln Tyr Cys Arg Leu Gln Lys Cys Phe Glu Val Gly Met Ser Lys
140 145
Glu Ser Val Arg Asn Asp Arg Asn Lys Lys Lys Lys Glu Val Pro
155 160
Lys Pro Glu Cys Ser Glu Ser Tyr Thr Leu Thr Pro Glu Val Gly
170 175
Glu Leu Ile Glu Lys Val Arg Lys Ala His Gln Glu Thr Phe Pro
185 190
Ala Leu Cys Gln Leu Gly Lys Tyr Thr Thr Asn Asn Ser Ser Glu
200 205
Gln Arg Val Ser Leu Asp Ile Asp Leu Trp Asp Lys Phe Ser Glu
215 220
Leu Ser Thr Lys Cys Ile Ile Lys Thr Val Glu Phe Ala Lys Gln
230 235
Leu Pro Gly Phe Thr Thr Leu Thr Ile Ala Asp Gln Ile Thr Leu
245 250
Leu Lys Ala Ala Cys Leu Asp Ile Leu Ile Leu Arg Ile Cys Thr
260 265
Arg Tyr Thr Pro Glu Gln Asp Thr Met Thr Phe Ser Asp Gly Leu
275 280
Thr Leu Asn Arg Thr Gln Met His Asn Ala Gly Phe Gly Pro Leu
290 295
Thr Asp Leu Val Phe Ala Phe Ala Asn Gln Leu Leu Pro Leu Glu
305 310
Met Asp Asp Ala Glu Thr Gly Leu Leu Ser Ala Ile Cys Leu Ile
320 325
Cys Gly Asp Arg Gln Asp Leu Glu Gln Pro Asp Arg Val Asp Met
335 340
Leu Gln Glu Pro Leu Leu Glu Ala Leu Lys Val Tyr Val Arg Lys
350 355
Arg Arg Pro Ser Arg Pro His Met Phe Pro Lys Met Leu Met Lys
365 370
Ile Thr Asp Leu Arg Ser Ile Ser Ala Lys Gly Ala Glu Arg Val
380 385
Ile Thr Leu Lys Met Glu Ile Pro Gly Ser Met Pro Pro Leu Ile
395 400
Gln Glu Met Leu Glu Asn Ser Glu Gly Leu Asp Thr Leu Ser Gly
410 415
Gln Pro Gly Gly Gly Gly Arg Asp Gly Gly Gly Leu Ala Pro Pro
425 430
Pro Gly Ser Cys Ser Pro Ser Leu Ser Pro Ser Ser Asn Arg Ser
440 445
Ser Pro Ala Thr His Ser Pro
455

```

```

<210> 2
<211> 663
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2319430CD1

```

```

<400> 2
Met Ala Ala Lys Glu Lys Leu Glu Ala Val Leu Asn Val Ala Leu
1 5 10

```

WO 02/057454

PCT/US02/01339

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

WO 02/057454

PCT/US02/01339

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

<210> 3
 <211> 504
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7291877CD1

<400> 3
 Met Lys Gly Ile Arg Lys Gly Glu Ser Arg Ala Lys Glu Ser Lys
 1 5 10 15
 Pro Trp Glu Pro Gly Lys Arg Arg Cys Ala Lys Cys Gly Arg Leu
 20 25 30
 Asp Phe Ile Leu Met Lys Lys Met Gly Ile Lys Ser Gly Phe Thr
 35 40 45
 Phe Trp Asn Leu Val Phe Leu Leu Thr Val Ser Cys Val Lys Gly
 50 55 60
 Phe Ile Tyr Thr Cys Gly Gly Thr Leu Lys Gly Leu Asn Gly Thr
 65 70 75
 Ile Glu Ser Pro Gly Phe Pro Tyr Gly Tyr Pro Asn Gly Ala Asn
 80 85 90
 Cys Thr Trp Val Ile Ile Ala Glu Glu Arg Asn Arg Ile Gln Ile
 95 100 105
 Val Phe Gln Ser Phe Ala Leu Glu Glu Glu Tyr Asp Tyr Leu Ser
 110 115 120
 Leu Tyr Asp Gly His Pro His Pro Thr Asn Phe Arg Thr Arg Leu
 125 130 135
 Thr Gly Phe His Leu Pro Pro Pro Val Thr Ser Thr Lys Ser Val
 140 145 150
 Phe Ser Leu Arg Leu Thr Ser Asp Phe Ala Val Ser Ala His Gly
 155 160 165
 Phe Lys Val Tyr Tyr Glu Glu Leu Gln Ser Ser Ser Cys Gly Asn
 170 175 180
 Pro Gly Val Pro Pro Lys Gly Val Leu Tyr Gly Thr Arg Phe Asp
 185 190 195

WO 02/057454

PCT/US02/01339

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

<210> 4
<211> 1114
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 1218126CD1

<400> 4
Met Ala Pro Thr Leu Phe Gln Lys Leu Phe Ser Lys Arg Thr Gly
1 5 10 15
Leu Gly Ala Pro Gly Arg Asp Ala Arg Asp Pro Asp Cys Gly Phe
20 25 30
Ser Trp Pro Leu Pro Glu Phe Asp Pro Ser Gln Ile Arg Leu Ile
35 40 45
Val Tyr Gln Asp Cys Glu Arg Arg Gly Arg Asn Val Leu Phe Asp
50 55 60
Ser Ser Val Lys Arg Arg Asn Glu Asp Ile Ser Val Ser Lys Leu
65 70 75

WO 02/057454

PCT/US02/01339

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

WO 02/057454

PCT/US02/01339

Ser Ser Leu Leu Phe Lys Glu Ser Glu Glu Ile Arg Thr Pro Asn
 545 550 555
 Cys Asn Cys Lys Tyr Cys Ser His Pro Leu Leu Gly Gln Asn Val
 560 565 570
 Glu Asn Ile Ser Gln Gln Glu Arg Glu Asp Ile Gln Asn Ser Ser
 575 580 585
 Lys Glu Leu Leu Gly Ile Ser Asp Glu Cys Arg Met Ile Ser Pro
 590 595 600
 Ser Asp Cys Gln Glu Glu Asn Ala Val Asp Val Lys Gln Tyr Arg
 605 610 615
 Asp Lys Leu Arg Thr Cys Phe Asp Ala Lys Leu Glu Thr Val Val
 620 625 630
 Cys Thr Gly Ser Val Pro Val Asp Lys Cys Ala Leu Ser Glu Ser
 635 640 645
 Gly Leu Glu Ser Thr Glu Glu Thr Trp Gln Ser Glu Lys Leu Leu
 650 655 660
 Asp Ser Asp Ser His Thr Gly Lys Ala Met Arg Ser Thr Gly Met
 665 670 675
 Val Val Glu Lys Lys Pro Pro Asp Lys Ile Val Pro Ala Ser Phe
 680 685 690
 Ser Cys Glu Ala Ala Gln Thr Lys Val Thr Phe Leu Ile Gly Asp
 695 700 705
 Ser Met Ser Pro Asp Ser Asp Thr Glu Leu Arg Ser Gln Ala Val
 710 715 720
 Val Asp Gln Ile Thr Arg His His Thr Lys Pro Leu Lys Glu Glu
 725 730 735
 Arg Gly Ala Ile Asp Gln His Gln Glu Thr Lys Gln Thr Thr Lys
 740 745 750
 Asp Gln Ser Gly Glu Ser Asp Thr Gln Asn Met Val Ser Glu Glu
 755 760 765
 Pro Cys Glu Leu Pro Cys Trp Asn His Ser Asp Pro Glu Ser Met
 770 775 780
 Ser Leu Phe Asp Glu Tyr Phe Asn Asp Asp Ser Ile Glu Thr Arg
 785 790 795
 Thr Ile Asp Asp Val Pro Phe Lys Thr Ser Thr Asp Ser Lys Asp
 800 805 810
 His Cys Cys Met Leu Glu Phe Ser Lys Ile Leu Cys Thr Lys Asn
 815 820 825
 Asn Lys Gln Asn Asn Glu Phe Cys Lys Cys Ile Glu Thr Val Pro
 830 835 840
 Gln Asp Ser Cys Lys Thr Cys Phe Pro Gln Gln Asp Gln Arg Asp
 845 850 855
 Thr Leu Ser Ile Leu Val Pro His Gly Asp Lys Glu Ser Ser Asp
 860 865 870
 Lys Lys Ile Ala Val Gly Thr Glu Trp Asp Ile Pro Arg Asn Glu
 875 880 885
 Ser Ser Asp Ser Ala Leu Gly Asp Ser Glu Ser Glu Asp Thr Gly
 890 895 900
 His Asp Met Thr Arg Gln Val Ser Ser Tyr Tyr Gly Gly Glu Gln
 905 910 915
 Glu Asp Trp Ala Glu Glu Asp Glu Ile Pro Phe Pro Gly Ser Lys
 920 925 930
 Leu Ile Glu Val Ser Ala Val Gln Pro Asn Ile Ala Asn Phe Gly
 935 940 945
 Arg Ser Leu Leu Gly Gly Tyr Cys Ser Ser Tyr Val Pro Asp Phe
 950 955 960
 Val Leu Gln Gly Ile Gly Ser Asp Glu Arg Phe Arg Gln Cys Leu
 965 970 975
 Met Ser Asp Leu Ser His Ala Val Gln His Pro Val Leu Asp Glu
 980 985 990
 Pro Ile Ala Glu Ala Val Cys Ile Ile Ala Asp Met Asp Lys Trp
 995 1000 1005

WO 02/057454

PCT/US02/01339

Thr Val Gln Val Ala Ser Ser Gln Arg Arg Val Thr Asp Asn Lys
 1010 1015 1020
 Leu Gly Lys Glu Val Leu Val Ser Ser Leu Val Ser Asn Leu Leu
 1025 1030 1035
 His Ser Thr Leu Gln Leu Tyr Lys His Asn Leu Ser Pro Asn Phe
 1040 1045 1050
 Cys Val Met His Leu Glu Asp Arg Leu Gln Glu Leu Tyr Phe Lys
 1055 1060 1065
 Ser Lys Met Leu Ser Glu Tyr Leu Arg Gly Gln Met Arg Val His
 1070 1075 1080
 Val Lys Glu Leu Gly Val Val Leu Gly Ile Glu Ser Ser Asp Leu
 1085 1090 1095
 Pro Leu Leu Ala Ala Val Ala Ser Thr His Ser Pro Tyr Val Ala
 1100 1105 1110
 Gln Ile Leu Leu

<210> 5
 <211> 479
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7479161CD1

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

WO 02/057454

PCT/US02/01339

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

<210> 6

<211> 1774

<212> FRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7722591CD1

<400> 6

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

WO 02/057454

PCT/US02/01339

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

WO 02/057454

PCT/US02/01339

Ser Gly Ala Gln Ser Gly Thr Pro Pro Thr Gly Leu Tyr Gly His
 650 655 660
 Ser Ala Val Tyr His Glu Ala Thr Asp Ser Leu Tyr Val Phe Gly
 665 670 675
 Gly Phe Arg Phe His Val Glu Leu Ala Ala Pro Ser Pro Glu Leu
 680 685 690
 Tyr Ser Leu His Cys Pro Asp Arg Thr Trp Ser Leu Leu Ala Pro
 695 700 705
 Ser Gln Gly Ala Lys Pro Arg Pro Arg Leu Phe His Ala Ser Ala
 710 715 720
 Leu Leu Gly Asp Thr Met Val Val Leu Gly Gly Arg Ser Asp Pro
 725 730 735
 Asp Glu Phe Ser Ser Asp Val Leu Leu Tyr Gln Val Asn Cys Asn
 740 745 750
 Ala Trp Leu Leu Pro Asp Leu Thr Arg Ser Ala Ser Val Gly Pro
 755 760 765
 Pro Met Glu Glu Ser Val Ala His Ala Val Ala Ala Val Gly Ser
 770 775 780
 Arg Leu Tyr Ile Ser Gly Gly Phe Gly Gly Val Ala Leu Gly Arg
 785 790 795
 Leu Leu Ala Leu Thr Leu Pro Pro Asp Pro Cys Arg Leu Leu Ser
 800 805 810
 Ser Pro Glu Ala Cys Asn Gln Ser Gly Ala Cys Thr Trp Cys His
 815 820 825
 Gly Ala Cys Leu Ser Gly Asp Gln Ala His Arg Leu Gly Cys Gly
 830 835 840
 Gly Ser Pro Cys Ser Pro Met Pro Arg Ser Pro Glu Glu Cys Arg
 845 850 855
 Arg Leu Arg Thr Cys Ser Glu Cys Leu Ala Arg His Pro Arg Thr
 860 865 870
 Leu Gln Pro Gly Asp Gly Glu Ala Ser Thr Pro Arg Cys Lys Trp
 875 880 885
 Cys Thr Asn Cys Pro Glu Gly Ala Cys Ile Gly Arg Asn Gly Ser
 890 895 900
 Cys Thr Ser Glu Asn Asp Cys Arg Ile Asn Gln Arg Glu Val Phe
 905 910 915
 Trp Ala Gly Asn Cys Ser Glu Ala Ala Cys Gly Ala Ala Asp Cys
 920 925 930
 Glu Gln Cys Thr Arg Glu Gly Lys Cys Met Trp Thr Arg Gln Phe
 935 940 945
 Lys Arg Thr Gly Glu Thr Arg Arg Ile Leu Ser Val Gln Pro Thr
 950 955 960
 Tyr Asp Trp Thr Cys Phe Ser His Ser Leu Leu Asn Val Ser Pro
 965 970 975
 Met Pro Val Glu Ser Ser Pro Pro Leu Pro Cys Pro Thr Pro Cys
 980 985 990
 His Leu Leu Pro Asn Cys Thr Ser Cys Leu Asp Ser Lys Gly Ala
 995 1000 1005
 Asp Gly Gly Trp Gln His Cys Val Trp Ser Ser Ser Leu Gln Gln
 1010 1015 1020
 Cys Leu Ser Pro Ser Tyr Leu Pro Leu Arg Cys Met Ala Gly Gly
 1025 1030 1035
 Cys Gly Arg Leu Leu Arg Gly Pro Glu Ser Cys Ser Leu Gly Cys
 1040 1045 1050
 Ala Gln Ala Thr Gln Cys Ala Leu Cys Leu Arg Arg Pro His Cys
 1055 1060 1065
 Gly Trp Cys Ala Trp Gly Gly Gln Asp Gly Gly Gly Arg Cys Met
 1070 1075 1080
 Glu Gly Gly Leu Ser Gly Pro Arg Asp Gly Leu Thr Cys Gly Arg
 1085 1090 1095
 Pro Gly Ala Ser Trp Ala Phe Leu Ser Cys Pro Pro Glu Asp Glu
 1100 1105 1110

WO 02/057454

PCT/US02/01339

Cys Ala Asn Gly His His Asp Cys Asn Glu Thr Gln Asn Cys His
 1115 1120 1125
 Asp Gln Pro His Gly Tyr Glu Cys Ser Cys Lys Thr Gly Tyr Thr
 1130 1135 1140
 Met Asp Asn Met Thr Gly Leu Cys Arg Pro Val Cys Ala Gln Gly
 1145 1150 1155
 Cys Val Asn Gly Ser Cys Val Glu Pro Asp His Cys Arg Cys His
 1160 1165 1170
 Phe Gly Phe Val Gly Arg Asn Cys Ser Thr Glu Cys Arg Cys Asn
 1175 1180 1185
 Arg His Ser Glu Cys Ala Gly Val Gly Ala Arg Asp His Cys Leu
 1190 1195 1200
 Leu Cys Arg Asn His Thr Lys Gly Ser His Cys Glu Gln Cys Leu
 1205 1210 1215
 Pro Leu Phe Val Gly Ser Ala Val Gly Gly Gly Thr Cys Arg Pro
 1220 1225 1230
 Cys His Ala Phe Cys Arg Gly Asn Ser His Ile Cys Ile Ser Arg
 1235 1240 1245
 Lys Glu Leu Gln Met Ser Lys Gly Glu Pro Lys Lys Tyr Ser Leu
 1250 1255 1260
 Asp Pro Glu Glu Ile Glu Asn Trp Val Thr Glu Gly Pro Ser Glu
 1265 1270 1275
 Asp Glu Ala Val Cys Val Asn Cys Gln Asn Asn Ser Tyr Gly Glu
 1280 1285 1290
 Lys Cys Glu Ser Cys Leu Gln Gly Tyr Phe Leu Leu Asp Gly Lys
 1295 1300 1305
 Cys Thr Lys Cys Gln Cys Asn Gly His Ala Asp Thr Cys Asn Glu
 1310 1315 1320
 Gln Asp Gly Thr Gly Cys Pro Cys Gln Asn Asn Thr Glu Thr Gly
 1325 1330 1335
 Thr Cys Gln Gly Ser Ser Pro Ser Asp Arg Arg Asp Cys Tyr Lys
 1340 1345 1350
 Tyr Gln Cys Ala Lys Cys Arg Glu Ser Phe His Gly Ser Pro Leu
 1355 1360 1365
 Gly Gly Gln Gln Cys Tyr Arg Leu Ile Ser Val Glu Gln Glu Cys
 1370 1375 1380
 Cys Leu Asp Pro Thr Ser Gln Thr Asn Cys Phe His Glu Pro Lys
 1385 1390 1395
 Arg Arg Ala Leu Gly Pro Gly Arg Thr Val Leu Phe Gly Val Gln
 1400 1405 1410
 Pro Lys Phe Thr Asn Val Asp Ile Arg Leu Thr Leu Asp Val Thr
 1415 1420 1425
 Phe Gly Ala Val Asp Leu Tyr Val Ser Thr Ser Tyr Asp Thr Phe
 1430 1435 1440
 Val Val Arg Val Ala Pro Asp Thr Gly Val His Thr Val His Ile
 1445 1450 1455
 Gln Pro Pro Pro Ala Pro Pro Pro Pro Pro Pro Ala Asp Gly
 1460 1465 1470
 Gly Pro Arg Gly Ala Gly Asp Pro Gly Gly Ala Gly Ala Ser Ser
 1475 1480 1485
 Gly Pro Gly Ala Pro Ala Glu Pro Arg Val Arg Glu Val Trp Pro
 1490 1495 1500
 Arg Gly Leu Ile Thr Tyr Val Thr Val Thr Glu Pro Ser Ala Val
 1505 1510 1515
 Leu Val Val Arg Gly Val Arg Asp Arg Leu Val Ile Thr Tyr Pro
 1520 1525 1530
 His Glu His His Ala Leu Lys Ser Ser Arg Phe Tyr Leu Leu Leu
 1535 1540 1545
 Leu Gly Val Gly Asp Pro Ser Gly Pro Gly Ala Asn Gly Ser Ala
 1550 1555 1560
 Asp Ser Gln Gly Leu Leu Phe Phe Arg Gln Asp Gln Ala His Ile
 1565 1570 1575

WO 02/057454

PCT/US02/01339

Asp Leu Phe Val Phe Phe Ser Val Phe Phe Ser Cys Phe Phe Leu
 1580 1585 1590
 Phe Leu Ser Leu Cys Val Leu Leu Trp Lys Ala Lys Gln Ala Leu
 1595 1600 1605
 Asp Gln Arg Gln Glu Gln Arg Arg His Leu Gln Glu Met Thr Lys
 1610 1615 1620
 Met Ala Ser Arg Pro Phe Ala Lys Val Thr Val Cys Phe Pro Pro
 1625 1630 1635
 Asp Pro Thr Ala Pro Ala Ser Ala Trp Lys Pro Ala Gly Leu Pro
 1640 1645 1650
 Pro Pro Ala Phe Arg Arg Ser Glu Pro Phe Leu Ala Pro Leu Leu
 1655 1660 1665
 Leu Thr Gly Ala Gly Gly Pro Trp Gly Pro Met Gly Gly Gly Cys
 1670 1675 1680
 Cys Pro Pro Ala Ile Pro Ala Thr Thr Ala Gly Leu Arg Ala Gly
 1685 1690 1695
 Pro Ile Thr Leu Glu Pro Thr Glu Asp Gly Met Ala Gly Val Ala
 1700 1705 1710
 Thr Leu Leu Leu Gln Leu Pro Gly Gly Pro His Ala Pro Asn Gly
 1715 1720 1725
 Ala Cys Leu Gly Ser Ala Leu Val Thr Leu Arg His Arg Leu His
 1730 1735 1740
 Glu Tyr Cys Gly Gly Gly Gly Gly Ala Gly Gly Ser Gly His Gly
 1745 1750 1755
 Thr Gly Ala Gly Arg Lys Gly Leu Leu Ser Gln Asp Asn Leu Thr
 1760 1765 1770
 Ser Met Ser Leu

<210> 7
 <211> 393
 <212> FMT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> incyte ID No: 2173285CD1

<400> 7
 Met Phe Ser Val Glu Ser Leu Glu Arg Ala Glu Leu Cys Glu Ser
 1 5 10 15
 Leu Leu Thr Trp Ile Gln Thr Phe Asn Val Asp Ala Pro Cys Gln
 20 25 30
 Thr Val Glu Asp Leu Thr Asn Gly Val Val Met Ala Gln Val Leu
 35 40 45
 Gln Lys Ile Asp Pro Ala Tyr Phe Asp Glu Asn Trp Leu Asn Arg
 50 55 60
 Ile Lys Thr Glu Val Gly Asp Asn Trp Arg Leu Lys Ile Ser Asn
 65 70 75
 Leu Lys Lys Ile Leu Lys Gly Ile Leu Asp Tyr Asn His Glu Ile
 80 85 90
 Leu Gly Gln Gln Ile Asn Asp Phe Thr Leu Pro Asp Val Asn Leu
 95 100 105
 Ile Gly Glu His Ser Asp Ala Ala Glu Leu Gly Arg Met Leu Gln
 110 115 120
 Leu Ile Leu Gly Cys Ala Val Asn Cys Glu Gln Lys Gln Glu Tyr
 125 130 135
 Ile Gln Ala Ile Met Met Met Glu Glu Ser Val Gln His Val Val
 140 145 150
 Met Thr Ala Ile Gln Glu Leu Met Ser Lys Glu Ser Pro Val Ser
 155 160 165
 Ala Gly Asn Asp Ala Tyr Val Asp Leu Asp Arg Gln Leu Lys Lys
 170 175 180

WO 02/057454

PCT/US02/01339

Thr Thr Glu Glu Leu Asn Glu Ala Leu Ser Ala Lys Glu Glu Ile
 185 190 195
 Ala Gln Arg Cys His Glu Leu Asp Met Gln Val Ala Ala Leu Gln
 200 205 210
 Glu Glu Lys Ser Ser Leu Leu Ala Glu Asn Gln Val Leu Met Glu
 215 220 225
 Arg Leu Asn Gln Ser Asp Ser Ile Glu Asp Pro Asn Ser Pro Ala
 230 235 240
 Gly Arg Arg His Leu Gln Leu Gln Thr Gln Leu Glu Gln Leu Gln
 245 250 255
 Glu Glu Thr Phe Arg Leu Glu Ala Ala Lys Asp Asp Tyr Arg Ile
 260 265 270
 Arg Cys Glu Glu Leu Glu Lys Glu Ile Ser Glu Leu Arg Gln Gln
 275 280 285
 Asn Asp Glu Leu Thr Thr Leu Ala Asp Glu Ala Gln Ser Leu Lys
 290 295 300
 Asp Glu Ile Asp Val Leu Arg His Ser Ser Asp Lys Val Ser Lys
 305 310 315
 Leu Glu Gly Gln Val Glu Ser Tyr Lys Lys Lys Leu Glu Asp Leu
 320 325 330
 Gly Asp Leu Arg Arg Gln Val Lys Leu Leu Glu Glu Lys Asn Thr
 335 340 345
 Met Tyr Met Gln Asn Thr Val Ser Leu Glu Glu Glu Leu Arg Lys
 350 355 360
 Ala Asn Ala Ala Arg Ser Gln Leu Glu Thr Tyr Lys Arg Gln Val
 365 370 375
 Lys Glu Thr Gln His Leu Asp Asp Gly Phe Arg Gln Ala Leu Ser
 380 385 390
 Tyr Asp Met

<210> 8
 <211> 311
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7487619CD1

<400> 8
 Met Ser Asn Ala Ser Leu Val Thr Ala Phe Ile Leu Thr Gly Leu
 1 5 10 15
 Pro His Ala Pro Gly Leu Asp Ala Pro Leu Phe Gly Ile Phe Leu
 20 25 30
 Val Val Tyr Val Leu Thr Val Leu Gly Asn Leu Leu Ile Leu Leu
 35 40 45
 Val Ile Arg Val Asp Ser His Leu His Thr Pro Met Tyr Tyr Phe
 50 55 60
 Leu Thr Asn Leu Ser Phe Ile Asp Met Trp Phe Ser Thr Val Thr
 65 70 75
 Val Pro Lys Met Leu Met Thr Leu Val Ser Pro Ser Gly Arg Ala
 80 85 90
 Ile Ser Phe His Ser Cys Val Ala Gln Leu Tyr Phe Phe His Phe
 95 100 105
 Leu Gly Ser Thr Glu Cys Phe Leu Tyr Thr Val Met Ala Tyr Asp
 110 115 120
 Arg Tyr Leu Ala Ile Ser Tyr Pro Leu Arg Tyr Thr Ser Met Met
 125 130 135
 Thr Gly Arg Ser Cys Thr Leu Leu Ala Thr Ser Thr Trp Leu Ser
 140 145 150
 Gly Ser Leu His Ser Ala Val Gln Ala Ile Leu Thr Phe His Leu
 155 160 165

WO 02/057454

PCT/US02/01339

Pro Tyr Cys Gly Pro Asn Trp Ile Gln His Tyr Leu Cys Asp Ala
 170 175 180
 Pro Pro Ile Leu Lys Leu Ala Cys Ala Asp Thr Ser Ala Ile Glu
 185 190 195
 Thr Val Ile Phe Val Thr Val Gly Ile Val Ala Ser Gly Cys Phe
 200 205 210
 Val Leu Ile Val Leu Ser Tyr Val Ser Ile Val Cys Ser Ile Leu
 215 220 225
 Arg Ile Arg Thr Ser Glu Gly Lys His Arg Ala Phe Gln Thr Cys
 230 235 240
 Ala Ser His Cys Ile Val Val Leu Cys Phe Phe Gly Pro Gly Leu
 245 250 255
 Phe Ile Tyr Leu Arg Pro Gly Ser Arg Lys Ala Val Asp Gly Val
 260 265 270
 Val Ala Val Phe Tyr Thr Val Leu Thr Pro Leu Leu Asn Pro Val
 275 280 285
 Val Tyr Thr Leu Arg Asn Lys Glu Val Lys Lys Ala Leu Leu Lys
 290 295 300
 Leu Lys Asp Lys Val Ala His Ser Gln Ser Lys
 305 310

<210> 9
 <211> 318
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <223> Incyte ID No: 7487607CD1

<400> 9
 Met Glu Trp Glu Asn His Thr Ile Leu Val Glu Phe Phe Leu Lys
 1 5 15
 Gly Leu Ser Gly His Pro Arg Leu Glu Leu Leu Phe Phe Val Leu
 20 25 30
 Ile Phe Ile Met Tyr Val Val Ile Leu Leu Gly Asn Gly Thr Leu
 35 40 45
 Ile Leu Ile Ser Ile Leu Asp Pro His Leu His Thr Pro Met Tyr
 50 55 60
 Phe Phe Leu Gly Asn Leu Ser Phe Leu Asp Ile Cys Tyr Thr Thr
 65 70 75
 Thr Ser Ile Pro Ser Thr Leu Val Ser Phe Leu Ser Glu Arg Lys
 80 85 90
 Thr Ile Ser Leu Ser Gly Cys Ala Val Gln Met Phe Leu Gly Leu
 95 100 105
 Ala Met Gly Thr Thr Glu Cys Val Leu Leu Gly Met Met Ala Tyr
 110 115 120
 Asp Arg Tyr Val Ala Ile Cys Asn Pro Leu Arg Tyr Pro Ile Ile
 125 130 135
 Met Ser Lys Asp Ala Tyr Val Pro Met Ala Ala Gly Ser Trp Ile
 140 145 150
 Ile Gly Ala Val Asn Ser Ala Val Gln Ser Val Phe Val Val Gln
 155 160 165
 Leu Pro Phe Cys Arg Asn Asn Ile Ile Asn His Phe Thr Cys Glu
 170 175 180
 Ile Leu Ala Val Met Lys Leu Ala Cys Ala Asp Ile Ser Asp Asn
 185 190 195
 Glu Phe Ile Met Leu Val Ala Thr Thr Leu Phe Ile Leu Thr Pro
 200 205 210
 Leu Leu Leu Ile Ile Val Ser Tyr Thr Leu Ile Ile Val Ser Ile
 215 220 225
 Phe Lys Ile Ser Ser Ser Glu Gly Arg Ser Lys Ala Ser Ser Thr
 230 235 240

WO 02/057454

PCT/US02/01339

Cys Ser Ala His Leu Thr Val Val Ile Ile Phe Tyr Gly Thr Ile
 245 250 255
 Leu Phe Met Tyr Met Lys Pro Lys Ser Lys Glu Thr Leu Asn Ser
 260 265 270
 Asp Asp Leu Asp Ala Thr Asp Lys Ile Ile Ser Met Phe Tyr Gly
 275 280 285
 Val Met Thr Pro Met Met Asn Pro Leu Ile Tyr Ser Leu Arg Asn
 290 295 300
 Lys Asp Val Lys Glu Ala Val Lys His Leu Leu Asn Arg Arg Phe
 305 310 315
 Phe Ser Lys

<210> 10
 <211> 311
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7487616CD1

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

WO 02/057454

PCT/US02/01339

Leu Lys Asn Gly Ser Val Phe Ala Gln Gly Glu
305 310

<210> 11
<211> 310
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7483204CD1

<400> 11
Met Asp Trp Glu Asn Cys Ser Ser Leu Thr Asp Phe Phe Leu Leu
1 5 10 15
Gly Ile Thr Asn Asn Pro Glu Met Lys Val Thr Leu Phe Ala Val
20 25 30
Phe Leu Ala Val Tyr Ile Ile Asn Phe Ser Ala Asn Leu Gly Met
35 40 45
Ile Val Leu Ile Arg Met Asp Tyr Gln Leu His Thr Pro Met Tyr
50 55 60
Phe Phe Leu Ser His Leu Ser Phe Cys Asp Leu Cys Tyr Ser Thr
65 70 75
Ala Thr Gly Pro Lys Met Leu Val Asp Leu Leu Ala Lys Asn Lys
80 85 90
Ser Ile Pro Phe Tyr Gly Cys Ala Leu Gln Phe Leu Val Phe Cys
95 100 105
Ile Phe Ala Asp Ser Glu Cys Leu Leu Leu Ser Val Met Ala Phe
110 115 120
Asp Arg Tyr Lys Ala Ile Ile Asn Pro Leu Leu Tyr Thr Val Asn
125 130 135
Met Ser Ser Arg Val Cys Tyr Leu Leu Leu Thr Gly Val Tyr Leu
140 145 150
Val Gly Ile Ala Asp Ala Leu Ile His Met Thr Leu Ala Phe Arg
155 160 165
Leu Cys Phe Cys Gly Ser Asn Glu Ile Asn His Phe Phe Cys Asp
170 175 180
Ile Pro Pro Leu Leu Leu Leu Ser Cys Ser Asp Thr Gln Val Asn
185 190 195
Glu Leu Val Leu Phe Thr Val Phe Gly Phe Ile Glu Leu Ser Thr
200 205 210
Ile Ser Gly Val Phe Ile Ser Tyr Cys Tyr Ile Ile Leu Ser Val
215 220 225
Leu Glu Ile His Ser Ala Glu Gly Arg Phe Lys Ala Leu Ser Thr
230 235 240
Cys Thr Ser His Leu Ser Ala Val Ala Ile Phe Gln Gly Thr Leu
245 250 255
Leu Phe Met Tyr Phe Arg Pro Ser Ser Ser Tyr Ser Leu Asp Gln
260 265 270
Asp Lys Met Thr Ser Leu Phe Tyr Thr Leu Val Val Pro Met Leu
275 280 285
Asn Pro Leu Ile Tyr Ser Leu Arg Asn Lys Asp Val Lys Glu Ala
290 295 300
Leu Lys Lys Leu Lys Asn Glu Ile Leu Phe
305 310

<210> 12
<211> 316
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature

WO 02/057454

PCT/US02/01339

<223> Incyte ID No: 7472099CD1

<400> 12

```

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

```

<210> 13

<211> 318

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7405443CD1

<400> 13

```

Met Glu Trp Glu Asn His Thr Ile Leu Val Glu Phe Phe Leu Lys
1 5 10 15
Gly Leu Ser Gly His Pro Arg Leu Glu Leu Leu Phe Phe Val Leu
20 25 30

```

WO 02/057454

PCT/US02/01339

Ile Phe Ile Met Tyr Val Val Ile Leu Leu Gly Asn Gly Thr Leu
 35 40
 Ile Leu Ile Ser Ile Leu Asp Pro His Leu His Thr Pro Met Tyr
 50 55
 Phe Phe Leu Gly Asn Leu Ser Phe Leu Asp Ile Cys Tyr Thr Thr
 60 70
 Thr Ser Ile Pro Ser Thr Leu Val Ser Phe Leu Ser Glu Arg Lys
 75 85
 Thr Ile Ser Phe Ser Gly Cys Ala Val Gln Met Phe Leu Gly Leu
 90 100
 Ala Met Gly Thr Thr Glu Cys Val Leu Leu Gly Met Met Ala Phe
 105 115
 Asp Arg Tyr Val Ala Ile Cys Asn Pro Leu Arg Tyr Pro Ile Ile
 120 130
 Met Ser Lys Asn Ala Tyr Val Pro Met Ala Val Gly Ser Trp Phe
 135 145
 Ala Gly Ile Val Asn Ser Ala Val Gln Thr Thr Phe Val Val Gln
 150 160
 Leu Pro Phe Cys Arg Lys Asn Val Ile Asn His Phe Ser Cys Glu
 165 175
 Ile Leu Ala Val Met Lys Leu Ala Cys Ala Asp Ile Ser Gly Asn
 180 190
 Glu Phe Leu Met Leu Val Ala Thr Ile Leu Phe Thr Leu Met Pro
 195 205
 Leu Leu Leu Ile Val Ile Ser Tyr Ser Leu Ile Ile Ser Ser Ile
 210 220
 Leu Lys Ile His Ser Ser Glu Gly Arg Ser Lys Ala Phe Ser Thr
 225 235
 Cys Ser Ala His Leu Thr Val Val Ile Ile Phe Tyr Gly Thr Ile
 240 250
 Leu Phe Met Tyr Met Lys Pro Lys Ser Lys Glu Thr Leu Asn Ser
 255 265
 Asp Asp Leu Asp Ala Thr Asp Lys Ile Ile Ser Met Phe Tyr Gly
 270 280
 Val Met Thr Pro Met Met Asn Pro Leu Ile Tyr Ser Leu Arg Asn
 285 295
 Lys Asp Val Lys Glu Ala Val Lys His Leu Pro Asn Arg Arg Phe
 300 310
 Phe Ser Lys 315

<210> 14

<211> 321

<212> FRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3090414CD1

<400> 14

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

WO 02/057454

PCT/US02/01339

Leu Lys Asp Ile Gly Phe Asp Glu Cys Leu Val Gln Met Phe Phe
 95 100 105
 Ile His Thr Phe Thr Gly Met Glu Ser Gly Val Leu Met Leu Met
 110 115 120
 Ala Leu Asp Arg Tyr Val Ala Ile Cys Tyr Pro Leu Arg Tyr Ser
 125 130 135
 Thr Ile Leu Thr Asn Pro Val Ile Ala Lys Val Gly Thr Ala Thr
 140 145 150
 Phe Leu Arg Gly Val Leu Leu Ile Ile Pro Phe Thr Phe Leu Thr
 155 160 165
 Lys Arg Leu Pro Tyr Cys Arg Gly Asn Ile Leu Pro His Thr Tyr
 170 175 180
 Cys Asp His Met Ser Val Ala Lys Leu Ser Cys Gly Asn Val Lys
 185 190 195
 Val Asn Ala Ile Tyr Gly Leu Met Val Ala Leu Leu Ile Trp Gly
 200 205 210
 Phe Asp Ile Leu Cys Ile Thr Ile Ser Tyr Thr Met Ile Leu Arg
 215 220 225
 Ala Val Val Ser Leu Ser Ser Ala Asp Ala Arg Gln Lys Ala Phe
 230 235 240
 Asn Thr Cys Thr Ala His Ile Cys Ala Ile Val Phe Ser Tyr Thr
 245 250 255
 Pro Ala Phe Phe Ser Phe Phe Ser His Arg Phe Gly Glu His Ile
 260 265 270
 Ile Pro Pro Ser Cys His Ile Ile Val Ala Asn Ile Tyr Leu Leu
 275 280 285
 Leu Pro Pro Thr Met Asn Pro Ile Val Tyr Gly Val Lys Thr Lys
 290 295 300
 Gln Ile Arg Asp Cys Val Ile Arg Ile Leu Ser Gly Ser Lys Asp
 305 310 315
 Thr Lys Ser Tyr Ser Met
 320

<210> 15
 <211> 422
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <223> Incyte ID No: 7503710CD1

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

WO 02/057454

PCT/US02/01339

Val Ser Ser Arg Gly Thr Pro Ile Leu Ile Leu Lys Arg Cys Ser
 155 160 165
 Thr Leu Leu Gly His Ile Ser Leu Phe Ser Gly Leu Leu Pro Thr
 170 175 180
 Pro Leu Phe Gly Val Val Glu Gln Met Leu Phe Lys Val Leu Pro
 185 190 195
 Gly Leu Leu Cys Pro Val Val Asp Ser Val Leu Gly Val Val Asn
 200 205 210
 Glu Leu Leu Gly Ala Val Leu Gly Leu Val Ser Leu Gly Ala Leu
 215 220 225
 Gly Ser Val Glu Phe Ser Leu Ala Thr Leu Pro Leu Ile Ser Asn
 230 235 240
 Gln Tyr Ile Glu Leu Asp Ile Asn Pro Ile Val Lys Ser Val Ala
 245 250 255
 Gly Asp Ile Ile Asp Phe Pro Lys Ser Arg Ala Pro Ala Lys Val
 260 265 270
 Pro Pro Lys Lys Asp His Thr Ser Gln Val Met Val Pro Leu Tyr
 275 280 285
 Leu Phe Asn Thr Thr Phe Gly Leu Leu Gln Thr Asn Gly Ala Leu
 290 295 300
 Asp Met Asp Ile Thr Pro Glu Leu Val Pro Ser Asp Val Pro Leu
 305 310 315
 Thr Thr Thr Asp Leu Ala Ala Leu Leu Pro Glu Val Met Thr Val
 320 325 330
 Arg Ala Gln Leu Ala Pro Ser Ala Thr Lys Leu His Ile Ser Leu
 335 340 345
 Ser Leu Glu Arg Leu Ser Val Lys Val Ala Ser Ser Phe Thr His
 350 355 360
 Ala Phe Asp Gly Ser Arg Leu Glu Glu Trp Leu Ser His Val Val
 365 370 375
 Gly Ala Val Tyr Ala Pro Lys Leu Asn Val Ala Leu Asp Val Gly
 380 385 390
 Ile Pro Leu Pro Lys Val Leu Asn Ile Asn Phe Ser Asn Ser Val
 395 400 405
 Leu Glu Ile Val Glu Asn Ala Val Ala Ala Leu Tyr Val Leu Val
 410 415 420
 Val Ala

<210> 15
 <211> 2192
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 71924779CB1

<400> 15
 aggtggcgcg taccaccaga gaaccagcga gtcgccagct gccctggcc tggcgggggc 60
 ggaaccgcgc gggatcccca cccccaccg gaatcctcgc cacggagaat ccttggagaa 120
 gccccggatc ccggctggg agggaggagt gctcgttgac ccccgaccgc gcgctgatcc 180
 gcccccggc ctccggactt ggggagccgc tgtactctgc ctccggacc acgagactct 240
 agacgggagt cccctcgagg tgaagccgct gagttcccg gccccggcag gcttccctgg 300
 gagagccgac ggaaccccc tcccagcaca caaaccttc ctgctttca ccgggactgg 360
 cggagccggc ggcggactta gacggggga cttcagggca gggggcgccc cctgccccgg 420
 tcaccagtcg gggcgagggg acgtctctc tcccagact gctctgctcg gatggcgccg 480
 ccggctgagt gacggggcgg gggcgagga ctcccagct cggactctt gccttogagg 540
 ggaagatgt accagagtc agaagtggg ggtcccacc ccaatcact cctagtggta 600
 gattttata accagaccg ggcctgttg ctcccagaga aggggtccc cggccgggt 660
 ccgtactca ccccgctccg gactccgct tggatggct caaacctc ccttgagacc 720
 nagagcgca gttctgaaga gatagtccc agcctccct cggcaccnc tctaccgcc 780
 atotacaage cttgatttgt ctgtcaggac aagtcctcag gctaccaeta tggggtcagc 840

WO 02/057454

PCT/US02/01339

```

gcctgtgagg gctgcaagg cttcttccgc cgcagcatcc agaagaacat ggtgtacacg 900
tgtcaaccggg acaagaactg caticatcaac aaggtgaccc ggaaccgctg ccagtlactgc 960
cgactgcaga agtgccttga agtgggcatg tccaaggagt ctgtgagaaa cgaccgaaac 1020
aagaagaaga agggagtgcc caagcccagag tgcctlgaga gctacacgtt gaccgcccag 1080
gtggggagca tcaatgagaa ggtgagcaaa gngnacagg aaacctccc gccccttgcg 1140
cagctgggca aafacactac gaacaaagc tcaaacacac gttctctctt ggaatlgac 1200
ctctgggaca agttcagtga actctccacc aagtgatca ttaagactgt ggaattgcg 1260
aagcagctgc ccgcttlcac caacctcacc atcggccacc agatcaccct cctcaagctt 1320
gccctgctgg acatcctgat cctggggatc tgcacgcygt acaogccoga gcaggacacc 1380
atgacctctt cggaggggtg gacctcgaac cggaccocaga tgcacaaccg tggcttcggc 1440
ccoctcaocg acctggtctt tgccttgccc aaccagctgc tgcctcggga gatggatgat 1500
ggggagacgg ggcctctcag gcgcatctgc ctcatctgcy gagaccocca ggaactggag 1560
cagccggaac ggggtggacat gctgcaaggag cgcctgctgg agggcctaaa ggtctacgtg 1620
cggaacggga gggcccagcc ccccccaatg tcccccaaga tgcataatgaa gattactgac 1680
ctgcaagaca tcagcgccaa gggggctgag cgggtgatca cgcctaagat ggagatcccg 1740
ggctcaatgc cgcctctcat ccaggaaatg ttggagaact cagagggccc ggaactctg 1800
agggacacag cgggggggtg ggggggggac ggggggtggc tggccccccc gccaggcagc 1860
tgtagcccca gctcagccc cagctccaac agaagcagcc cggccaccca cccccctga 1920
ccgcccacac caatgagaa cagcctccgc cctccgccc ggtttttctc tgccttttac 1980
cagacattgt gaccocgac cagcctcggc cccactgccc cggggcagta ctggcgactt 2040
cccggggac gggggggag gagaagatct tggacagagg ctggcctcag ggatgcctgt 2100
cccagctggt gaataaagc aggcgaagat gagccggcg gttcaaaagt cgggocgttc 2160
aaacctgccc ccatttaaga aggggcccac aa 2192

```

<210> 17

<211> 3614

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2319430CB1

<400> 17

```

cgttgagggc aaccagagg ccattgccc ctcctccct ctatcccct cggctgtta 60
agggctcggc ccagagcgc tctgcccgc gacgatggt accgtacgg cgggccact 120
cgcgctgccc ccgctcccc agagagccac atccgaggt cggcgagaa gagcccccg 180
tgtgaccgtg ccgtaccggc ccctgccc cgcggagga gaacgggag ggggcccaga 240
gagccgggga gtlggggag ccgcccccc gcagcggcc tccccggga gggagtcct 300
cagcctgagg tctctccca gaaaaaaaa aagaaaaaa aaaacaact ggtgcaasg 360
gagaaactgg aggcagtggt aaatgtggc ctgaggggt caagcatcat gctgttggat 420
gtcctgtaca gatgggatgt cagctccttt tccagcaga tccaaagaag tagccttagt 480
aataaccctc tttccagta taagtattg gctttaata tgcattatg aggttatata 540
tlaagtggg tgcctcaac attgcccag cagcalctgg tccagctta tctalatttt 600
ttgactgctc tgcctccta tgcctgacat caatlttca gggactatgt tccgagtag 660
ctggagtttg cctatgaggg acaatgtat ttagaacctc tctctatgaa tgggtttacc 720
acagccttaa tagttagtt ggtgtgtgt actltatgt cctgtgtcat gaaacaaga 780
cagatttggc tgtttcagc tcaatgctt cctctgtag cagcactctg cctgttctct 840
ttggagacaa ttgttatcat caataaatt gctatgatt ttaactggat ggaagtctc 900
tattttctgt ggtcctaatc ttggtaacct tataacctg ctaaatctgc atacagaga 960
ttggttcagg tagtggaggt atatggcctt ctcgcccgg gaattgccc gtggaaatca 1020
ctggtagctc ctgtctctt catggtttc tggctctct tattgtctc tcaagtttac 1080
tctctattca gtaactgaga tcaactgca tcaactgaga ggtctctctt ccttttctgt 1140
acaagtattg cggaaatgct cagcaactct taactcttt tgggtttgg cttaocgggt 1200
tctttgtgt ccttgggtgt tctcaactc tgaagtttt actgcaagg ttaactgagc 1260
tcatgaatg atcctgcat gaatcggggc atgacagaag gagtaacgt gttaatcctg 1320
gcagtgacga ctggcctgat agaactgag gttgttcatc gggcattctt gctcagttat 1380
atccttttca ttgtctgag tcttatccta cagttatgt tagaacttgc agatctctat 1440
gthttggccc tgggagcaac tagagacaag agcttgtgga aacacttccg tgcctgaagc 1500
ctttgtttat ttttattggt atctcctgtc tatabgctt atatgattg ccagtttttc 1560
caatggattc ttggcttctc tateattatt tccagcagca ttctacote tcttcaggtt 1620
ctgggaacac tttttattha tgtcttatt atggttgag aatcagaaa agagccagtg 1680
gaaaacatgg atgatgcat ctactatgtg aatggcactt accgctgctt ggaatttctt 1740

```

WO 02/057454

PCT/US02/01339

```

gtggccctct gtgtggtggc ctatgggctc tcagagacca tctttggaga atggacagtg 1800
atgggctcaa tgatcaactt cattcaattc tactataacg tgtggetteg ggcccagctg 1860
gggtggaaga gctttctctc ccgcagggat gctgtgaata agattaaabc gttaccocatt 1920
gctacgaaag agcagcttga gaaacacaat gatalttttg ccactctgta ccaggacatg 1980
aaatctgctg tgaatcagcc ttgcagcaat ttttccatg caggetgtct taagaatcgg 2040
ctgtatgtcc agggagcttg gctctgtgct caactgocac tgaanaatbc ctcccagctt 2100
ccagattag gaactgagcc agttctacag cctcatgctg gagctgagca aaactcctatg 2160
tttcaggaaag gtaactgaacc cccagccocag gacalactc caggaccag gatcacggaa 2220
ggltccaggg acaataatga gtacattgcc agacaccag ataaccagga aggggctttt 2280
gaccoccaaag aatcactca cagtgcgaaa gatgaagcac atcctgttga atcagcctag 2340
aggagaagca gcaggaatga tgccttgata ctctggagga gaagttaact caagatggaa 2400
tctatgttct gatttgagga atgaaaatga gatgacag caggaaaactg acattccaag 2460
gatctaactc aggaagtact ctcaagtggg accacctgct ttcacccctt gacattgtgg 2520
gagaaatctt gcaatgtatc tcaatcaaaa tgtatttata tgttctctgc tgatgtttta 2580
tagaggtttg tgaagaaaaa tcaacctcag caacttcaga aactgcccct gatacgtgtg 2640
agagagaaat aaaaatcagat tttgagtgtt gaagggactg aggaagtggag gataaagagc 2700
atagagacag catggaaga agggagccaga agtggaaactg aacttcaact ctccatggga 2760
cagatcaatc tcaatataca gctctgaatag caaccagccc tctcctccac cccgtttctc 2820
ctcagttaat tggagctcag tcaagtgatt attgagctt gtaacagcact gaaatgaatc 2880
caaatgaa caagcatga tagattcaaa agattgaagc ccgctcaaac tttgatgtgt 2940
ctttaggaaa ggggtgggtg ggcactgggg ccttgccggtt gcattoactg aatctgagac 3000
tctttcaact tatgacggag tcttcaatal tttgatgtat atgaaacttt tgttaaatc 3060
gttgataact tctgtgctg tgtgaagtaa actaaactc tgaatgacac tttggagctt 3120
gcttttagtga aggagaccaa agtgggaagg gctttagggc actgatagag gccctgggtg 3180
tacttttcaa tctgtgtaa ttttaaatc ttgcaactga atcaaacag ttttaaatc 3240
tggcaatatt tgcactttgg gaatgagtac ataactgtat gatcacactc tgcacaaagc 3300
acttttaaa cgttbaatag actttgcaac tttcttttga caaggaatgt tcaatattaa 3360
atctttacat tcaatcagtc tacaggtaga actggggagg ggggaatgta attttttatg 3420
ggaaatltga tatgaaaga aactagtcac ttatttatac aataggcttg gctcaaaaag 3480
tgtttttcag acctgggtat tctcaatgtg ggatgtgact ttattttatt tttagtagca 3540
aaatggatg tagactgaca gacatagctg aatgtcttaa taaattttaa tttgaagata 3600
aaaaaaaaaaaa aaaa 3614

```

```

<210> 18
<211> 1585
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 7291877CB1

```

```

<400> 18
cgcgacccaa caacagcaac aactccactg cgccgggctg agggagcagga attagagct 60
cggaataat atgaaggga tccgcaagg gaaagcga gcaaaagga ccaaacctc 120
ggagcctggc aagcgaagat gcgctaaatg tggccgcta gacttcaatc tgatgaaga 180
aaagggaatt aaagttgata ttacgttttg gaactctgct tttttatga cggltctctg 240
tgtgaaagga tttattata catgtggttg aactttaa ggaactaatg gcactataga 300
aagccctggg tttccatag gatatacaaa tgggtcaaac tgcacatggg taataataga 360
agaagaacga aatgaaatac aaattgtttt tcagtcattt gctctagaag aagaatacga 420
ctactatca ttaatgatg gacatcctca tctcaaac tttaggacaa ggttaacagg 480
atccatctg ccacctocag tgacaagtac caaatctgt ttctcaactc gtttgaccag 540
tgattttgca gttgatgctc atggatataa ggtatatac gaagaattgc agatgagctc 600
tttggaaat cctggtgttc cacccagaag tgtattata ggcacaagat tgaagctgg 660
ggacaagat cgcacagct gtgtaactgt atacatcctt gatggccacc ctcaagctac 720
ctgcatagcc aattcagtta atacagcttc gttggatttt cctgttccta tctgtagagc 780
tgaagatgct tgtgagggaa caatgagagg atccagtgcc atcatalcca gccctagttt 840
tctcaatgag tcaataaaga atgtcagtg caactggacc atgtgagcag agcctggagg 900
cacnattca ctcatatta ctgatattca aatggagag aatatgat tcttgaat 960
agaaggtctt gggccactca ccatatggtt atctggaaat aatataccac caccaatlat 1020
cagcaacaaa aactggctca gactgcattt tgttacagac agcaatcac gataccctg 1080
atltagctct ccatacaag gttctctcac attgacccac actacatcca ctggtgagtt 1140
agagagcat aacaggacta ccactggtgc tattgtctgt gctagcaaac ctgcaagatt 1200

```

WO 02/057454

PCT/US02/01339

```

tactgtatcc agtgttacag ctgtccaccat ccatagactt tccgaggaac agcagtgca 1260
agttacagat ctacagaaat caggtctgga cccaacacg tccaaggaag ggclctctcc 1320
tcataccagca gatcacaaaa gtaccaggag aagaccnaga catgctgaac agatagaag 1380
aaactaaagag cttagcagttg ttactcatag aggcattgca aatagagtcg agnacataga 1440
aaaaccattt cttagtggtac aagatagatt ctgtaaaaag aattctgata aaagtactaa 1500
agaagtacaa gtgtgtatgc agagagtgag tcttttaagt tactttttca atgagttggt 1560
aaacaaccga aaaccaatbg cttaa

```

<210> 19

<211> 5618

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1218126CB1

<400> 19

```

gatatgactg agggcccatg gggggggggg gggggggctg agggagcagg gectagcaag 60
cgccacagct agcgaccctt gcttggcctt ggctagactg gccctacgdc ttttccagaa 120
gctcttcagc aagaggaacg ggctggggcg gcccgccgcg gacgcccggg acccagattg 180
cgggttcagt tggcctttac cagagtttga tccaagccag attcgactga ttgtatatca 240
agactgtgaa agacagggga gaaatgtttt gtttgactcc agtgttaaga gaagaaatga 300
ggacataatca gtatcgaaac tctgcagtga tcttcaagtt aaagtctttg ggaaatgctg 360
ccaactgaaa cctggaggag acagttcttc ctctttagat agttctgtga ctctactctc 420
tgatataaaa gaccagtgct ttaagtacca gggttctcgg tgcctctctg atgccaatat 480
gcttggagag atgatgtttg gctcagtagc aatgagctac aaaggaacca ccttaaaaa 540
tcatacagatt cgttcccctc cacagctcat gctcagcaaa gtgtttactg ctccggactg 600
caacagattt tgtggggagt tcaatacgtc acaagatagt ctgtgaatca tcaatcagga 660
caacaatata taaaggctg ataataacac agttattaat ggactgctg gaaatatagc 720
atctctcagc agettgtgca tcactccatt tcttcccca aactcctcac ttaccggag 780
ttgtgccagc agctaccagc gacgttggcg acgcaacca acaacaagtt tggaaatgg 840
ggatattcct agatggcta taaagaagaa cttaaatctc tccagtgaaa gcttggccc 900
taaccocgga attgctggga aaagaagat tccaattggg gtaatctttt cttgtccaa 960
agatgaagat gaaataaca aatlaatga attctttttt tcaattttc ctctctttga 1020
aagctacatg acaaatata agagtgcaat agaacagct atgaaaatga gccggagatc 1080
agctgatgcc agtcagagaa gtttggcata taatgaata gttgatgcc taaatgaatt 1140
cagaacaaca atttgtaac ttaccagat gccacgaatt ggagaacctg tctggcttac 1200
aatgatgtcg gggactccag aaagaacca ctttctctat cgtttcatga aggagttcac 1260
ctttctaagt gaaatgctt ccaaaaatca attcttgcca gctctcatca ctgcagttct 1320
gaccaatcat cttgctctgg ttccaacagt catgccaat ggacaaccac atataaaaa 1380
atlltagaaa aagcattctc ctacagatgt ggacalgttg gcaaaagctc atccatata 1440
cccactttgg gcacaactgg gagacttgta tggcgctatt ggcctcccc lacggttac 1500
aaggactgtg gtagttggca aacgacnaga catggtccag aggetacttt attttttac 1560
ttatltata agatcctcag aactcaaga aacgactct ttgaaaaatg ggaagatga 1620
agcaatcgtt atccagga cagtaattac taccacttta gagaagtg agaataga 1680
atcagagtat gctctgtca caatgcata aacaanaagc agtttctct taaagagtc 1740
agagaatatt agaactccc attgtaactg taatatgca agtcaaccac tcttgggca 1800
aatgtagagc aacattcac aacagagagc agagatatt caaaaagct ctaaggagct 1860
gttaggaatt tcaatgagc gccgatgat ttctcctct gactgccag aagaataagc 1920
tgttgatggt aaacagta gagataaatt aagaacttgc tttgaccca agtttagagc 1980
agttgtttgc acaggatctg ttccagtaga caaatgtgca ttgtcagagt caggcttaga 2040
gtcaacagag gaaactggc agagttagaa gtgtcggat tcagacagtc acacaggcaa 2100
agcaatgaga tccacaggaa tggttgtgga aaaaaaacct ccagataaga ttgtgctgc 2160
ttcattttct tgtgaggctc ccagacaaa ggttactctc ctgattgggg atctatgct 2220
acctgattca gatactgagc ttcaagatca ggcagtggcg gatcagatta ccagacatca 2280
caccaacca ttgaaggaag aaagaggggc tattgatcag catcaagaaa ctaaaacca 2340
aaccagagc caatctggag agtctgatac acagaaatg gtttctgaa agcctctga 2400
acttccctgt tggatcatt cagaccaga aagatgagc ttattcagc aatattttac 2460
tgatgattca atcgaacca ggactattga tgatgttcaa tttaaacaa gtacagtag 2520
taagacctat tgcctgatgt tagagtttcc aaaaatag tgtaaaaaa abaacagca 2580
gaacaatgaa ttttctaatt gtatgaaac agttcccaca gattcatgta aaacctgct 2640
tctctcagc gaccaagag atacactctc cattctgtc cccatggg ataagaagag 2700

```

WO 02/057454

PCT/US02/01339

```

ttcagataaa aaaattgctg taggaaactga atgggacatt ccaagaaatg aaagttcaga 2760
cagtgccott ggggatagtg aaagtgaaga tacaggtcat gatatgacta gacaagttag 2820
cagttattat ggggagagagc aagaagatfg ggcagaagag gatgagatag cttttccctg 2880
gtcaaaagta atcgaagtga gtgctgttca gcccaacatl gcccaactcg ggaggctcct 2940
gctgggtggc tactgctgat ctatgtggcc tgactttgtt ctccaaggaa ttgggagtga 3000
tgaagagttc cgtcagtgtc tgatgtcaaa ttatctocat gctgtgcagc atccagtttt 3060
ggatgaacca atggcagaag ctgtctgtat tatagctgac atggataaat ggcctgttca 3120
agtgccagct agccagagac gagtgcagca taataaattg ggaaggaagc tattgttttc 3180
cagttcttgt tccaatctgc ttcaatccac aottcagott tataagcata actgtctccc 3240
aaatltttgt gtaatgcatc ttgaagaccg gttgcaaggag ctatactcca aaagtataat 3300
gctgtctgaa taccgtgagg ggcagatgct gttcaatgct aaggagctgg gagtggtttc 3360
ggggattgaa tccagtgatc ttccactctc ggctgtgtga gcaagcactc actctccata 3420
tgttgcaaca atactccttt aatataccta aaaattgtta gaaattgggt ggaaaatagc 3480
tagaaaacca ggaagcagac acaacaatga ttatagaga ttctttttcc cttttagact 3540
tccatctgaa tgaatcagtc accaggggat tctgcatagc atgtatattt ctgtgtatgt 3600
cagatggctt tttctttttg actggacttt tgggtgtggg tagattttta aacaaatgaa 3660
atlaaagcaa caataaatttt gaagcatttg aaaaagocaa agtgtacggg agaaatttct 3720
acaaaatgaa tatbatcaag agttctcatg gatacctgca gctgtgtcac agctcaataa 3780
tagcaacagt gttctcagat ttaattggctc agaaatagtt atccatlagt tttaaatlct 3840
taatttctaa ggtcacagaga tctataaacac cttagattat tgttagtttt gcaattcaaa 3900
acagcaaatg tctgtttatt tctcaaatga agtattttaa acagcctgtt aattataaga 3960
aacctagaat atgagtgta aatgtgttat gttatccacc caagtgtaca tatgtacota 4020
ttttttttta aaaagcagaa atagaataac agactgggta aacatgctct taanaatata 4080
tatatttcca actagtatg tctataatgc tgaatatta cttatgggtg atttttctgt 4140
ttcaacacct ctaaaatata agtaaaagcca accctttttt taaggctgag attcccaaaa 4200
tgagaatact actttatacc atttgtttat aagtatgaac tgttcttata aatattaata 4260
tttacaatatt caactattta acataaatga aaattaggat taanaattgc accaaagcat 4320
cggcaaaaaa aatacatata tctttaaaag tgcctaggta gccaaaggcc tgccttttgg 4380
taccacacct catgaaccca taggagctga tattgttttc actgcttaat aatcccaaat 4440
ttacactatt cataactctt aaaaatttct tctttttttc taagagoccc tcccttccaa 4500
aagtgatttt ttttcaaaag ttttccactc tcaattgttg cctttgtaca taactatagag 4560
tgttgcctgt aagaagagct aatataagac caaactcttg taagtaagtt aactagaagc 4620
gtgggtgatg aagtttttca atacttctca ctactcagc ttaactttagt actactaat 4680
agtttattct ttgcttatct ggtctaaagc acttttaagc ctagtatgaa agtgggttcc 4740
tgccttcatt gactatttct atcabaattt catcattgat taanaaaga aaaccacttt 4800
tttattcagt tattaaatat atttactata taacaatcc attcttgctg tttaaatlct 4860
caatagttaa tggaaagtgt tctttgacct tgaatttaca gcaattgggt acattttgcc 4920
ttgctgtgta tgtatccaag agacttccaa ctagacaaag aaaaatttgt tgttttaatg 4980
gaalgtaaac ctgaaattgg tgtgtctgca actgttttgg cccatgacct ttacactagt 5040
cccagttatt acctgagctc cccatggagc actgtgtgcc aaggaggttt tgtggatata 5100
tttcttttgg cttaatttcc ttattctgtg cattaacaaa attatccagt tgcctgattt 5160
tggaaattca tgaatcaato tttttggcag aattoagaat attaaaagt tcaatacatt 5220
gcccggcatt gtaccctttt tttttttttt ttttttttga cggagtttca ctctgtttac 5280
cccagcttga gtgcaatgg tgcgatctca gctcactgca acctccgctc cccagttcaa 5340
gtgalltccc gtcctcagc ccccaagtta gctgggatta caatttgggc gcccaacacc 5400
cccagctaat ttttattttt agtagagatg agtttcaacc atgttttggc caggtctgct 5460
ttgaactcct gaactcaagt gatccaccgt gctctgaact cccaagttac tgggattaca 5520
ggcgtgagcc atgtgtgcc agcctgtgac tttttttttt tttaattgta gctctgttta 5580
gcattggggc atatgtcggg gtgtctcttt aacottaa 5618

```

<210> 20

<211> 1641

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7479161CB1

<400> 20

```

ccgacaggag gccgggagcc cccaactacc ccttctggag ctgcaggagc aaggcactgc 60
agccagctcat gctgcccctg tggctccctg ttctgctctg gggcctggcg actccatgcc 120
aggagctgct agagacggtg ggcacgctcg ctcgatttga caaggatgaa ctcgcaaaag 180

```

WO 02/05744

PCT/US02/01339

```

ccatccagaa ctaactgggt ggggagccca ttctgcagaa tggctgggga tgggtcacag 240
ctgtgaaocg gggocctctg ggctcaggag ggctgcttgg aggaggcggc ttgctgggoc 300
acggaggggt ttttggcggt gtgcaggagc tctctggtct gaagattggag gagctcacgc 360
tgccaaagggt gttgctgaag ctgctgcccg gatttggggt gcaactgagc ctgcaaccca 420
aagtgggcat gcattgctct gggccccttg gtggccttct gcagctggct ccggaggtga 480
acgtgacatc ggggtggcg ctggscgtga gctcaagggg caaccaccac ctatctcca 540
agggctgagc caagctctcg ggcacatca gctgttctc agggctgctg ccaacacac 600
tcttgggggt cgtggaacag atgctctca agtactctcc gggactgctg tgcccctgg 660
tggacagtgct gctgggtgtg gtgaatgagc tctgggggc tgtctgggc ctgggtccc 720
ltgggctctc tgggtccggt gaattctctc tggccacatt gctctctc tcacaaccgt 780
acatagaact ggacatcaac cctatcgtga agagtgtagc tgggatatac attgacttcc 840
ccaagctccg tgcccagccg aaggtgccc ccaagaaggga ccaacacatcc caggtgatgg 900
lgccactgta cctctcaac accaogtttg gactcctgca gaccaacggc gccctgaca 960
tggacatcac cctgagctg gtccccagcg atgtcccact gacaactaca gaactggcag 1020
ctttgctccc tgagcccctg gggagactgc cctgcacca gcaactccta ctgttctgc 1080
gggtgaggga agctcccacg gtcacactcc acaacaagaa ggccttgctc tccctccagc 1140
ccaacatcca tgtgctgttc latgtcccta aggggaaccc tgaatccctc tttgactga 1200
actccgtcat gactgtgctg gccagctgg ctcctctggc taccagctg cacatctccc 1260
tgtccctgga accgtcagc gtcaggctgg cctccctctc taccatgctc ttgacggat 1320
cgcttttaga agaatgctc agcaatctgg tggggcagc gctatgcacca aagcttacc 1380
tggccctgga tgttggaaat cccctgctca agttcttaa tacaatattt tcaatctag 1440
ttctggagat gctagagaat gctgtggcgc ctctctatgt cctttagta gcatagaaga 1500
tgggtgtctt ctacagtagc tggactatgc calgtlatit tgttcttga ctaaggccct 1560
gtgaggtgca actggtccc tttcattttt ggtcagagat ggagataag gaatttatg 1620
ttggtactag cactgggata g 1641

```

```

<210> 21
<211> 6056
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7722591CB1

```

```

<400> 21
gactgacctg tgaggactgc ctggccaact ctagecagtg cgcctgggtc cagtccacc 60
aacactgctt cctgtttgct gctactgttg cccggtacc ccaagggggc tgtcagagct 120
ggagacagag gtatggtccc tggggcaggg ctaacagagc aagattcccc accggcaagg 180
ggctggggct ctgaccacca cccctgccat cctgcagtg acactcggag ccaaggtgcc 240
ggagctgoga tggcttctcg acctgcccag agtctctgca gagccacgag tgggctggt 300
gtggcaatga ggacaacccc acactgggac ggtgagcccg ggcaggtggg tgggcaaggt 360
gcccgctgtg gtccttctcc catgaccggt cactctaagc gccctcttgc tctctgccc 420
tcttgcccac cccatggcca ccttcccctt tccgtactgt tttctgttg ttatltttac 480
cctcccgctc ttttctccc attttctctc ctttccggc tctctgggat tctgtttctt 540
tctcctgttc tgttctgttc tctccgctcc cctctcaact gcaactctgt ctatgtctct 600
ctctgtcttt ccaaaattgt tttgtctgc gacttctccc ggtttctctg tctctcttcc 660
caaatitgta tttgcatccc tctctctcca attggccccc tctctcccgt tcaatgttcc 720
tatgtatggc tcaatgttct atgttctccc ccgcttcttc actctccacc cctgcaaggtg 780
cctacagggg gaettctcag ggcccctcgg tgggggtaac tgcctccctg ggtgtggggg 840
agggcctggg tgcctcccgt ggcccctgcc tcccctctg ggcataccgc cgtctgctcg 900
acgtggatga gtgtgcctg ggccctggccc ggtgccacc gggggcgacc tgcctgaaca 960
cgcccctcag ctacagatgt cactgcaccg ggggtaccac ggggtgatggc atctcaact 1020
gcaaccgcac tgacttggag gactgtggcc atgggtgtgt cagtgggccc cgggaactta 1080
cctgcgtgtg tgacctagc tggacatcag acctgcccgc lcccacacct gccccgggtc 1140
cgccagcccc ccgctgctcc ogggactgtg gctgcagctt ccaacagccc tgcggcaagg 1200
ggggccctgg cttctgggac gactgcccag actgacatg gggggagcac tgcgaacgat 1260
gcggcccggy cagctctggc aacggccacg gctctagggg ctgcccggcc tgcgaatgca 1320
acggacagcg ggcccacgc cgtggccact gcgacaacc cagtgggctc tcttcttccc 1380
aggaacacac cggaggtgoc cactgcaccg tctgctccc aggtatctat ggggatccc 1440
ggcccggtgg tcccctctt ogggaggtg ggggtggccc cctcctccc aactgtctc 1500
nagtggcact gggctcacgc cgggtcgggg ggtctgccc tccaggtggc gggcctgca 1560
gagccgggoc tggcctgtcc tactgtgtgt ggttgtctc ggcactgag gagctacagc 1620

```

WO 02/05744

PCT/US02/01339

```

cctgtgctcc cgggaccctc tgtccccac tcaccctcac cttctcccc gacagcagca 1680
ccccctgca cctgagctac gtcctggcgt ttgatggatt cccacgcttc ctggcaactg 1740
gtgtlgtcca gtcggaccgc agcctcatag ctgctctctg cggccagcga cgggacagcc 1800
ccctcaactg tcaggccctg tctgggactg tctgtctgca cggggaggcc aatggctcct 1860
catctgggg cttcaatgt tgggtgggct ctgcccctg agagcgtgc tgggtcaggg ggcgccggga 1920
gctgtccctc ccccacgaaa tgggtgccc agacagctg tgaaggtgc gggctctgct 1980
gatgtctca gggctgggct gcccacact gcccatgac tetgttctc gagaactgca 2040
atgcccacac tgggcagga acttgaacc agagctcgg tgtgtgalc lgtgcccagg 2100
gcttcgggg ccccgaactg gccaccaagc tggatggcgg ccaactlgtc tgggagacc 2160
tcattggacg cgcctctca gccgacactg ccagcgcctt cctgcaaccg ctgggocaca 2220
ccatggtgga tggaccgcat gccacottgt ggalgtlgtg gggcctgggg ctgcccocag 2280
ggctgtctgg aaactgtac agtactacg tgaagtacg ggggtggaca cagatctgg 2340
cgggagcga ggaacggggc ccaggcccat cggcccctc ctccatgca gctgcatatg 2400
tgcccctgg cgtgtgtcc atgtatctg tgggggact taccctgga ggcgtcacc 2460
gtgatttct ggtctcaac ctacaccctc tgaatggcg gcaaggaga gcccocaga 2520
ccgtggagct gccagcctt cctgttcaac ccttactgc ccgcccgggc ctgtctctgc 2580
tctgtgtgg cgttactcc ccgaaaaatg gcttcaacca gcaactgctg gactcaccg 2640
tggcaaccg cacclyggg tcaggagccc agagtggac acccccaca ggtctctatg 2700
gtcaactgt tgtctaacac gaagccacg actccctca cgtgttggg ggtttccgat 2760
tcaatgtgga gctgggggc ccatccccg agctctact cctgactgt cctgacccga 2820
cctggagtct gctggccctc tctcagggg caaagccccc ccccggctt tccaccctc 2880
caccocctgt aggggacac atggtgttc ttggggggc ctcggacct gacagtlca 2940
gcagcagct tetgtctac caggtcaact gcaatgcctg gcttctgccc gactcacc 3000
gctcggcctc lgtggggccc ccaatggagg agtctgtgg ccaatgctg gcagcagctg 3060
ggagccgct gatatctct ggggttctg ggggagtgcc cctggccgc ctgctggacc 3120
tgaacctgc cccgaccoc tgcgcctgc tgtctcacc tgaagctgt aaccagctg 3180
ggcctgcac ctgtgcaat gggcctcctg tctcgggga tcaagccacc aggcctggct 3240
cgggggctc cccctgctc ccaatgctc gctcccggg ggaatgctg cgtctccga 3300
cctgagtgat gtcctggcc cgcactctc ggaacctga acctggagat ggaagcctg 3360
ccaccoccc cgttaagtgg tgtaccaact gccccaggg tcttgcatt ggaacccag 3420
ggtctgca ccttgagaa gactgtcga tcaaccagc agaggtcttc tgggcaagga 3480
actgctccg gctctgctg ggggtcctg actgcagca ctgcacggg gaggcaagt 3540
gatgtggac gggcagttc aagaggacg ggggacccg cgcactctc tctgtcagc 3600
ccactatga ctggagctg ttaagccact ctctgtgaa lgtgtcccc atgcccgtg 3660
aatcatcac cccactgccc tgcocacc cttgtcaact cctaccacac tgtacctct 3720
gctggactc taaggagca gatggggct gccagcactg tgtttggag agcagctgc 3780
agcagtgct gagccttcc taactgccc tgcgatgat ggcgggagg tggggcggc 3840
tgctccggg acctgagag ctgctccctg gctgtctca ggaactcag tgcctctgt 3900
gctcggcgg ccccatgct gctggtgtg cctgggggg ccagatagg ggtggcctg 3960
gcattgggg tggactcag gcccctcctg atgggtgac atgtggcct cgggggctc 4020
cctgggctc cctgtcctg cccctgagg acgagtgtc aacggggcc cagactgca 4080
acgagacga gaattgccc gaccagccc accgtatga gtcagctgc aagaaccgct 4140
ataccatga caacatgaca gggctgtgct gccctgtgt gcccagggc tgcgtgaacg 4200
gctcaagtg ggaaccgac cactgcctc gccacttgg ctttggggc cgcacatgct 4260
ccacggatg ccgtctgac cgcacactg aatggctgg tgttggggc cgtgacact 4320
gctgtctct cgcacaacc accaaggga gccactlga gcaagtctc cctgttctg 4380
tgggtcagc lgtcggagg ggaactcgc gccctgcca cgcctltgt cgtgaaata 4440
gccactctg catctcagc aaggagtac aatgtcaaa gggagagca aagaagtae 4500
cactggacc agaggagat gaaaactgg tgacagagg tctagttaa gacagggcc 4560
tgtgctgaa ctgcccaga aacagctatg gggagaaatg ccagagctgc ctgacggct 4620
acttctctc ggaacggga tgcacaaat gcaagttaa tggccacgg gacacatga 4680
acgagcagga tgggacgggc tglccalgtc agaataaac agagacgggc acatgcccg 4740
gcagctcccc cagtccctg cagactctc acaagtaca gtyggccaa gtcgggaaat 4800
catttcaag agtccctg ggggcccac agtctacc cctcatctc gggagcag 4860
agtctgctc ggaaccacc tcccagaca actgcttca tggcccaaa cgcggggcc 4920
tggcccggg ccgactgtc ctcttgggg tgcagccaa attcaacca gtcgacatc 4980
gctcaagct ggaactgacc tlogggggc tggacctca tgtctcacc tctatlgaca 5040
cctgtgtgt ccgtgtggc cccagactg gctcactac tgtcaacc cagcaacc 5100
cagcccacc acctcaca ccccctgag atggtgggc cggggggct ggggatccag 5160
gaggacagg ggcacagct gggccgggg cccacagca gccacggta cggagat 5220
ggcccgggg cctgattac taccgtacg tgacggacc gtcggcagt ctggtgtc 5280
gcccgtgct ggaaccgct gctacacat accacaaga gcaocattg ctaagtca 5340
cgcctctca cctgctgtg ctgggctgg gacaccaag tgggcccgc gcaaccgct 5400

```

WO 02/057454

PCT/US02/01339

```

cagccgactc gcaggccctg ctctctctcc gccaggacca gcccacatt gacctgttg 5460
tcctctcttc cgtctctctc tcctgctctc tcctctctcc ctcactctgt gtgctctctc 5520
ggaaaggccaa gcaggctctg gccagcgggc aggagcagcg cgggcacttg caggagatga 5580
ccaaagatggc cagcgcgccc ttcgccaagg tcacctctg ctccaccct gacctactg 5640
ccccgctctc cgcctggaag cgggtgggc tcccactcc cgcctcngc cgcctctagc 5700
ccttctctggc acccctctgct ctgaccgggg ccgggtggcc ctggggaacc atgggaggg 5760
gctgctgccc accagccatc ccgcaccaca ctgctggctc gcgagctggg ccaatcactc 5820
tcagccccc acagatggc atggctggcg tggccacct gotgctccag ctgctggcg 5880
ggcccactgc acccaacggc gcctgcctgg ggtcagccct cgtcacactg cgcacacggc 5940
tgcacagata ctgtgggggt ggtgggggtg ctggggcagc tgggcatggg actggtggcg 6000
gccggaaggg actgctgagc caggacaacc tcaccagact gtcctcttga catgoc 6056

```

```

<210> 22
<211> 1699
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2173285C81

```

```

<400> 22
ggggcgtctgg gggtagcgtg gccggagccgc tgcagcgtc gggcgagagt cggcggccgg 60
atccgaggag cagcggggcc lgaggccgag tcagctggcg gggcccccgg atcccccgac 120
agagcggcgg cggctctctg ccagccggta ggcctgctc ggcggcggcg gggaaagatg 180
tcagcgtaga gtctctggag cggcgggagc tctgagagag cctcctcact tggatccaga 240
catttaatgt ggatccacca tgcagaccgc tggaaagatt aacgaatggg gttgtgatgg 300
cccaggtctc tcaaaaagata gatcctgcat attttgatga aaattggcta aacagaaaca 360
aaactgaagt agggataaat tggaggctaa agataagcaa ttaaaagaaa attttaaag 420
gaactctggg ttataatcat gagattttag gacagcaaat taatgacttt acccttctgt 480
atctgaacct tatctgggag cactctgctg cagcagagct tggaaaggatg ctctcagctca 540
tcttagctgt tgcctgtaac tctgaaacga agcaagatga catccaaagc attatgatga 600
tggaggaatc ttttcaacat gttgtcatga cagccattca agcgtgatg agtaaaagaa 660
ctcctgtctc tcttggaaat gatcctctat ttagccttga tctctcagctg aagaaaacta 720
cagaggactc aaatgaagct ttgtcagcaa aggaagaaat tgctcaaaaga tgccatgaac 780
tggatgatga ggttgcagca ttgcaggaag agaaagatag tttgttgcca gagaatcagg 840
tattaaatgga aagactcaat caatctgatt ctatagaaga ccccaacagt ccagcaggaa 900
gaaggcattt gcagctccag actcaaatag aacagctcca agaaagaaaca ttcagactag 960
aagcagccaa agatgattat cgaatacgtt gtagaagatg agaaagaggat atctctgaac 1020
ttcggcaaca gaatgatgaa ctgacacctt tggcagatga agctcagctc ctgaaagatg 1080
agatgcagct gctgagacat tcttctgata aagtalctaa actagaaggt caagtagaat 1140
cttataaaaa gaagctagaa gacctgggtg atttaaagcg ccaggtttaa ctcttagaag 1200
agaagaatac catgtatgat cagaatactg tcagctctaga ggaagagtta agaaaggcca 1260
acgcagcggc aagtcacctc gaaacctaca agagacaggt aaaaagaaaca cagcactctg 1320
atgatggtt caggagagct ctgagctatg acatgtaget taccaaatt actaatttgg 1380
ttctatgata tctctgtttt tacttttct tballgtat gatloatta ggagactgag 1440
tctcactctg tcaccagcgc tggagtgagc tggcatgatc tcagctcact gcaacctcca 1500
ectcccaggc tcaagctatt ctctctcccc agcctctcga gttagctgaa ctacagcagc 1560
atgctgccac acctggctaa ttttttgat tttggtaaa acagggtttc actgtgttgc 1620
ccaggctggt cttgactcct tgagctcaag tgatccaca gccctcagct tccaaagttg 1680
taggattaca agcgtgagc

```

```

<210> 23
<211> 1661
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7487619C81

```

```

<400> 23
ctattttggtc cagcettatc gccccggact cgtaaccttc ggccatgccc tattataggg 60

```

WO 02/057454

PCT/US02/01339

```

tcccccaacc agtagtacag acattgcaat taactggcag agtcttttcc ctgattgaga 120
agtaaccctg tcaatgcatg cgaaggccag tgaaggagga ttttaggacc ataggtgacc 180
tcgttaaaat gagattgata tgcctgcca aaaccagcaa aagacagaga cctclgcccc 240
aaactgcaag aatgggatac tgcacaacc tgaatagac ttggaagagg actclggtcc 300
tccatgaga ttccagcttc tggggacaca ttgatagaga cctgttgaga cctctagacc 360
aggttccagc tgcctgggc ccaagagagg ttgagataa anttttaca gttttatgga 420
actaagtttg tgcctatttt ttaatatatc aatggaatc taatatagtg tttctctact 480
tctctctgca tgtgtgcttc tgtgtgtgtg cacctgtgtg catgtgtgtg agagagctgt 540
aaataatttc atcatcatct ctgtgaggga agctttgtaa caagcgaagt gcaagataac 600
tccagaatta tctaccctgt tgatgcagct tccacataga gaatggattc tcaatctcca 660
attaagtgtc aaatgctggg tgccttttat atcccagay ggagagagac caaggtgtgag 720
aagaaatgct caacccagc ctctgacagc cgttcaatcc caagggcctt ccccatgccc 780
cagggctgga cgcctccctc ttgggaatc tctgtgtgtg ttaagtgctc actgtgctgg 840
ggaacctcct catcctgctg gtgatcaggg tggatttcca cctccacacc cccatgtact 900
acttctctac caacctgtcc ttcattgaca tgggtttctc cactgtcagg gtgccaama 960
tgcctgatgac cttgtgtccc ccaagggaca gggctatctc cttccacagc tgcgtgctc 1020
agctctatct tttccacttc ctggggagca cagagtgttt cctotacaca gtcactggct 1080
acgatcgtca cctggcacc agtctaccgc tcaagtaacc cagcatgatg acggggcgt 1140
cgtctactct tctggcacc agcaattggc tcaagtgctc tctgcaactc gctgtccagg 1200
ccatattgac ttccatttg cctactctgg gaccacaact gatccagcac tattttgtgt 1260
atgcaccgcc cctcctgaaa ctggcctgtg cagacacctc agcctatagag actgtacttt 1320
tltgtactgt tggaaatgtg gctcgggctc gctttgtctc gatagtctct tccatgtctg 1380
ccatgctctg ttccactctg cggatccgca cctcagaggg gaagcagaga gctttcaga 1440
cctgtgctcc cnaactgtatc gtggtctctt gcttctttgg cctgtgtctt ttcatttacc 1500
tgaggccagc ctcaggaaa gctgtgtgat gagttgtggc cgtttctcac actgtctga 1560
cgccctctc caacctgtt gtgtacacc tgaggaaaca ggaggtgaag aaagctctgt 1620
tgaagctgaa agacaaagta gcacattctc agagcaata g 1661

```

```

<210> 24
<211> 2033
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7487607CB1

```

```

<400> 24
caatttcagg acctcgtggt tctgcccgcc tgcctctccc aaagtactgg gatttacagg 60
tcacctgccc tgcctggtag tatagaaaag gacatgattc aggtggggat gtcagagata 120
caaaagtgaa ggaagaagaa tggctgtctt tctttctatt cagaaaagat cccagaatc 180
tacatttatt cagcagaaac tatgcatgtt atactatgtg agggaaagat cagcagaaca 240
gctgttcgag cagaggaagc tgaactctct gcaggaagaa ggcaatagt attttctcat 300
acttagagat aatataccct gttcctcat ggcttccaaa cagtaatccc ttcttcagaa 360
atctacagc ttccacagaa attttgtga acagctaaa aacagaagat caagcamaa 420
caatctgctg tgtattgcaa cctaaagagt gagctgacct tccatttaga ggttaaatag 480
agagtaaaat ggaatgggaa aaccacacca tctgtgtgga atttttctg aagggacttt 540
ctggctacc cagaacttgag ttaecttttt ttgtgctcat cttcctaag tctgtgtgca 600
tcctctctgg gaatggtact ctctttttaa tcaagctctt ggacctcacc cttcaccacc 660
ctatgtactt cttctggggg aacctctcct tcttgacat ctgctacacc accaccteta 720
ttcctctcac gctgatgagc ttctttcagc aagaaagac catttccctt tctgctctgt 780
cagtgcagat gttcctgggc ttggccatgg ggacaacaga gctgtgtctt ctgggcatga 840
tggcctatga cctctatgct gctatctgca accctctgag atatcccatc atcatgagta 900
aggatgccta tgtaccatg gcagctgggt cctggatcat aggagctgtc aattctgagc 960
tacaatcagt gtttgtgta caattgcttt tctgaggaa taacatcacc aatcatttca 1020
cctgtgaaat tctggctgtc atgaaactgg cctgtgtgca catctcagc aatgagttca 1080
tcatgctgtg ggcacaaca ttgtctatat tgcaccttt gctattatc atgtctctt 1140
accgttaat cctgtgagc atcttcaaaa ttactcttc cgaagggaga agcaagctt 1200
cctcaactgt tccagccat ctgactgtgg tcaataatc ctatgggacc atctcttca 1260
tgaactatgc ccccaagctc aaagagacac ttaattogga tgaactggat gctaccgaca 1320
aaattatata catgtctatc ggggtgatga ctcccatgat gaatccttca atctacagtc 1380
ttgaaaccaa ggtgtgaaa gaggcagtaa aacctctact gaacagaggg tcttttagca 1440
agtgaatgca aatgtactg gaatatgac acacttgata ttgttgaac ttcagaatta 1500

```

WO 02/057454

PCT/US02/01339

```

tgtagaatt ttgggtactt ttactatitt tatgcatttt catatattat gttaaaaata 1550
tgagatacag catttcaaaa ttattgcatg tccactctag agaatttgca agatacaagg 1620
cagtaggatg aagaagaag aggggttacc tattactcta ctagtgggaa atggccocgt 1680
tbaacatttt tgaacagtaa ctltcatatt atgggttttt tttctgcat tggaaattggg 1740
tgggatggc ctlttcatgt tcaacttttt ccataactgtt atttcatagg caacatttca 1800
tajeatcttt caaatataat aaagccctct gttagagaaa aagcaaaaaa gaaaaaccc 1860
aacatagttt actcacattt tccagggaca agcctgtgtt atagtttccac attaatctcc 1920
agatcctgtt aaagcacta aataaccagt tttttttctt gtatttaaat tttgggtgct 1980
ggtgtccagc ctccaggttt ctggggacca tcccaaggg ggggaaata atg 2033

```

```

<210> 25
<211> 1659
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7487616CB1

```

```

<400> 25
ctattcacc tgagcaata ctgaccocatt tttcagcttc aacagagctt tctttaccct 60
cttgtttctc aggggtgtao caacaggggtt gaaaagagga gtcagcgtgg tgltagaaac 120
ggccacaacc ccatgcaagg cgtccctggc gcttggcctc aggtaaatga aagaccagg 180
gccaaagagc caaaggacca cgtacacatg ggaggcacag gcttgaagtg ctctgtgctt 240
ccctctctgag gtggcgatcc ccaggatgga acagacgatg gacacatagg acagcactat 300
caggacaagc cagcccgagg ccaactagccc aatattcaca aagatgacca tetcgttggc 360
tgagttgctt gcacagcca gtttccagat gggcggctgg tccagaaagt agtggctgat 420
ctggttgggt ccacagtagg gcaaatggaa agtcaatag gcttggacag cagagtgcag 480
agagccactg agccagtagc cgtgtggcag gaggccacac gaggccccag tcatcatgtt 540
gggtacactg agcgggtaac tgatggccag gtaggcatca taggacatga ctgttagag 600
gaaacactcg gtgtcccca ggaagtggaa aaaatagagc tgagccaagc agctgtgaa 660
ggagatagtc ctgccccttg gggacacca ggtcactcagc atlttgggca ccgtgacagt 720
ggagaaccac atgtcaatga aggacagctt ggttaggaag tagtaactgg cgtgtggag 780
gtagagatcc accctgatca ccagcagatg gaggagtttc cccagcacag tagcactga 840
aacccaccag agatctocaa agaggggggc gtccagcctt ggggcatggg gaaggccctt 900
gagatgaaac gctgtcacga ggcctggcctt ggacatttct tctcaccctt ggtctctctc 960
cctctgggga tataaagagc acccagcatt tagcacttaa ttgagaaatg agaatccatt 1020
ctctatgttg aaactgcatc aaccaggtag ataattcttg agttatcctg cacttccgtt 1080
gttacaagaag ttccctcaca gagatgatga tgaatttatt tcagcctctc tcaacacatc 1140
gcacaacagc gcacacacac agagacacac atgcagaaga aagtagagaa acactatatt 1200
agatttccat tgatatatta aaaaaataac acaacttag tgccataaaa cgtgtaaaat 1260
ttttatctca cctctctctt gggccaggtc agctggaccc tctgtgaga gtcctcaacg 1320
gtctcaatca atgtgtcccc agaagctgaa atctcatggg aggaccacag tctcttcca 1380
agtcatttgc aggtgttggc agratcccat tcttgcagtt ttgggcagag gtcctctgtt 1440
tctcgttttt ggcaggacat atcaatctca tttcacagag gtcacactatg gtcctaaaat 1500
ctctctccac tggcctctgc atgcaatgac aggttactt ctcaatcag gaaaagactc 1560
tgccagttaa ttgcaatgtc tgtactactg gatggggag cctataatac ggcattggcc 1620
aagtttaaga gtccggggcg ataaggctgg acccaatag 1659

```

```

<210> 26
<211> 1175
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7483204CB1

```

```

<400> 26
cccttttagg gttttttttt ttaataacag taataatgoc tctttgaaag tgacactcac 60
ctggataact ttttgaggtt aaagaagata atttacataa accaatcttg tttactttta 120
cagataaatt tttttttgat atcttggaaa gtcaagttct agcctgtcat tctcgttaat 180
atltctgtag cagtttgaaa caagaacaag gaagaatgga ctgggaaat tgctctctca 240

```

WO 02/057454

PCT/US02/01339

```

taactgattt tttctcttg ggaattacca ataaccaga gatgaaagt accctatttg 300
ctgtattctt ggctgtttat atcattaatt tctcagocaa tcttggaatg atagttttaa 360
tcagaatgga ttaccaactt caacacocaa tgtattctct cctcagtcac ctgtctttct 420
gtgatctctg ctattctact gcaactgggc ccaagatgct ggtagateta cttgocaaaga 480
acaagtcaat accctcttat ggctgtgctc tgcattctct ggtctctctg atctttgca 540
atctcagatg tctactcctg ccagtgatgg cctttgatcg gtaaaaggcc atcatcaac 600
cctcctctta tacagtcaac atgtctagca ggtg-ggcta tctactcttg actggggttt 660
atctcgtggg aatagcagat gctttgatac atatgacact ggctctccgc ctatgctctc 720
gtgggtctaa tggatataat cacttctctc gtgatctccc tctctcttta ttaactctct 780
gctcagatcc acaggtcaat gagttagtgt taltccocct ctctgggttt altgaaactg 840
gtaccatttc aggagtttct atttctatt gttatatac cctatcagtc ttgagatcc 900
actctcgtga gggggagctc aaagtctctc ctacatgca tccocacta tctcgggttg 960
caatlttcca ggggaactct ctctttatgt atttcggcc aagttctctc tattctctag 1020
atacagataa aatgacctca ttgttttaca cctctgtggt tccatctgtg aacccctga 1080
tttatagcct gaggaacaag gatgtgaaag aggcctgaa aaaactgaaa aatgaaattt 1140
tattttaagg aaatagtaaa aatacatggt tatac 1175

```

```

<210> 27
<211> 1737
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7472099CB1

```

```

<400> 27
tcagaattga gctgggggag gcttcagagg ctctcagtc ccaggaggct totctgagac 60
tgggaaggga gtggaaactc agtagccctt ggtcctggtt gcggcctccc tcccatgtct 120
tcattatctg atttggtaaa atatctacaa gacgggggag tccgggctgc tttctttgac 180
acgtgtcctc actccacacc tagaogctat gtccatgtct ggaagatgct ttaaggaag 240
ttttgatgca gagaagagtg gcgggtgctc tccctcgtct gccaccctc gcttactctc 300
agaagctaat tctctctctg tgactgatgt gactaacttt tttcaagcc tggcaggttc 360
ttatgacctc gcaaatgttc atgttaactg agatatacca gttggaagga aagaaatgct 420
ggttgctata gtacctctt gctacagagt caaatcttgt cctggaaga aatgtacaa 480
tattaact taaaacagta tttctacta agaagagaat tttattctga taaggtgag 540
gagccctatga agatcaacaa ggagaatttc caagagtcat gtcagcctcc agtatccct 600
caacacatcc aactctctc ttgttgatgt ggaltccagc cctggagcac ctgcacatct 660
ggatctccat cccctctca gcatatacac tggcctgctc tggaaactgc accctctctc 720
tcatcatcca gctgatgca gccctccatg agccatata cctctttctg gccatgttg 780
cagccatcga cctgctctt lcctctcag cctgcccaca aatgcttccc atattctgtt 840
tcaggatcgt ggagatcaac tttttgctt gctctgcca gatgttctc cttaactctc 900
tctccatcat ggagtcagca gtgctgctgg coactggcct tgaccgctat gtggccatct 960
gcaagccact gcactacacc acggtctctg ctgggtccct catcaccag attggcatgg 1020
ctctgtggc cgggctgtg acactaatga ctccactccc ctctctgctg agatgtttcc 1080
actactgccc agccctagtg atgcccct gctactctga acacatggct tggctcaggc 1140
tgcctgtggg aacactagtc ttcacacata tctatggcat tgcctgctc atgtttattg 1200
gaggtgtgga tctattctt atactctat ctatattct tctctctcag gcagttctac 1260
aacctctctc tcaggaggcc cgtacaaag cacttgggac atgtctctc cacataggtg 1320
ccatcttagc cttctacaca ccttcagtc tctctcagc catgcaccgt gtggcccctg 1380
gtgctgtgcc acagctccac attctctctg ccaattctc tctgctctc ccacccatgg 1440
tcaatcccat catctacggc gtttaagacca agcagatccg tgacagtctt gggagtattc 1500
ccgagaagag atgtgtgaat agagagttag gaataagtg aaaaagagtg gggcaagtg 1560
aatgctgtag tgggcccagg ctgtgctgag agtagatggg tgcagactc cacgtttatg 1620
tctttctctg tattatgaaa agataaaat atgtcctgaa gctcagctcc acagtctgtt 1680
aagaattgtg ggtctttgca ctggtaact ctggattgaa ctggctgact tgcggtc 1737

```

```

<210> 28
<211> 972
<212> DNA
<213> Homo sapiens

```

<220>

WO 02/057454

PCT/US02/01339

<221> misc_feature
<223> Incyte ID No: 7485443CB1

<400> 28
kaaataagaga gtaaaaaggc atgggaaaaa cacaccattc tgggtgaatt tttttotgag 60
ggactttctg gtcacccaag acttgggta ctcttttttg tgcctatctt cataaagtal 120
gtgtgtaatc ttctggggaa tggtaactctc attttaatac gactcttggc cctccacott 180
caacccccga tgtactctct tctggggaaa ctctctctct tggacatctg ctacacacc 240
acctctattc cctccacact agtggacttc ctttcagaaa gaaagaccat ttccttttct 300
ggctgtgcag tgcgaatgtt ccttggcttg gcaatgggga caacagagtg tgtgctcttg 360
ggcaatgatg cctttgacgg ctatgtggct atctgcaacc cctcgagata tcccatcacc 420
atgagcaaga atgcctatgt acccatggct gttgggtctt ggtttgcagg gattgtcaac 480
tctgcagtac aaactacatt tgtagtacaa ttgcctttct gcaggaagaa tgcctcaact 540
catttctcat gtgaaattct agctgtcatg aagttggcct gtgctgacat ctcaaggcaat 600
gagttctcca tgcctgttgg cacaatattg ttcacattga tgcactgct cttgatagtt 660
atctcttact cattaatcat ttccagcacc ctcaagattc actcctctga ggggagaagc 720
aaagctttct ctactgtctc agcccatctg actgtggtea taatattcta tgggaccatc 780
ctcttctagt atctgaagcc caagtctaaa gagacactta atctcaga-ga ctctggatgct 840
accgcaaaaa ttatctccat gttctatggg gtgatgactc ccatgatgaa tcccttaate 900
taccgtctta gaaacaagga tbtgaaagag gcagtaaac accctaccga cagaaggttc 960
tttagcaagt ga 972

<210> 29
<211> 1592
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3090414CB1

<400> 29
tagactacag gccacagagc tggaaaacttt tccaogctag gtcccagctt gagctgtgtc 60
ataacccaagt tctctaagat tcaactataa aaacttacta aggtalggag aagaagaaaa 120
cattaaatctg cagaagaacc agacaaattt tgagctattt cataacctac cagactctac 180
atgctaaccac tgaataaaac agaccataa ccagctctat ttattctgaa tggagtccca 240
ggactggaaag acacacaact ctggatttcc ttcccatctt gctctabgta tggttggctt 300
atggtaggga atgtggactt cctctactct attcaactat aggatgccct gcacaaaacc 360
atgtactact tcttggcoat gcttctcttt actgaccttg ttatgtctc tagtacaate 420
cctaaagccc tctgcatctt ctggtttcat ctcaagaca ttgatttga tgaatgctt 480
gtccagatgt tcttcatoca caecttca caagtgaggc ctgggtgtct tatgtttatg 540
gcccggatc gctatgtggc catctgctac ccttacgctt attcaactat cctcaccat 600
cctgtaattg caaagggttg gactgcccct ttccctgagag gggatattct cattattccc 660
tttactttcc tccccaagcg cctgcccctac tgcagagcca atatacttcc ccataacctac 720
tgtgaccaca tgtctgttag caaattgtcc tbtggtatg tcaagtcnaa tgcactctat 780
ggtgtgatgg ttgcccctctt gatttggggc tttsacatac tbtgtatcac catctctat 840
accatgatct tccgggcagt ggtccgectc tctcagcag atgctcgcca gaaggccttt 900
aatacctgca atgcccacat ttgtgcaatt gttttctctt ataectcage tttctctccc 960
ttcttttccc accgctttgg ggaacaata atccccctt cttgccaat cattgtgacc 1020
aataatttacc tgcctctacc acccaactatg aacctattg tctatggggt gaaaacccaa 1080
cagatcagag actgtgctat aaggatcttt tcaggttcta aggatccaaa atcctacagc 1140
atgtgaatga acacttgcca ggatggagaa gagaaggaaa gaattacttc tatttgcttc 1200
ttatgcagga gttcataaaa tctttctgga agtaactgat tgaacacaaa atggagtttg 1260
tgaactggtg cattctcaat aagtacctg gaaactcnaa catcattgga aggccaccca 1320
acacttctat aaatttttta ccttctact catgtgaagg accagctcaa taattaaacc 1380
atattttatt cgacaaaaaa aaaaaaaaaa aaaaaacggg ggggggcccg caactatgac 1440
gcccccaacc cgggaatata ctccggcagc ggaacaaca gggcgtaat ctgcacaaa 1500
ttttggccc cnaatggggc ccccggttg ggtccacct tbtactccca tcttgtgggg 1560
cgcccacggc ggggaaacct ccggccagga tg 1592

<210> 30
<211> 1480
<212> DNA

WO 02/057454

PCT/US02/01339

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7503710C81

<400> 30

```

ggtggaggaa cgcacaggag gccggggagcc cccacctacc ccttggtggag ctgcaggagc 60
aagggcattgc agccagtcatt gctggccctg ttggtccctgc ttctgctctg gggcctggcg 120
actccatgcc aggagctgct agagacgctg gccacgctcg ctgggattga caaggatgaa 180
ctcgcacaag ccattccagaa ctcaactggt ggggagccca ttctgcagaa tgtgctggga 240
tcggtcacag ctgtgaaccg gggcctcttg ggtcaggagc ggtctcttgg aggagcggcg 300
ttgctgggccc acggaggggt ttttggcgtt gtcgaggagc tctctggtct gaagattgag 360
gagctcaacc tgccaaaggt gttgtgagc ctgctgcccg gatttggggt gcagctgagc 420
ctgcaaccca aagtgggcat goattgctct ggccccctg gtggccttct gcagctggct 480
gcggagtgga acgtgacatc gcgggtggcg ctggccgtga gctcaaggcg cacaccocat 540
cttatctcca agcgtctcag caccgtctctg gcccacatca gctctgtctc agggctgctg 600
cccacaccac tctttggggt cgtggaaccg atgctctcca aggtgcttc gggactgctg 660
tgcccctggg tggacagtggt gctgggtgtg gtgaatgagc tcttggggcg tctgctggcg 720
cljgtgtccc ttgggctctt tgggtcctgt gaattctctc tggccacatt gctctctatc 780
tccaaccagt acatagaact ggacatcaac cctatcttga agagtgtagc tgggtatctc 840
attgacttcc ccaagtcccg tgcccacgcc aagtgctccc ccaagaagga ccacaatccc 900
caggtgatgg tggcaactga cctcttcaac accacgtttg gactctctga gaccaacggc 960
gcctctgaca tggacatcac cctgagctg gttcccagcg atgtcccact gacaactaca 1020
gacctggcag ctttctctcc tgaggtcatg actgtgctg ccaagctggc tccctcggtg 1080
accaagctgc acatctccct gtcccctgaa cggctcagtg tcaagtggtc ctctctcttt 1140
accaatgctt ttgacggatc gctgttagaa gaattgctca gccatgtggt cggggcagtg 1200
tatgaccaca agcttaacct ggcccctggt gttggaaatc ccttgcctaa ggttcttaat 1260
atcaattttt ccaattcagt tctaggatc gttagaantg ctgtggcagc tctctatgct 1320
ctgttagag catagaagat ggtgtcttc tcagatcagt ggaclatgoc atgttatttt 1380
gtctctggac taaggccctg tgaggtgcaa ctggtccact ttcatttttt gtcagagatg 1440
gagaataagg aattatatgt ttgtactagc actggaatag 1480

```

【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 July 2002 (25.07.2002)

PCT

(10) International Publication Number
WO 2002/057454 A3(51) International Patent Classification: C12N 15/12,
C07K 14/705, C12N 5/10, A01K 67/027, C07K 16/28,
C12Q 1/68, A61K 38/17, G01N 33/68, A61K 39/395(21) International Application Number:
PCT/US2002/001339

(22) International Filing Date: 16 January 2002 (16.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/262,838 19 January 2001 (19.01.2001) US
60/265,927 2 February 2001 (02.02.2001) US
60/271,196 23 February 2001 (23.02.2001) US
60/274,549 9 March 2001 (09.03.2001) US
60/334,179 28 November 2001 (28.11.2001) US(71) Applicant (for all designated States except US): INCYTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): LEE, Ernestine,
A. [US/US]; 624 Kains Street, Albany, CA 94706 (US).
CHAWLA, Narinder, K. [US/US]; 33 Union Square,
#712, Union City, CA 94587 (US). BAUGHN, Mariah,
R. [US/US]; 14244 Santiago Road, San Leandro, CA
94577 (US). AZIMZAI, Yalda [US/US]; 5518 Boulder
Canyon Drive, Castro Valley, CA 94552 (US). TANG,
Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA
95118 (US). YUE, Henry [US/US]; 826 Lois Avenue,
Sannyvale, CA 94087 (US). THANGAVELU, Kavitha
[IN/US]; 1950 Montecito Avenue, #23, Mountain View,
CA 94043 (US). XU, Yuming [US/US]; 1739 Walnut
Drive, Mountain View, CA 94040 (US). ARVIZU, Chan-
dra [US/US]; 490 Sherwood Way, #1, Menlo Park, CA
94025 (US). WARREN, Bridget, A. [US/US]; 10130
Parkwood Drive, #2, Cupertino, CA 95014 (US). YAO,
Monique, G. [US/US]; 1189 Woodgate Drive, Carmel, IN
40633 (US). AU-YOUNG, Janice [US/US]; 233 Golden
Eagle Lane, Brisbane, CA 94005 (US). HAFALLA, April,
J., A. [US/US]; 2227 Calle de Primavera, Santa Clara,
CA 95054 (US). ELLIOTT, Vicki, S. [US/US]; 3770Polton Place Way, San Jose, CA 95121 (US). KALLICK,
Deborah, A. [US/US]; 900 Olive Street, Menlo Park, CA
94025 (US). GANDHI, Ameena, R. [US/US]; 705 5th
Avenue, San Francisco, CA 94118 (US). RICHARDSON,
Thomas, W. [US/US]; 616 Canyon Road, #107, Redwood
City, CA 94062 (US). KHAN, Farrah, A. [IN/US]; 9445
Harrison Street, Des Plaines, IL 60016 (US). LI, Yan
[CN/US]; 3885 Corrina Way, Palo Alto, CA 94303 (US).
SWARNAKAR, Anita [CA/US]; 8 Locksley Avenue,
#5D, San Francisco, CA 94122 (US). RAMKUMAR,
Jayalaxmi [IN/US]; 34359 Maybird Circle, Fremont,
CA 94555 (US). NGUYEN, Dannie, B. [US/US]; 1403
Ridgewood Drive, San Jose, CA 95118 (US). GRAUL,
Richard, C. [US/US]; 682-29th Avenue, San Francisco,
CA 94121 (US). LI, Dyung, Aina, M. [US/US]; 233 Coy
Drive, San Jose, CA 95123 (US).(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, 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, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:
12 February 2002For 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: RECEPTORS AND MEMBRANE-ASSOCIATED PROTEINS

(57) Abstract: The invention provides human receptors and membrane-associated proteins (REMAP) and polynucleotides which identify and encode REMAP. 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 REMAP.

WO 2002/057454 A3

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 02/01339
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C07K14/705 C12N5/10 A01K67/027 C07K16/28 C12Q1/68 A61K38/17 G01N33/68 A61K39/395		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C07K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EMBL, EPO-Internal, BIOSIS, WPI Data, PAJ, SEQUENCE SEARCH		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] 15 June 1998 (1998-06-15) "Rattus norvegicus retinoic acid receptor alpha 2 isoform (RAR) mRNA, complete cds" Database accession no. U15211 XP002221752 cited in the application 87.3% identity with SEQ ID No 16 in 1680 bp overlap the whole document	1-20,23, 26-56,71
X	-& DATABASE SMALL [Online] 1 May 2000 (2000-05-01) "Retinoic acid receptor alpha 2 isoform" Database accession no. Q9QWJ1 XP002221753 cited in the application 97.4% identity with SEQ ID No 1 over entire length (459 aa) the whole document -/--	1-20,23, 26-56,71
<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 : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "Z" document member of the same patent family		
Date of the actual completion of the international search 28 March 2003		Date of mailing of the international search report 04.04.03
Name and mailing address of the ISA European Patent Office, P.B. 6518 Patentplan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax. (+31-70) 340-3016		Authorized officer Weikl, M

Form PCT/ISA210 (second sheet) (July 1999)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 02/01339

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of documents, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>AKMAL K M ET AL: "REGION-SPECIFIC LOCALIZATION OF RETINOIC ACID RECEPTOR-ALPHA EXPRESSION IN THE RAT EPIDIDYMS" BIOLOGY OF REPRODUCTION, SOCIETY FOR THE STUDY OF REPRODUCTION, CHAMPAIGN, IL, US, vol. 54, 1996, pages 1111-1119, XP000901245 ISSN: 0006-3363 cited in the application the whole document</p>	1-20, 23, 26-56, 71
A	<p>AKMAL KARIN M ET AL: "Ligand-dependent regulation of retinoic acid receptor alpha in rat testis: In vivo response to depletion and repletion of vitamin A." ENDOCRINOLOGY, vol. 139, no. 3, March 1998 (1998-03), pages 1239-1248, XP002221751 ISSN: 0013-7227 cited in the application the whole document</p>	1-20, 23, 26-56, 71
A	<p>WO 95 35505 A (UNIV LELAND STANFORD JUNIOR) 28 December 1995 (1995-12-28) cited in the application the whole document</p>	46-55

Form PCT/ISA/210 (continuation of second sheet) (July 1997)

INTERNATIONAL SEARCH REPORT	International application No. PCT/US 02/01339
Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)	
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:	
1. <input checked="" type="checkbox"/> Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: see FURTHER INFORMATION sheet PCT/ISA/210	
2. <input checked="" type="checkbox"/> Claims Nos.: 21, 22, 24, 25, 48, 52-55 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: see FURTHER INFORMATION sheet PCT/ISA/210	
3. <input type="checkbox"/> Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).	
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)	
This International Searching Authority found multiple inventions in this international application, as follows:	
see additional sheet	
1. <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.: 1-55 (insofar as they relate to SEQ ID Nos 1 and 16), 56, 71	
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 02/01339

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 30, 33 and 35 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Although claim 19 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.2

Claims Nos.: 21, 22, 24, 25, 48, 52-55

Present claims 21, 22, 24 and 25 relate to a compound defined by reference to a desirable characteristic or property, namely agonistic/antagonistic activity concerning a polypeptide having SEQ ID No 1.

The claims cover all compositions and methods incorporating such compounds having this characteristic or property, whereas the application provides no support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for these compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of 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 compound 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, the no search has been carried out for claims 21, 22, 24 and 25.

Present claims 48 and dependent claims 52-55 relate to a nucleotide sequence which is only defined by reference to a desirable property namely hybridization with at least 30 contiguous nucleotides of the polynucleotide of claim 12 (the polynucleotide of the invention). Due to the lack of hybridization conditions and further structural or functional definitions of this nucleotide, basically every nucleotide could fall within the definition as set out by claim 48. Any nucleotide sequence can be imagined to hybridize under at least one particular condition to an arbitrary 30 nucleotide stretch of SEQ ID No 1. This lack of clarity is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been restricted to nucleotide sequences which are complementary to stretches of SEQ ID No 1.

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 02/01339

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-55 (insofar as they relate to SEQ ID Nos 1 and 16), 56, 71
2. Claims: 1-55 (insofar as they relate to SEQ ID Nos 2 and 17), 57, 72
3. Claims: 1-55 (insofar as they relate to SEQ ID Nos 3 and 18), 58, 73
4. Claims: 1-55 (insofar as they relate to SEQ ID Nos 4 and 19), 59, 74
5. Claims: 1-55 (insofar as they relate to SEQ ID Nos 5, 15, 20 and 30), 60, 70, 75, 85
claims relating to splice variants 5 and 15
6. Claims: 1-55 (insofar as they relate to SEQ ID Nos 6 and 21), 61, 76
7. Claims: 1-55 (insofar as they relate to SEQ ID Nos 7 and 22), 62, 77
8. Claims: 1-55 (insofar as they relate to SEQ ID Nos 8 and 23), 63, 78
9. Claims: 1-55 (insofar as they relate to SEQ ID Nos 9 and 24), 64, 79

International Application No. PCT/US 02/01339

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

10. Claims: 1-55 (insofar as they relate to SEQ ID Nos 10 and 25), 65, 80
11. Claims: 1-55 (insofar as they relate to SEQ ID Nos 11 and 26), 66, 81
12. Claims: 1-55 (insofar as they relate to SEQ ID Nos 12 and 27), 67, 82
13. Claims: 1-55 (insofar as they relate to SEQ ID Nos 13 and 28), 68, 83
14. Claims: 1-55 (insofar as they relate to SEQ ID Nos 14 and 29), 69, 84

INTERNATIONAL SEARCH REPORT
Information on patent family members

International Application No
PCT/US 02/01339

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9535505 A	28-12-1995	US 5807522 A	15-09-1998
		AT 180570 T	15-06-1999
		AU 709276 B2	26-08-1999
		AU 2862995 A	15-01-1996
		CA 2192095 A1	28-12-1995
		DE 69509925 D1	01-07-1999
		DE 69509925 T2	09-12-1999
		DK 804731 T3	08-11-1999
		EP 0804731 A1	05-11-1997
		EP 0913485 A1	06-05-1999
		ES 2134481 T3	01-10-1999
		GR 3030430 T3	30-09-1999
		JP 18503841 T	07-04-1998
		JP 3272365 B2	08-04-2002
		JP 2002243736 A	28-08-2002
		US 6110426 A	29-08-2000
		WO 9535505 A1	28-12-1995
US 2003012695 A1	16-01-2003		

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード(参考)
A 6 1 P 1/18	A 6 1 P 3/06	4 B 0 6 5
A 6 1 P 3/06	A 6 1 P 3/10	4 C 0 8 4
A 6 1 P 3/10	A 6 1 P 5/40	4 H 0 4 5
A 6 1 P 5/40	A 6 1 P 7/04	
A 6 1 P 7/04	A 6 1 P 7/06	
A 6 1 P 7/06	A 6 1 P 7/08	
A 6 1 P 7/08	A 6 1 P 9/10	
A 6 1 P 9/10	A 6 1 P 11/00	
A 6 1 P 11/00	A 6 1 P 11/06	
A 6 1 P 11/06	A 6 1 P 17/02	
A 6 1 P 17/02	A 6 1 P 17/06	
A 6 1 P 17/06	A 6 1 P 19/02	
A 6 1 P 19/02	A 6 1 P 19/10	
A 6 1 P 19/10	A 6 1 P 21/02	
A 6 1 P 21/02	A 6 1 P 21/04	
A 6 1 P 21/04	A 6 1 P 25/00	
A 6 1 P 25/00	A 6 1 P 25/08	
A 6 1 P 25/08	A 6 1 P 25/14	
A 6 1 P 25/14	A 6 1 P 25/16	
A 6 1 P 25/16	A 6 1 P 25/22	
A 6 1 P 25/22	A 6 1 P 25/28	
A 6 1 P 25/28	A 6 1 P 29/00	1 0 1
A 6 1 P 29/00	A 6 1 P 31/04	
A 6 1 P 31/04	A 6 1 P 31/10	
A 6 1 P 31/10	A 6 1 P 31/12	
A 6 1 P 31/12	A 6 1 P 31/18	
A 6 1 P 31/18	A 6 1 P 33/00	
A 6 1 P 33/00	A 6 1 P 33/02	
A 6 1 P 33/02	A 6 1 P 35/00	
A 6 1 P 35/00	A 6 1 P 35/02	
A 6 1 P 35/02	A 6 1 P 37/02	
A 6 1 P 37/02	A 6 1 P 37/08	
A 6 1 P 37/08	C 0 7 K 14/47	
C 0 7 K 14/47	C 0 7 K 14/705	
C 0 7 K 14/705	C 0 7 K 16/18	
C 0 7 K 16/18	C 0 7 K 16/28	
C 0 7 K 16/28	C 0 7 K 16/42	
C 0 7 K 16/42	C 0 7 K 19/00	
C 0 7 K 19/00	C 1 2 M 1/00	A
C 1 2 M 1/00	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 P 21/02	C
C 1 2 N 5/10	C 1 2 P 21/08	
C 1 2 P 21/02	C 1 2 Q 1/02	
C 1 2 P 21/08	C 1 2 Q 1/68	A
C 1 2 Q 1/02	G 0 1 N 33/15	Z

C 1 2 Q 1/68	G 0 1 N 33/50	Z
G 0 1 N 33/15	G 0 1 N 33/53	D
G 0 1 N 33/50	G 0 1 N 33/53	M
G 0 1 N 33/53	G 0 1 N 33/566	
G 0 1 N 33/566	G 0 1 N 33/577	B
G 0 1 N 33/577	G 0 1 N 37/00	1 0 2
G 0 1 N 37/00	C 1 2 N 15/00	F
	C 1 2 N 5/00	A
	A 6 1 K 37/02	

- (31)優先権主張番号 60/274,549
 (32)優先日 平成13年3月9日(2001.3.9)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/334,179
 (32)優先日 平成13年11月28日(2001.11.28)
 (33)優先権主張国 米国(US)

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

- (72)発明者 チョーラ、ナリンダー・ケイ
 アメリカ合衆国カリフォルニア州9 4 5 8 7・ユニオンシティ・# 7 1 2・ユニオンスクエア 3
 3
 (72)発明者 ボーゲン、マライア・アール
 アメリカ合衆国カリフォルニア州9 4 5 7 7・サンレアンドロ・サンティアゴロード 1 4 2 4 4
 (72)発明者 アジムザイ、ヤルダ
 アメリカ合衆国カリフォルニア州9 4 5 5 2・カストロパレー・ボールダーキャニオンドライブ
 5 5 1 8
 (72)発明者 タング、ワイ・トム
 アメリカ合衆国カリフォルニア州9 5 1 1 8・サンノゼ・ランウィックコート 4 2 3 0
 (72)発明者 ユエ、ヘンリー
 アメリカ合衆国カリフォルニア州9 4 0 8 7・サニーベイル・ルイスアベニュー 8 2 6
 (72)発明者 サンガベル、カピサ
 アメリカ合衆国カリフォルニア州9 4 0 4 3・マウンテンビュー・# 2 3・モンテシトアベニュー
 1 9 5 0
 (72)発明者 スー、ユーミング
 アメリカ合衆国カリフォルニア州9 4 0 4 0・マウンテンビュー・ウォルナットドライブ 1 7 3
 9
 (72)発明者 アーピズ、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 ワレン、ブリジット・エイ
 アメリカ合衆国カリフォルニア州9 5 0 1 4・クーペルティノー・# 2・パークウッドドライブ
 1 0 1 3 0
 (72)発明者 ヤオ、モニック・ジー
 アメリカ合衆国インディアナ州4 6 0 3 3・カーメル・ウッドゲートドライブ 1 1 8 9
 (72)発明者 オウ・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・プリズペーン・ゴールデンイーグルレーン 2 3 3

- (72)発明者 ハファリア、エープリル・ジェイ・エイ
アメリカ合衆国カリフォルニア州9 5 0 5 4・サンタクララ・コーレデブリマベータ 2 2 2 7
- (72)発明者 エリオット、ビッキー・エス
アメリカ合衆国カリフォルニア州9 5 1 2 1・サンノゼ・ポルトンプレイスウェイ 3 7 7 0
- (72)発明者 カリック、デボラー・エイ
アメリカ合衆国カリフォルニア州9 4 0 2 5・メンロパーク・オリーブストリート 9 0 0
- (72)発明者 ガンディー、アミーナ・アール
アメリカ合衆国カリフォルニア州9 4 1 1 8・サンフランシスコ・フィフスアベニュー 7 0 5
- (72)発明者 リチャードソン、トーマス・ダブリュ
アメリカ合衆国カリフォルニア州9 4 0 6 2・レッドウッドシティ・# 1 0 7・キャニオンロード
6 1 6
- (72)発明者 カーン、ファラ・エイ
アメリカ合衆国イリノイ州6 0 0 1 6・デブレインズ・ハリソンストリート 9 4 4 5
- (72)発明者 リュ、ヤン
アメリカ合衆国カリフォルニア州9 4 3 0 3・パロアルト・コリーナウェイ 3 8 8 5
- (72)発明者 スウォーナカール、アニータ
アメリカ合衆国カリフォルニア州9 4 1 2 2・サンフランシスコ・# 5ディー・ロックスリーアベ
ニュー 8
- (72)発明者 ランクマール、ジャヤラクシミ
アメリカ合衆国カリフォルニア州9 4 5 5 5・フレモント・メイバードサークル 3 4 3 5 9
- (72)発明者 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 グラール、リチャード・シー
アメリカ合衆国カリフォルニア州9 4 1 2 1・サンフランシスコ・トゥエンティナインスアベニ
ュー 6 8 2
- (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
- F ターム(参考) 2G045 AA34 AA35 AA40 BA11 BB50 DA13 DA36 FB02 FB03
4B024 AA01 AA11 AA12 BA43 BA63 CA02 CA04 CA09 HA14 HA17
4B029 AA07 FA15
4B063 QA01 QA05 QA18 QQ43 QQ52 QR55 QR62 QR80 QS25 QS34
4B064 AG20 AG27 CA19 CC24 DA01 DA05 DA07
4B065 AB01 BA02 CA24 CA25 CA44 CA46
4C084 AA01 AA07 AA16 NA14 ZA021 ZA051 ZA061 ZA151 ZA161 ZA451
ZA541 ZA551 ZA591 ZA661 ZA681 ZA751 ZA891 ZA941 ZA961 ZA971
ZB051 ZB071 ZB081 ZB131 ZB151 ZB261 ZB331 ZB351 ZB371 ZB381
ZC081 ZC331 ZC351
4H045 AA10 AA11 AA20 AA30 BA10 BA41 CA40 DA50 DA76 EA21
EA22 EA27 EA28 EA50 EA51 EA54 FA74

專利名稱(譯)	受体和膜結合蛋白		
公開(公告)号	JP2005500008A	公開(公告)日	2005-01-06
申請号	JP2002558506	申請日	2002-01-16
[標]申請(專利權)人(譯)	洞察Genomics公司		
申請(專利權)人(譯)	洞察基因组公司		
[標]發明人	リーアーンステーンエイ チョーラナリンダーケイ ボーグンマライアアール アジムザイヤルダ タングワイトム ユエヘンリー サンガベルカピサ スーユーミング アービズチャンドラ ワレンブリジットエイ ヤオモニークジー オウヤングジャンニス ハファリアエープリルジェイエイ エリオットビッキーエス カリックデボラーエイ ガンディーアミーナアール リチャードソントーマスダブリュ カーンファアラエイ リュヤン スウォーナカールアニータ ランクマールジャヤラクシミ ニュエンダニエルビー グラールリチャードシー リュデユングアイナエム		
發明人	リー、アーンステーン・エイ チョーラ、ナリンダー・ケイ ボーグン、マライア・アール アジムザイ、ヤルダ タング、ワイトム ユエ、ヘンリー サンガベル、カピサ スー、ユーミング アービズ、チャンドラ ワレン、ブリジット・エイ ヤオ、モニーク・ジー オウ・ヤング、ジャンニス ハファリア、エープリル・ジェイ・エイ エリオット、ビッキー・エス カリック、デボラー・エイ ガンディー、アミーナ・アール リチャードソン、トーマス・ダブリュ カーン、ファラ・エイ リュ、ヤン スウォーナカール、アニータ		

ランクマール、ジャヤラクシミ
ニュエン、ダニエル・ビー
グラール、リチャード・シー
リュ、デュング・アイナ・エム

IPC分类号	G01N33/50 A61K38/00 A61K45/00 A61P1/04 A61P1/16 A61P1/18 A61P3/06 A61P3/10 A61P5/40 A61P7/04 A61P7/06 A61P7/08 A61P9/10 A61P11/00 A61P11/06 A61P17/02 A61P17/06 A61P19/02 A61P19/10 A61P21/02 A61P21/04 A61P25/00 A61P25/08 A61P25/14 A61P25/16 A61P25/22 A61P25/28 A61P29/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P33/02 A61P35/00 A61P35/02 A61P37/02 A61P37/08 C07K14/47 C07K14/705 C07K16/18 C07K16/28 C07K16/42 C07K19/00 C12M1/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N5/09 C12N5/12 C12P21/02 C12P21/08 C12Q1/02 C12Q1/68 G01N33/15 G01N33/53 G01N33/566 G01N33/577 G01N37/00
CPC分类号	A61K38/00 A61P1/04 A61P1/16 A61P1/18 A61P11/00 A61P11/06 A61P17/02 A61P17/06 A61P19/02 A61P19/10 A61P21/02 A61P21/04 A61P25/00 A61P25/08 A61P25/14 A61P25/16 A61P25/22 A61P25/28 A61P29/00 C07K14/705
FI分类号	C12N15/00.ZNA.A A61K45/00 A61P1/04 A61P1/16 A61P1/18 A61P3/06 A61P3/10 A61P5/40 A61P7/04 A61P7/06 A61P7/08 A61P9/10 A61P11/00 A61P11/06 A61P17/02 A61P17/06 A61P19/02 A61P19/10 A61P21/02 A61P21/04 A61P25/00 A61P25/08 A61P25/14 A61P25/16 A61P25/22 A61P25/28 A61P29/00.101 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P33/02 A61P35/00 A61P35/02 A61P37/02 A61P37/08 C07K14/47 C07K14/705 C07K16/18 C07K16/28 C07K16/42 C07K19/00 C12M1/00.A 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 G01N33/577.B G01N37/00.102 C12N15/00.F C12N5/00.A A61K37/02
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/AA40 2G045/BA11 2G045/BB50 2G045/DA13 2G045/DA36 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/AA12 4B024/BA43 4B024/BA63 4B024/CA02 4B024/CA04 4B024/CA09 4B024/HA14 4B024/HA17 4B029/AA07 4B029/FA15 4B063/QA01 4B063/QA05 4B063/QA18 4B063/QQ43 4B063/QQ52 4B063/QR55 4B063/QR62 4B063/QR80 4B063/QS25 4B063/QS34 4B064/AG20 4B064/AG27 4B064/CA19 4B064/CC24 4B064/DA01 4B064/DA05 4B064/DA07 4B065/AB01 4B065/BA02 4B065/CA24 4B065/CA25 4B065/CA44 4B065/CA46 4C084/AA01 4C084/AA07 4C084/AA16 4C084/NA14 4C084/ZA021 4C084/ZA051 4C084/ZA061 4C084/ZA151 4C084/ZA161 4C084/ZA451 4C084/ZA541 4C084/ZA551 4C084/ZA591 4C084/ZA661 4C084/ZA681 4C084/ZA751 4C084/ZA891 4C084/ZA941 4C084/ZA961 4C084/ZA971 4C084/ZB051 4C084/ZB071 4C084/ZB081 4C084/ZB131 4C084/ZB151 4C084/ZB261 4C084/ZB331 4C084/ZB351 4C084/ZB371 4C084/ZB381 4C084/ZC081 4C084/ZC331 4C084/ZC351 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/BA41 4H045/CA40 4H045/DA50 4H045/DA76 4H045/EA21 4H045/EA22 4H045/EA27 4H045/EA28 4H045/EA50 4H045/EA51 4H045/EA54 4H045/FA74
優先権	60/262838 2001-01-19 US 60/265927 2001-02-02 US 60/271196 2001-02-23 US 60/274549 2001-03-09 US 60/334179 2001-11-28 US
外部リンク	Espacenet

摘要(译)

本发明提供了鉴定和编码REMAP的人受体和膜相关蛋白 (REMAP) 以及多核苷酸。 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REMAP异常表达有关的疾病的方法。

Image Image ID	Image Image ID	Image Image ID	Image Image ID	Image Image ID
104479	16	132479021	16	132479021
211940	2	211940021	17	211940021
723167	3	723167021	18	723167021
122416	4	122416021	19	122416021
147061	5	147061021	20	147061021
772454	6	772454021	21	772454021
207186	7	207186021	22	207186021
194763	8	194763021	23	194763021
146107	9	146107021	24	146107021
146765	10	146765021	25	146765021
146204	11	146204021	26	146204021
147299	12	147299021	27	147299021
146543	13	146543021	28	146543021
139414	14	139414021	29	139414021
152701	15	152701021	30	152701021