

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-511208

(P2004-511208A)

(43) 公表日 平成16年4月15日(2004.4.15)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A 2 B O 3 O
A O 1 H 5/00	A O 1 H 5/00	A 2 G O 4 5
A O 1 K 67/027	A O 1 K 67/027	4 B O 2 4
A 6 1 K 38/00	A 6 1 K 39/395	D 4 B O 6 3
A 6 1 K 39/395	A 6 1 K 39/395	N 4 B O 6 4
	審査請求 未請求 予備審査請求 有	(全 322 頁) 最終頁に続く

(21) 出願番号	特願2001-580948 (P2001-580948)	(71) 出願人	301005050
(86) (22) 出願日	平成13年4月27日 (2001.4.27)		インサイト・ゲノミックス・インコーポレイテッド
(85) 翻訳文提出日	平成14年10月25日 (2002.10.25)		アメリカ合衆国カリフォルニア州94304・パロアルト・ポータードライブ 3160
(86) 国際出願番号	PCT/US2001/013862		
(87) 国際公開番号	W02001/083524	(74) 代理人	100089266
(87) 国際公開日	平成13年11月8日 (2001.11.8)		弁理士 大島 陽一
(31) 優先権主張番号	60/200,184	(72) 発明者	ラル、ブリーティ
(32) 優先日	平成12年4月28日 (2000.4.28)		アメリカ合衆国カリフォルニア州95056・サンタクララ・ビーオーボックス 5142
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/201,875		
(32) 優先日	平成12年5月4日 (2000.5.4)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/202,090		
(32) 優先日	平成12年5月4日 (2000.5.4)		
(33) 優先権主張国	米国 (US)		
			最終頁に続く

(54) 【発明の名称】 RNA代謝タンパク質

(57) 【要約】

本発明は、ヒトRNA代謝タンパク質 (RMEP) と、RMEPを同定及びコードするポリヌクレオチドとを提供する。本発明はまた、発現ベクター、宿主細胞、抗体、アゴニスト、およびアンタゴニストを提供する。更に、本発明は、RMEPの異常発現に関連する疾患の診断・治療・予防方法を提供する。

【特許請求の範囲】

【請求項 1】

単離されたポリペプチドであって、

- (a) SEQ ID NO: 1 及び SEQ ID NO: 47 (SEQ ID NO: 1 - 47) からなる群から選択されたアミノ酸を含むポリペプチドと、
- (b) SEQ ID NO: 1 - 47 からなる群から選択されたアミノ酸配列と少なくとも 90% の同一性を有するアミノ酸配列を含む天然のポリペプチドと、
- (c) SEQ ID NO: 1 - 47 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 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されることを特徴とする単離されたポリペプチド。

10

【請求項 2】

SEQ ID NO: 1 - 47 からなる群から選択された請求項 1 の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 4】

請求項 2 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 5】

SEQ ID NO: 48 - 94 からなる群から選択された請求項 4 の単離されたポリヌクレオチド。

20

【請求項 6】

請求項 3 のポリヌクレオチドに機能的に結合されたプロモーター配列を含む組換えポリヌクレオチド。

【請求項 7】

請求項 6 の組換えポリヌクレオチドで形質転換された細胞。

【請求項 8】

請求項 6 の組換えポリヌクレオチドを含む遺伝子組換え生物。

【請求項 9】

請求項 1 のポリペプチドを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結合されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養するステップと、
- (b) そのように発現したポリペプチドを回収するステップとを含むことを特徴とする請求項 1 のポリペプチドの生産方法。

30

【請求項 10】

請求項 1 のポリペプチドに特異的に結合する単離された抗体。

【請求項 11】

単離されたポリヌクレオチドであって、

- (a) SEQ ID NO: 48 - 94 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、
- (b) SEQ ID NO: 48 - 94 からなる群から選択されたポリヌクレオチド配列と少なくとも 90% の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、
- (c) 前記 (a) のポリヌクレオチドに相補的なポリヌクレオチドと、
- (d) 前記 (b) のポリヌクレオチドに相補的なポリヌクレオチドと、
- (e) 前記 (a) 乃至 (d) の RNA 等価物とで構成される群から選択された単離されたポリヌクレオチド。

40

【請求項 12】

50

請求項 1 1 のポリヌクレオチドの少なくとも 6 0 個の連続する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 1 3】

サンプルにおける請求項 1 1 に記載のポリヌクレオチド配列を有する標的ポリヌクレオチドを検出する方法であって、

(a) 前記サンプルをプローブでハイブリダイズするステップであって、前記プローブが、前記サンプル内の前記標的ポリヌクレオチドに相補的な配列を含む少なくとも 2 0 個の連続するヌクレオチドを含み、前記プローブと前記標的ポリヌクレオチドまたはその断片との間でハイブリダイゼーション複合体が形成される条件下で、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズする、該ステップと、

10

(b) 前記ハイブリダイゼーション複合体が存在するか否かを検出し、存在する場合には必要に応じてその収量を測定するステップとを含むことを特徴とする標的ポリヌクレオチドの検出方法。

【請求項 1 4】

前記プローブが少なくとも 6 0 個の連続するヌクレオチドを含むことを特徴とする請求項 1 3 に記載の方法。

【請求項 1 5】

サンプルにおける請求項 1 1 のポリヌクレオチド配列を有する標的ポリヌクレオチドを検出する方法であって、

(a) ポリメラーゼ連鎖反応増幅法を用いて、前記標的ポリヌクレオチドまたはその断片を増幅するステップと、

20

(b) 増幅された前記標的ポリヌクレオチドまたはその断片が存在するか否かを検出し、存在する場合には必要に応じてその収量を測定するステップとを含むことを特徴とする標的ポリヌクレオチドの検出方法。

【請求項 1 6】

請求項 1 のポリペプチド及び医薬的に容認できる賦形剤を含む組成物。

【請求項 1 7】

前記ポリペプチドが、S E Q I D N O : 1 - 4 7 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 6 の組成物。

【請求項 1 8】

30

機能的な R M E P (新規の R N A 代謝タンパク質) の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 1 6 の組成物を投与することを含むことを特徴とする治療方法。

【請求項 1 9】

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

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝露するステップと、

(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 2 0】

40

請求項 1 9 の方法によって同定されたアゴニスト化合物及び医薬的に容認できる賦形剤を含む組成物。

【請求項 2 1】

機能的な R M E P の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 2 0 の組成物を投与することを含むことを特徴とする治療方法。

【請求項 2 2】

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

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝露するステップと、

(b) 前記サンプルにおけるアンタゴニスト活性を検出するステップとを含むことを特徴

50

とするスクリーニング方法。

【請求項 23】

請求項 22 の方法によって同定されたアンタゴニスト化合物及び医薬的に容認できる賦形剤を含む組成物。

【請求項 24】

機能的な R M E P の過剰な発現に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 23 の組成物を投与することを含むことを特徴とする治療方法。

【請求項 25】

請求項 1 のポリペプチドに特異的に結合する化合物をスクリーニングする方法であって、
(a) 請求項 1 のポリペプチドを好適な条件下で少なくとも 1 つの試験化合物と結合させるステップと、
(b) 請求項 1 のポリペプチドと前記試験化合物との結合を検出して、請求項 1 のポリペプチドと特異的に結合する化合物を同定する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。

10

【請求項 26】

請求項 1 のポリペプチドの活性を変化させる化合物をスクリーニングする方法であって、
(a) 請求項 1 のポリペプチドを、そのポリペプチドの活性が許容される条件下で少なくとも 1 つの試験化合物と結合させるステップと、
(b) 前記試験化合物の存在下での請求項 1 のポリペプチドの活性を評価するステップと、
(c) 前記試験化合物の存在下での請求項 1 のポリペプチドの活性と、前記試験化合物の非存在下での請求項 1 のポリペプチドの活性とを比較するステップとを含み、
前記試験化合物の存在下での請求項 1 のポリペプチドの活性の変化が、請求項 1 のポリペプチドの活性を変化させる化合物の存在を示唆することを特徴とするスクリーニング方法。

20

【請求項 27】

請求項 5 の配列を含む標的ポリヌクレオチドの発現を変化させるのに有効な化合物をスクリーニングする方法であって、
(a) 前記標的ポリヌクレオチドの発現に好適な条件下で、前記標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、
(b) 前記標的ポリヌクレオチドの発現の変化を検出するステップと、
(c) 様々な量の前記化合物の存在下での前記標的ポリヌクレオチドの発現と、前記化合物の非存在下での前記標的ポリヌクレオチドの発現とを比較する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。

30

【請求項 28】

試験化合物の毒性を評価する方法であって、
(a) 核酸を含む生体サンプルを前記試験化合物で処理するステップと、
(b) 処理した前記生体サンプルの核酸を、請求項 11 のポリヌクレオチドの少なくとも 20 個の連続するヌクレオチドを含むプローブとハイブリダイズさせるステップであって、このハイブリダイゼーションが、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われ、前記標的ポリヌクレオチドが、請求項 11 のポリヌクレオチドのポリヌクレオチド配列またはその断片を含むポリヌクレオチドである、前記ステップと、
(c) ハイブリダイゼーション複合体の収量を定量するステップと、
(d) 前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処理の生体サンプルにおけるハイブリダイゼーション複合体の収量と比較するステップとを含み、
前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量の差が試験化合物の毒性を示唆することを特徴とする試験化合物の毒性評価方法。

40

【請求項 29】

50

生体サンプルにおける R M E P の発現に関連する病態または疾患の診断テストであって、
 (a) 請求項 1 0 の抗体と前記ポリペプチドとの結合に好適な条件下で、前記生体サンプルを請求項 1 0 の抗体と結合させて、抗体 / ポリペプチド複合体を形成するステップと、
 (b) 前記複合体を検出するステップとを含み、
 前記複合体の存在が、前記生体サンプルにおける前記ポリヌクレオチドの存在と相関性を有することを特徴とする診断テスト。

【請求項 3 0】

前記抗体が、

- (a) キメラ抗体、
- (b) 一本鎖抗体、
- (c) F a b 断片、
- (d) F (a b ') ₂ 断片、または
- (e) ヒト化抗体であることを特徴とする請求項 1 0 に記載の抗体。

10

【請求項 3 1】

請求項 1 0 の抗体及び許容される賦形剤を含む組成物。

【請求項 3 2】

患者において、R M E P の発現に関連する病態または疾患を診断する方法であって、請求項 3 1 の組成物を前記患者に有効量投与することを含むことを特徴とする方法。

【請求項 3 3】

前記抗体が標識されていることを特徴とする請求項 3 1 に記載の組成物。

20

【請求項 3 4】

患者において、R M E P の発現に関連する病態または疾患を診断する方法であって、請求項 3 3 の組成物を前記患者に有効量投与することを含むことを特徴とする方法。

【請求項 3 5】

請求項 1 0 の抗体の特異性を有するポリクローナル抗体を作製する方法であって、

- (a) 抗体反応が起こる条件下で、S E Q I D N O : 1 - 4 7 からなる群から選択されたアミノ酸配列またはその免疫原性断片を有するポリペプチドで動物を免疫化するステップと、
- (b) 前記動物から抗体を単離するステップと、
- (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て、S E Q I D N O : 1 - 4 7 からなる群から選択されたアミノ酸配列を有するポリペプチドに特異的に結合するポリクローナル抗体を同定するステップとを含むことを特徴とするポリクローナル抗体作製方法。

30

【請求項 3 6】

請求項 3 5 の方法により作製された抗体。

【請求項 3 7】

請求項 3 6 の抗体及び適当な担体を含む組成物。

【請求項 3 8】

請求項 1 0 の抗体の特異性を有するモノクローナル抗体を作製する方法であって、

- (a) 抗体反応が起こる条件下で、S E Q I D N O : 1 - 4 7 からなる群から選択されたアミノ酸配列またはその免疫原性断片を有するポリペプチドで動物を免疫化するステップと、
- (b) 前記動物から抗体を産生する細胞を単離するステップと、
- (c) 前記抗体産生細胞を不死化細胞と融合し、モノクローナル抗体産生ハイブリドーマ細胞を形成するステップと、
- (d) 前記ハイブリドーマ細胞を培養するステップと、
- (e) S E Q I D N O : 1 - 4 7 からなる群から選択されたアミノ酸配列を有するポリペプチドに特異的に結合する前記培養モノクローナル抗体から単離するステップとを含むことを特徴とするモノクローナル抗体作製方法。

40

【請求項 3 9】

50

請求項 38 の方法で作製したモノクローナル抗体。

【請求項 40】

請求項 39 の抗体及び適当な担体を含む組成物。

【請求項 41】

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

【請求項 42】

組換え免疫グロブリンライブラリのスクリーニングにより得られた請求項 10 の抗体。

【請求項 43】

サンプルにおける S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドを検出する方法であって、

10

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

(b) 特異的な結合を検出するステップとを含み、

前記特異的な結合が、前記サンプルに S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドが存在することを示唆することを特徴とする方法。

【請求項 44】

S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドをサンプルから精製する方法であって、

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

20

(b) 前記サンプルから前記抗体を分離して、S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸配列を有する精製されたポリペプチドを得るステップとを含むことを特徴とする方法。

【請求項 45】

S E Q I D N O : 1 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 46】

S E Q I D N O : 2 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 47】

S E Q I D N O : 3 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

30

【請求項 48】

S E Q I D N O : 4 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 49】

S E Q I D N O : 5 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 50】

S E Q I D N O : 6 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 51】

S E Q I D N O : 7 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 52】

S E Q I D N O : 8 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

40

【請求項 53】

S E Q I D N O : 9 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 54】

S E Q I D N O : 10 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 55】

S E Q I D N O : 11 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 56】

S E Q I D N O : 12 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 57】

S E Q I D N O : 13 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

50

【請求項 58】

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

【請求項 59】

SEQ ID NO : 15 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 60】

SEQ ID NO : 16 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 61】

SEQ ID NO : 17 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 62】

SEQ ID NO : 18 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

10

【請求項 63】

SEQ ID NO : 19 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 64】

SEQ ID NO : 20 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 65】

SEQ ID NO : 21 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 66】

SEQ ID NO : 22 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 67】

SEQ ID NO : 23 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

20

【請求項 68】

SEQ ID NO : 24 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 69】

SEQ ID NO : 25 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 70】

SEQ ID NO : 26 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 71】

SEQ ID NO : 27 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 72】

SEQ ID NO : 28 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

30

【請求項 73】

SEQ ID NO : 29 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 74】

SEQ ID NO : 30 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 75】

SEQ ID NO : 31 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 76】

SEQ ID NO : 32 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 77】

SEQ ID NO : 33 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

40

【請求項 78】

SEQ ID NO : 34 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 79】

SEQ ID NO : 35 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 80】

SEQ ID NO : 36 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 81】

SEQ ID NO : 37 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 82】

SEQ ID NO : 38 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

50

【請求項 83】

SEQ ID NO : 39 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 84】

SEQ ID NO : 40 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 85】

SEQ ID NO : 41 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 86】

SEQ ID NO : 42 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 87】

SEQ ID NO : 43 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 88】

SEQ ID NO : 44 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 89】

SEQ ID NO : 45 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 90】

SEQ ID NO : 46 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 91】

SEQ ID NO : 47 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 92】

SEQ ID NO : 48 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 93】

SEQ ID NO : 49 のポリヌクレオチド配列を含む請求項 11 のポリヌクレオチド。

【請求項 94】

SEQ ID NO : 50 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 95】

SEQ ID NO : 51 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 96】

SEQ ID NO : 52 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 97】

SEQ ID NO : 53 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 98】

SEQ ID NO : 54 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 99】

SEQ ID NO : 55 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 100】

SEQ ID NO : 56 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 101】

SEQ ID NO : 57 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 102】

SEQ ID NO : 58 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 103】

10

20

30

40

50

SEQ ID NO : 59 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 104】

SEQ ID NO : 60 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 105】

SEQ ID NO : 61 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 106】

SEQ ID NO : 62 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。 10

【請求項 107】

SEQ ID NO : 63 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 108】

SEQ ID NO : 64 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 109】

SEQ ID NO : 65 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。 20

【請求項 110】

SEQ ID NO : 66 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 111】

SEQ ID NO : 67 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 112】

SEQ ID NO : 68 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 113】

SEQ ID NO : 69 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。 30

【請求項 114】

SEQ ID NO : 70 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 115】

SEQ ID NO : 71 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 116】

SEQ ID NO : 72 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。 40

【請求項 117】

SEQ ID NO : 73 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 118】

SEQ ID NO : 74 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 119】

SEQ ID NO : 75 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。 50

【請求項 1 2 0】

SEQ ID NO : 76 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 2 1】

SEQ ID NO : 77 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 2 2】

SEQ ID NO : 78 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 2 3】

SEQ ID NO : 79 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 2 4】

SEQ ID NO : 80 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 2 5】

SEQ ID NO : 81 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 2 6】

SEQ ID NO : 82 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 2 7】

SEQ ID NO : 83 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 2 8】

SEQ ID NO : 84 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 2 9】

SEQ ID NO : 85 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 3 0】

SEQ ID NO : 86 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 3 1】

SEQ ID NO : 87 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 3 2】

SEQ ID NO : 88 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 3 3】

SEQ ID NO : 89 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 3 4】

SEQ ID NO : 90 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 3 5】

SEQ ID NO : 91 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 1 3 6】

SEQ ID NO : 92 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオ

10

20

30

40

50

チド。

【請求項 137】

SEQ ID NO: 93 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 138】

SEQ ID NO: 94 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【発明の詳細な説明】

【0001】

(技術分野)

本発明は、RNA 代謝タンパク質の核酸配列及びアミノ酸配列に関する。本発明はまた、これらの配列を利用した神経系の疾患、自己免疫/炎症性の疾患、細胞増殖異常、及び発生または発達障害の診断・治療・予防に関する。本発明は更に、RNA 代謝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関する。

【0002】

(発明の背景)

リボ核酸 (RNA) は、ATP 及び CTP、UTP、GTP の 4 つのヌクレオチドから成る直線状の一本鎖重合体である。大抵の生物では、RNA は生物の遺伝物質であるデオキシリボ核酸 (DNA) の複製物として転写される。レトロウイルスでは、DNA ではなく RNA が遺伝物質として働く。遺伝物質である RNA 複製物はタンパク質をコードしたり、生物における種々の構造的、触媒的、または調節的な役割を果たす。RNA はその細胞内の局在性及び機能によって分類される。メッセンジャー RNA (mRNA) はポリペプチドをコードする。リボソーム RNA (rRNA) はリボソームタンパク質と共に、mRNA をポリペプチドに翻訳する細胞質粒子であるリボソームを構成する。トランスファー RNA (tRNA) は、mRNA コドンとそのコドンに一致するアミノ酸の両方を認識して mRNA の翻訳において機能する細胞質アダプター分子である。ヘテロ核 RNA (hnRNA) には、mRNA 前駆体やその他の様々な大きさの核 RNA が含まれる。核内低分子 RNA (snRNA) は、非コード介在配列を排除して、mRNA 前駆体のエキソンを再結合する核のスプライセオソーム複合体の一部である。

【0003】

DNA からの転写、RNA プロセッシング、及び mRNA のタンパク質への翻訳の間、タンパク質が RNA に会合している。構造的、触媒的、または調節的な目的で用いられる時もタンパク質が RNA に会合している。

【0004】

RNA プロセッシング

様々なタンパク質が、核における転写 RNA のプロセッシングに必要である。mRNA 前駆体のプロセッシングステップには、メチルグアノシンでの 5' 末端のキャッピングや 3' 末端のポリアデニル化、及びイントロンを除去するスプライシングが含まれる。スプライセオソーム複合体は、U1 及び U2、U4、U5、U6 と呼ぶ 5 つの低分子核内リボ核タンパク粒子 (snRNP) からなる。それぞれの snRNP には、一種類の snRNA と約 10 のタンパク質が含まれる。snRNA の中には、その RNA 成分がイントロンコンセンサス配列を認識してその配列と塩基対を形成するものもある。snRNA のタンパク質成分が、スプライセオソームの構築及びスプライシング反応を仲介する。

【0005】

mRNA 前駆体の切断の初期段階には、切断因子 Im (CF Im) が関係する。ヒト CF Im タンパク質は、3' 末端プロセッシング複合体を構成するプロセッシング因子の補充及び構築を助ける (Rueggsegger, U. 他 (1998) Mol. Cell. 1: 243 - 253)。マウスホルミン結合タンパク質 (FBP) である FBP11 及び FBP12 は、イントロンの 5' 及び 3' 末端の架橋を促す mRNA 前駆体スプライシング複合体の成分である。これらのタンパク質は、U1 snRNP 及び U2 snRN

10

20

30

40

50

Pを伴う架橋相互作用において機能する。snRNPタンパク質に対する自己抗体が、全身性エリテマトーデスを患っている患者の血中に見られる (Stryer, L. (1995) Biochemistry W. H. Freeman and Company, New York NY, p. 863)。

【0006】

ヘテロ核内リボ核タンパク質 (hnRNP: heterogeneous nuclear ribonucleoprotein) は、スプライシング、成熟RNAの細胞質への輸送、及びmRNAの翻訳において機能することが確認された (Biamonti, G. 他 (1998) Chin. Exp. Rheumatol. 16: 317 - 326)。hnRNPの例には、RNAの3'末端の切断及びポリアデニル化に關与する酵母タンパク質Hrp1p、RNAの5'末端のキャッピングに關与するCbp80p、核からのmRNAの輸送に關与する哺乳動物hnRNP A1の相同体であるNp13pがある (Shen, E. C. 他 (1998) Genes Dev. 12: 679 - 691)。hnRNPは、リウマチ性疾患における自己免疫反応の重要な標的であることが分かった (Biamonti, 前出)。

【0007】

多くのsnRNP及びhnRNPは、RNA認識モチーフ (RRM) によって特徴付けられる (Birney, E. 他 (1993) Nucleic Acids Res. 21: 5803 - 5816を参照)。RRMはアミノ酸約80個の長さであり、4つの鎖及び2つのヘリックスからなる - サンドイッチ構造を形成している。RRMは、コアRNP - 1オクタペプチドモチーフと周囲の保存された配列を含む。snRNPタンパク質に加えて、上記のモチーフを含むRNA結合タンパク質の例には、選択的スプライシングを調節する因子や新生RNAを安定化させる異核リボ核タンパク質 (heteronuclear ribonucleoprotein) がある。選択的スプライシング因子には、発生的に調節されるタンパク質があり、この具体的な例がショウジョウバエ (*Drosophila melanogaster*) 及び線虫 (*Caenorhabditis elegans*) などの下等真核生物で同定された。これらのタンパク質は、パターン形成及び性決定などのそれぞれの発生プロセスにおいて重要な役割を果たす (例えば、Hodgkin, J. 他 (1994) Development 120: 3681 - 3689を参照)。

【0008】

RNAの安定化及び分解

RNAヘリカーゼは、ATP加水分解によるエネルギーを用いてRNAの高次構造及び二次構造を変化させ調節し、二重鎖RNAを不安定にして巻き戻す。最もよく特徴付けられた遍在性のRNAヘリカーゼファミリーには、DEADボックスファミリーがある。このDEADボックスファミリーという名称は、このファミリーのタンパク質の特徴である保存されたB型ATP結合モチーフからついた。40を超えるDEADボックス型ヘリカーゼが、細菌から昆虫、酵母、両生類、哺乳類、及び植物に至る多様な生物において同定された。DEADボックス型ヘリカーゼは、翻訳開始、スプライシング、リボソーム形成、並びにRNAの編集、輸送、及び安定化などの多様なプロセスにおいて機能する。或るDEADボックス型ヘリカーゼは、精子形成及び胚形成において組織特異的及び段階特異的に役割を果たす。全てのDEADボックス型ヘリカーゼは、約420のアミノ酸上に分散した幾つかの保存された配列モチーフを含む。これらのモチーフには、A型ATP結合モチーフ、DEADボックス/B型ATP結合モチーフ、機能が既知のセリン/アルギニン/トレオニントリペプチド、並びに基質結合及び巻き戻しにおいて役割を果たす可能性のあるC末端グリシンリッチモチーフが含まれる。更に、異なるDEADボックス型ヘリカーゼ配列間のアラインメントから、37のアミノ酸残基がこれらの配列間で一致していることが分かり、これらの残基の保存がヘリカーゼ機能に重要であると思われる (Linder, P. 他 (1989) Nature 337: 121 - 122を参照)。

【0009】

DEADボックス1タンパク質(DDX1)の過剰な発現が、神経芽細胞腫(Nb)及び網膜芽細胞腫(Rb)の進行に役割を果たし得る。これらの所見から、DDX1が、癌細胞のRNAの正常な二次構造及び発現レベルを変えて癌の進行を促進すると思われる。他のDEADボックス型ヘリカーゼは、紫外線誘発性の癌、B細胞リンパ腫、及び骨髄悪性腫瘍に直接或いは間接的に関与し得る(Godbout, R.ら(1998) J. Biol. Chem. 273: 21161-21168を参照)。

【0010】

リボヌクレアーゼ(RNアーゼ)は、RNA鎖のホスホジエステル結合の加水分解を触媒してRNAを切断する。例えば、RNアーゼPは、成熟プロセスの一部としてtRNA前駆体の5'末端を切断するリボ核タンパク質酵素である。RNアーゼHは、RNA/DNAハイブリッドのRNA鎖を消化する。このようなハイブリッドはレトロウイルスが侵入した細胞で生じ、RNアーゼHはレトロウイルス複製サイクルにおける重要な酵素である。RNアーゼHドメインは、逆転写酵素に関連するドメインとして見出される場合が多い。血清及び細胞抽出物におけるRNアーゼ活性が、様々な癌や感染症において上昇する(Schein, C.H. (1997) Nat Biotechnol. 15: 529-536)。RNアーゼ活性の調節が、腫瘍血管形成、アレルギー反応、ウイルス感染、ウイルス複製、及び真菌感染を制御するための手段として研究されている。

10

【0011】

未成熟な終止コドン即ちナンセンスコドンを有するmRNAの分解は、ナンセンス仲介性mRNA崩壊(NMD: nonsense-mediated mRNA decay)と呼ばれる監視機構によって行われる。この機構は、機能しないポリペプチドや有害なポリペプチドをコードし得る欠陥mRNAの排除を助ける。様々なNMD成分が、酵母RNA代謝障害及びヒトRNA代謝障害の両方に関係する(Hentze, M. and Kulozik, A. (1999) Cell 96: 307-310)。

20

【0012】

翻訳

リボソームRNA(rRNA)はリボソームタンパク質と共にリボソームを構成する。リボソームは、メッセンジャーRNA(mRNA)をポリペプチドに翻訳する細胞質粒子である。真核細胞のリボソームは、60Sサブユニット(大サブユニット)と40Sサブユニット(小サブユニット)からなる80Sの粒子である。リボソームは18SrRNA、28SrRNA、5SrRNA、及び5.8SrRNAに加えて、生物によって異なるが50種類から80種類のタンパク質を含む。リボソームタンパク質は、属するサブユニットによって分類される(すなわち、60Sの大サブユニットを構成するものはL、40Sの小サブユニットを構成するものはS)。大腸菌のリボソームが最も研究され、50のタンパク質を含み、その多くが全ての生活型に保存されていることが分かっている。9つのリボソームタンパク質(すなわち、S5、S6、S17、L1、L6、L9、L12、L14、及びL30)の構造が 3×10^{-8} cm(3.0オングストローム)未満の分解能で解析され、鎖間に位置する酸性及び塩基性のRNA結合モチーフに加えて、タンパク質フォールドなどの共通モチーフを含むことが明らかになった(Liljas, A. and Garber, M. (1995) Curr. Opin. Struct. Biol. 5: 721-727, または Woodson, S. A. and Leontis, N. B. (1998) Curr. Opin. Struct. Biol. 8: 294-300、Ramakrishnan, V. and White, S. W. (1998) Trends Biochem. Sci. 23: 208-212を参照)。

30

40

【0013】

リボソームタンパク質は翻訳後修飾を受けたり、他のリボソーム関連タンパク質と相互作用して翻訳を調節し得る。例えば、相同性の高い40Sリボソームタンパク質であるS6キナーゼ(S6K1及びS6K2)は、リボソームタンパク質を含むタンパク質合成装置を形成する翻訳成分の生合成を調節することで、細胞成長の調節において重要な役割を果

50

たす。S 6 K 1 の場合、少なくとも 8 個のリン酸化部位が階層的にキナーゼの活性化を仲介すると考えられている (D u f n e r a n d T h o m a s (1 9 9 9) E x p . C e l l . R e s . 2 5 3 : 1 0 0 - 1 0 9) 。また、L 1 を含むリボソームタンパク質の中には、リボソームタンパク質をコードする多シストロン性メッセンジャー R N A に結合して翻訳リプレッサーとして機能するものもある (L i l j a s , A . 前出、及び G a r b e r , M . 前出を参照) 。

【 0 0 1 4 】

近年の研究により、幾つかのリボソームタンパク質が、タンパク質合成における役割とは別に第 2 の機能を有するということが報告された。これらのタンパク質は、細胞増殖の調節因子として、また場合によっては細胞死の誘導物質としても機能する。例えば、ヒトリボソームタンパク質 L 1 3 a の発現が、細胞周期の G 2 / M 期における細胞成長を停止させてアポトーシスを誘導することが分かった。また、L 1 3 a の発現を抑制すると標的細胞におけるアポトーシスが誘導されることから、細胞の生存にはこのタンパク質が適量必要であると考えられる。同様の結果が酵母でも得られた。すなわち、L 1 3 a 、 r p 2 2 、及び r p 2 3 の酵母相同体を不活性化すると、著しい成長の遅れ及び細胞死が起こった。密接に関連するリボソームタンパク質 L 7 は G 1 期において細胞を停止させ、アポトーシスを誘導する。従って、リボソームタンパク質のサブセットは、細胞周期チェックポイントとして働き、細胞増殖調節因子の新規のファミリーを構成すると思われる。

10

【 0 0 1 5 】

特定のリボソームタンパク質に対する抗体の上昇が視覚化される 3 D 免疫低温電子顕微鏡法 (i m m u n o c r y o e l e c t r o n m i c r o s c o p y) を用いて、無傷のリボソームの表面上における個々のリボソームタンパク質をマッピングした。L 1 、L 7 、及び L 1 2 のマッピングは進んだが、無傷のリボソームの構造が $2 \sim 2.5 \times 10^{-7}$ c m (2 0 - 2 5 オングストローム) の低い分解能でしか解析されていないため、異なった粗構造間に不整合が存在する (F r a n k , J . (1 9 9 7) C u r r . O p i n . S t r u c t . B i o l . 7 : 2 6 6 - 2 7 2) 。

20

【 0 0 1 6 】

リボソームにおいて 3 つの特徴的な部位が同定されている。アミノアシル - t R N A 受容部位 (A 部位) が、開始 t R N A を除く荷電 t R N A を受け取る。A 部位からアミノ酸が細長い鎖に付加されつ時に、新生ポリペプチドがペプチジル - t R N A 部位 (P 部位) に結合する。脱アシル化した t R N A が、リボソームから離脱する前に出口部位 (E 部位) に結合する。リボソームの構造については、S t r y e r , L . (1 9 9 5) B i o c h e m i s t r y W . H . F r e e m a n a n d C o m p a n y , N e w Y o r k N Y p p . 8 8 8 - 9 0 8 1 、L o d i s h , H . ら (1 9 9 5) M o l e c u l a r C e l l B i o l o g y S c i e n t i f i c A m e r i c a n B o o k s , N e w Y o r k N Y p p . 1 1 9 - 1 3 8 、及び L e w i n , B (1 9 9 7) G e n e s V I O x f o r d U n i v e r s i t y P r e s s , I n c . N e w Y o r k , N Y を参照されたい。

30

【 0 0 1 7 】

t R N A の荷電

遺伝子コードの正確な翻訳は、適切な転移 R N A (t R N A) との結合を形成するアミノ酸による。アミノアシル - t R N A シンターゼ (a a R S) は全ての生物に見られる必須タンパク質である。a a R S は、タンパク質の生合成の第 1 のステップとして、アミノ酸を活性化し、あるアミノ酸とそれに対応する t R N A との正しい結合を担う。原核生物は少なくとも 2 0 種類の a a R S を有し、その 1 つずつが異なったアミノ酸に対応する。一方、真核生物は、通常、細胞質型とミトコンドリア型の 2 種類の a a R S を有し、それぞれが異なったアミノ酸に対応する。2 0 種類の a a R S 酵素は、構造により 2 つのクラスに分類することができる。クラス I 酵素はアミノ酸を t R N A の 3 ' 末端における 2 ' ヒドロキシルに付加し、一方、クラス II 酵素はアミノ酸を t R N A の 3 ' 末端における 3 ' ヒドロキシルに付加する。それぞれのクラスは、触媒ドメインの固有のトポロジによ

40

50

り特徴付けられている。クラス I 酵素はヌクレオチド結合ロスマンフォールドに基づいて触媒ドメインを含む。具体的には、コンセンサストラペプチドモチーフが高度に保存されている (Prosite Document PDOC00161、アミノアシル転移 RNA シンテターゼクラス I シグネチャ)。クラス II 酵素は、7 本鎖逆平行 シートドメインからなる中央の触媒ドメイン、N 末端調節ドメイン、及び C 末端調節ドメインを含む。クラス II 酵素は、そのヘテロ二量体構造またはホモ二量体構造に基づいて 2 つのグループに分けられる。後者のグループは更に、N 末端調節ドメイン及び C 末端調節ドメインの構造により細分できる (Hartle in, M. and Cusack, S. (1995) J. Mol. Evol. 40: 519-530)。異なった aaRS の存在は、多様な進化の結果であり、遺伝子複製時に生じたと思われる。注目すべきは、グルタミン (Gln) などのアミノ酸は最後に出現したアミノ酸であり、進化の研究から Gln-RS が真核生物に初めに出現し、後に原核生物に水平に移入されたと思われる (Lamour, V. ら (1994) Proc. Natl. Acad. Sci. U. S. A. 91: 8670-74 and Siatecka, M. ら (1998) Eur. J. Biochem. 256: 80-7)。Gln-RS 及び Gln-tRNA^{Gln} の重要性を後述する。

【0018】

特定の アミノアシル-tRNA シンテターゼは、タンパク質の生合成の役割に加えて、細胞の忠実性、RNA スプライシング、RNA 輸送、アポトーシス、並びに転写及び翻訳の調節に役割を果たす。例えば、ヒトチロシル-tRNA シンテターゼは、異なるサイトカイン活性を有する 2 つの断片に蛋白分解的に切断することができる。カルボキシ末端ドメインは、単球活性及び白血球走化性活性を有し、ミエロペルオキシダーゼ、腫瘍壊死因子、及び組織因子の産生を刺激する。N 末端ドメインは、インターロイキン-8 A 型受容体に結合し、インターロイキン-8 様サイトカインとして機能する。ヒトチロシル-tRNA シンテターゼは、アポトーシス腫瘍細胞から分泌され、アポトーシスを加速させ得る (Wakasugi, K. and Schimmel, P. (1999) Science 284: 147-151)。ミトコンドリア Neurospora crassa TyrRS 及び S. cerevisiae LeuRS は、特定のグループイントロンスプライシング活性に必須の因子であり、ヒトミトコンドリア LeuRS は、酵母 null 株 (yeast null strain) において酵母 LeuRS と置換できる。幾つかの aaRS は、細胞増殖、分化、及びアポトーシスにおいて機能するシグナル伝達分子のクラスである diadenosine oligophosphate を合成することができる (Kisselev, L. L ら (1998) FEBS Lett. 427: 157-163; Vartanian, A. ら (1999) FEBS Lett. 456: 175-180)。

【0019】

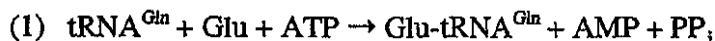
最適な条件下では、1 秒当たり約 40 個のアミノ酸の速度でポリペプチド合成が進行する。翻訳中に誤ったアミノ酸が組み込まれる割合は 10^{-4} のオーダーであり、主にアミノアシル-tRNA が正しくないアミノ酸で荷電されて生じる。不適切に荷電した tRNA は、伸長しているポリペプチドに正しくないアミノ酸が組み込まれることになるため、細胞にとって有害である。翻訳の速度は、最適な伸長速度と翻訳の忠実性との双方の妥協によるものと思われる。数学的な計算から、 10^{-4} が生物系におけるタンパク質の生合成のエラーとして許容できる最大値であると推定される (Stryer, L. 前出, 及び Watson, J. ら (1987) The Benjamin/Cummings Publishing Co., Inc. Menlo Park, CA)。特にエラーが生じやすいのは、グルタミンで tRNA^{Gln} を荷電する場合のアミノアシル-tRNA の荷電である。誤った荷電を修正する機構が存在するが、これは進化中に生じたものと思われる。グルタミンは、ポリペプチド合成に用いられる 20 の天然のアミノ酸の中で最後に自然界に出現したアミノ酸である。グラム陽性の真正細菌、ラン藻類、Archaea、及び真核生物細胞小器官は、酵素 Glu-tRNA^{Gln} アミドトランスフェラー

ゼ (Glu - AdT) を用いる Glu - tRNA^{Gln} (Glu - tRNA シンテターゼ 即ち GluRS により合成される) の変換に基づいて Gln - tRNA^{Gln} の合成のための非正準経路を有する。アミド基転移経路に関与する反応は以下のとおりである (Curnow, A. W. ら (1997) Nucleic Acids Symposium 36: 2 - 4)。

【0020】

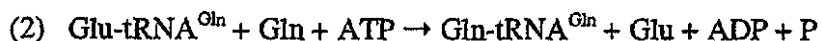
【化1】

GluRS



10

Glu-AdT



類似の酵素である Asp - tRNA^{Asn} アミドトランスフェラーゼは Archaea に存在し、Asp - tRNA^{Asn} を Asn - tRNA^{Asn} に変換する。ホルミラーゼは、真性細菌において Met - tRNA^{fMet} を fMet - tRNA^{fMet} に変換する酵素であり、関連酵素であると思われる。加水分解活性もまた、誤って荷電した Val - tRNA^{Ile} を破壊することが確認された (Schimmel, P. ら (1998) FASEB J. 12: 1599 - 1609)。原始生活型における Glu - AdT の進化は、特定のグルタミル - tRNA シンテターゼ (GlnRS) の非存在により Gln - tRNA^{Gln} の合成の代替経路が必要になったことによるとと思われる。事実、グラム陽性細菌における Glu - AdT オペロンの枯渇は致命的である (Curnow, A. W. ら (1997) Proc. Natl. Acad. Sci. U. S. A. 94: 11819 - 11826)。GluRS が天然の翻訳装置において高度に保存されていることから他の生物における GluRS 活性の存在が推定されるが、GluRS はヒトを含む全ての生物において同定されたというわけではない。このような酵素が、翻訳の忠実性を高めること、及び欠陥ポリペプチドの合成を減少させることに関与しているであろう。

20

30

【0021】

アミノアシル - tRNA に対する自己抗体が、リュウマチ様関節炎、皮膚筋炎、及び多発性筋炎などの自己免疫疾患の患者で産生され、間質性肺疾患 (ILD) 合併症と密接に関連する (Freist, W. ら (1999) Biol. Chem. 380: 623 - 646; Freist, W. ら (1996) Biol. Chem. Hoppe Seyler 377: 343 - 356)。これらの抗体がウイルス感染にตอบสนองして産生されることが分かったため、コクサッキーウイルスを用いて動物において実験的ウイルス筋炎を発症させた。

【0022】

ヒトと病原体との間の aaRS 構造の比較は、新規の抗生物質のデザインに有用である (Schimmel, 前出)。遺伝子組換え aaRS を用いることにより、in vitro でのタンパク質への非天然のアミノ酸の部位特異的組み込みが可能となる (Liu, D. R. ら (1997) Proc. Natl. Acad. Sci. USA 94: 10092 - 10097)。

40

【0023】

翻訳開始

翻訳の開始は3つの段階に分けることができる。第1の段階では、開始転移RNA (Met - tRNA_f) が40Sリボソームサブユニットと共に43S開始前複合体を形成する。第2の段階で、43S開始前複合体がmRNAに結合し、この後この複合体が正しいAUG開始コドンに移動する。第3の段階では、60Sリボソームサブユニットが40Sサ

50

ブユニットに結合し、開始コドンにおいて80Sリボソームを形成する。転写の調節は、主に開始プロセスにおける第1の段階及び第2の段階で行われる(V. M. Pain (1996) Eur. J. Biochem. 236: 747-771)。

【0024】

幾つかの開始因子は、その多くが複数のサブユニットと含み、開始tRNAと40Sリボソームサブユニットとの結合に関与する。ある真核生物の開始因子(eIF) eIF5Aは、ユニークなアミノ酸残基であるhypusine(N-(4-amino-2-hydroxybutyl)lysine)を含む18kDのタンパク質である(Rinaudo, M. ら (1993) Gene 137: 303-307)。グアニンヌクレオチド結合タンパク質eIF2は、開始tRNAを40Sリボソームサブユニットに結合させる。eIF2がGTPに会合している時にのみ開始tRNAに結合する。グアニンヌクレオチド交換タンパク質eIF2Bは、eIF2をGDP結合不活性型からGTP結合活性型に変換する。他の2つの因子であるeIF1A及びeIF3が、18SリボソームRNA及び特定のリボソーム構造タンパク質と相互作用することで、40Sサブユニットに結合して安定化させる。eIF3はまた、40SリボソームサブユニットとmRNAとの会合に関与する。Met-tRNA_f、eIF1A、eIF3、及び40Sリボソームサブユニットから43S開始前複合体が形成される(Pain, 前出)。

10

【0025】

43S開始前複合体とmRNA分子との結合には、更なる因子が必要であり、このプロセスは幾つかのレベルで調節される。eIF4Fは、eIF4E、eIF4A、及びeIF4Gの3つのタンパク質からなる複合体である。eIF4EはmRNA 5'末端m⁷GTPキャップを認識してそれに結合し、eIF4Aは双方向性RNA依存性ヘリカーゼであり、eIF4Gは足場ポリペプチド(scaffolding polypeptide)である。eIF4Gは3つの結合ドメインを有する。eIF4GのN末端側の3分の1がeIF4Eと相互作用し、真中の3分の1がeIF4Aに結合し、C末端側の3分の1が43S開始前複合体と結合したeIF3と結合する。従って、eIF4Gが、40SリボソームサブユニットとmRNAとの間の橋として働く(M. W. Hentze (1997) Science 275: 500-501)。

20

【0026】

eIF4Fの43S開始前複合体を結合を開始させる能力は、mRNAの構造的な特性により調節される。mRNA分子は、5'キャップとAUG開始コドンとの間に非翻訳領域(UTR)を有する。mRNAの中には、この領域が43S開始前複合体の結合を妨げる二次構造を形成しているものもある。eIF4Aのヘリカーゼ活性が、この二次構造を排除して43S開始前複合体の結合を促進すると思われる(Pain, 前出)。

30

【0027】

翻訳伸長

伸長は、開始メチオニンにアミノ酸を次々に付加していき、完全長のポリペプチド鎖を形成するプロセスである。伸長因子EF1、EF1、及びEF2は、開始後のポリペプチド鎖の伸長に関与する。EF1はGTP結合タンパク質である。EF1のGTP結合型は、アミノアシル-tRNAをリボソームのA部位に結合させる。新しく到着したアミノアシル-tRNAに結合したアミノ酸が開始メチオニンとペプチド結合する。EF1上のGTPがGDPに加水分解され、EF1-GDPがリボソームから離脱する。EF1がEF1-GDPに結合してGDPをEF1から解離させることにより、EF1のGTPへの結合及び次のサイクルが可能となる。

40

【0028】

続くアミノアシル-tRNAがリボソームに輸送され、別のGTP結合タンパク質であるEF-Gが、tRNAのリボソームのA部位からP部位、さらに最終的にはE部位へのトランスローケーションを触媒する。

【0029】

翻訳停止

50

終結因子 e R F は翻訳を停止させる。e R F は m R N A 上の停止コドンを認識して、リボソームからポリペプチド鎖を離脱させる。

【0030】

新規の R N A 代謝タンパク質、およびそれらをコードするポリヌクレオチドの発見により、新規の組成物を提供することで当分野の要望に応えることができる。この新規の組成物は、神経系の疾患、自己免疫/炎症性の疾患、細胞増殖異常、及び発生または発達障害の診断・治療・予防において有用であり、また、R N A 代謝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。

【0031】

(発明の要約)

本発明は、総称して「R M E P」、個別にはそれぞれ「R M E P - 1」、「R M E P - 2」、「R M E P - 3」、「R M E P - 4」、「R M E P - 5」、「R M E P - 6」、「R M E P - 7」、「R M E P - 8」、「R M E P - 9」、「R M E P - 10」、「R M E P - 11」、「R M E P - 12」、「R M E P - 13」、「R M E P - 14」、「R M E P - 15」、「R M E P - 16」、「R M E P - 17」、「R M E P - 18」、「R M E P - 19」、「R M E P - 20」、「R M E P - 21」、「R M E P - 22」、「R M E P - 23」、「R M E P - 24」、「R M E P - 25」、「R M E P - 26」、「R M E P - 27」、「R M E P - 28」、「R M E P - 29」、「R M E P - 30」、「R M E P - 31」、「R M E P - 32」、「R M E P - 33」、「R M E P - 34」、「R M E P - 35」、「R M E P - 36」、「R M E P - 37」、「R M E P - 38」、「R M E P - 39」、「R M E P - 40」、「R M E P - 41」、「R M E P - 42」、「R M E P - 43」、「R M E P - 44」、「R M E P - 45」、「R M E P - 46」、及び「R M E P - 47」と呼ぶ R N A 代謝タンパク質である精製されたポリペプチドを提供する。本発明の一実施態様では、(a) S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸と少なくとも 90 % の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) S E Q I D N O : 1 - 47 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単離されたポリペプチドを提供する。別法では、S E Q I D N O : 1 - 47 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。

【0032】

更に本発明は、(a) S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸と少なくとも 90 % の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドをコードする単離されたポリヌクレオチドを提供する。別法では、このポリヌクレオチドは、S E Q I D N O : 1 - 47 からなる群から選択されたポリペプチドをコードする。更なる別法では、このポリヌクレオチドは、S E Q I D N O : 48 - 94 からなる群から選択される。

【0033】

更に、本発明は、(a) S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸と少なくとも 90 % の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) S E Q I D N O : 1 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドをコードするポリヌクレオチドと機能的に結合されたプロモータ

10

20

30

40

50

ー配列を含む組換えポリヌクレオチドを提供する。別法では、本発明は、この組換えポリヌクレオチドで形質転換された細胞を提供する。更なる別法では、本発明は、この組換えポリヌクレオチドを含む遺伝子組換え生物を提供する。

【0034】

また、本発明は、(a) SEQ ID NO: 1 - 47 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1 - 47 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1 - 47 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 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドの生産方法を提供する。この方法は、(a) このポリペプチドの発現に好適な条件下で、このポリペプチドをコードするポリヌクレオチドと機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養するステップと、(b) このように発現したポリペプチドを回収するステップとを含む。

10

【0035】

更に、本発明は(a) SEQ ID NO: 1 - 47 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1 - 47 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1 - 47 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 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドに特異的に結合する単離された抗体を提供する。

20

【0036】

更に、本発明は、(a) SEQ ID NO: 48 - 94 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(b) SEQ ID NO: 48 - 94 からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(c) 前記(a)のポリヌクレオチドに相補的なポリヌクレオチドと、(d) 前記(b)のポリヌクレオチドに相補的なポリヌクレオチドと、(e) 前記(a) (d)のRNA等価物とで構成される群から選択された単離されたポリヌクレオチドを提供する。別法では、このポリヌクレオチドは、少なくとも60個の連続するヌクレオチドを含む。

30

【0037】

更に本発明は、(a) SEQ ID NO: 48 - 94 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(b) SEQ ID NO: 48 - 94 からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(c) 前記(a)のポリヌクレオチドに相補的なポリヌクレオチドと、(d) 前記(b)のポリヌクレオチドに相補的なポリヌクレオチドと、(e) 前記(a) (d)のRNA等価物とで構成される群から選択されたポリヌクレオチド配列を有するサンプルにおける標的ポリヌクレオチドを検出する方法を提供する。この方法は、(a) 前記サンプル内の標的ポリヌクレオチドと相補的な配列を含む少なくとも20個の連続するヌクレオチドを含むプローブと前記サンプルをハイブリダイズさせるステップであって、前記プローブと前記標的ポリヌクレオチドまたはその断片とでハイブリダイゼーション複合体が形成される条件下で、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズする、該ステップと、(b) 前記ハイブリダイゼーション複合体の存在するか否かを検出し、存在する場合には必要に応じてその収量を測定するステップとを含む。別法では、前記プローブは、少なくとも60個の連続するヌクレオチドを含む。

40

【0038】

更に本発明は、(a) SEQ ID NO: 48 - 94 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(b) SEQ ID NO: 48 - 94 からなる

50

群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(c)前記(a)のポリヌクレオチドに相補的なポリヌクレオチドと、(d)前記(b)のポリヌクレオチドに相補的なポリヌクレオチドと、(e)前記(a)(d)のRNA等価物とで構成される群から選択されたポリヌクレオチド配列を有するサンプルにおける標的ポリヌクレオチドを検出する方法を提供する。この方法は、(a)ポリメラーゼ連鎖反応増幅法を用いて、前記標的ポリヌクレオチドまたはその断片を増幅するステップと、(b)増幅された前記標的ポリヌクレオチドまたはその断片が存在するか否かを検出し、存在する場合には必要に応じてその収量を測定するステップとを含む。

【0039】

更に本発明は、(a)SEQ ID NO: 1-47からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b)SEQ ID NO: 1-47からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c)SEQ ID NO: 1-47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: 1-47からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択された有効量のポリペプチド及び好適な医薬用賦形剤を含む組成物を提供する。一実施例では、SEQ ID NO: 1-47からなる群から選択されたアミノ酸配列を含む組成物を提供する。本発明は更に、患者にこの組成物を投与することを含む、機能的RMEP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。

【0040】

更に本発明は、(a)SEQ ID NO: 1-47からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b)SEQ ID NO: 1-47からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c)SEQ ID NO: 1-47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: 1-47からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドのアゴニストとして有効な化合物をスクリーニングする方法を提供する。この方法は、(a)このポリペプチドを含むサンプルを化合物に曝露するステップと、(b)このサンプルのアゴニスト活性を検出するステップとを含む。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RMEP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。

【0041】

更に、本発明は、(a)SEQ ID NO: 1-47からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b)SEQ ID NO: 1-47からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c)SEQ ID NO: 1-47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: 1-47からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドのアнтаゴニストとして有効な化合物をスクリーニングする方法を提供する。この方法は、(a)このポリペプチドを含むサンプルを化合物に曝露するステップと、(b)このサンプルのアнтаゴニスト活性を検出するステップとを含む。別法では、本発明は、この方法によって同定されたアンタゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的RMEPの過剰な発現に関連した疾患やその症状の治療方法を提供する。

【0042】

更に本発明は、(a)SEQ ID NO: 1-47からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b)SEQ ID NO: 1-47からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと

、(c) SEQ ID NO: 1 - 47 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 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドに特異的に結合する化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを好適な条件下で少なくとも1つの試験化合物と結合させるステップと、(b) このポリペプチドとこの試験化合物との結合を検出して、このポリペプチドと特異的に結合する化合物を同定するステップとを含む。

【0043】

更に本発明は、(a) SEQ ID NO: 1 - 47 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1 - 47 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1 - 47 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 - 47 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドの活性を調節する化合物をスクリーニングする方法を提供する。このスクリーニング方法は、(a) このポリペプチドを、その活性が許容される条件下で少なくとも1つの化合物と結合させるステップと、(b) この試験化合物の存在下でのこのポリペプチドの活性を評価するステップと、(c) この試験化合物の存在下でのこのポリペプチドの活性と、この試験化合物の非存在下でのこのポリペプチドの活性とを比較するステップとを含み、この試験化合物の存在下でのこのポリペプチドの活性の変化が、このポリペプチドの活性を調節する化合物の存在を示唆する。

【0044】

更に本発明は、SEQ ID NO: 48 - 94 からなる群から選択された配列を含む標的ポリヌクレオチドの発現を変化させるのに有効な化合物をスクリーニングする方法であって、(a) この標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、(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を含む、該スクリーニング方法を提供する。

【0045】

本発明はさらに、試験化合物の毒性を評価する方法を提供する。この方法は、(a) 核酸を含む生体サンプルを前記試験化合物で処理するステップと、(b) 処理した前記生体サンプルの核酸をプローブとハイブリダイズするステップと、(c) ハイブリダイゼーション複合体の収量を測定するステップと、(d) 前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処理の生体サンプルにおけるハイブリダイゼーション複合体の収量とを比較するステップとを含み、前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量の差異が試験化合物の毒性を示唆する。この方法における前記プローブが、(1) SEQ ID NO: 48 - 94 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(2) SEQ ID NO: 48 - 94 からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(3) 前記(1)に相補的な配列を有するポリヌクレオチドと、(4) 前記(2)のポリヌクレオチドに相補的なポリヌクレオチドと、(5) 前記(1) (4)のRNA等価物とで構成される群から選択されたポリヌクレオチドの少なくとも20個の連続するヌクレオチドを含む。また、前記ハイブリダイゼーションは、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われる。また、前記標的ポリヌクレオチドが、(1) SEQ ID NO: 48 - 94 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(2) SEQ ID NO: 48 - 94 からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(3) 前記(1)のポリヌクレオチドに相補的なポリヌクレオチドと、(4) 前記(2)のポリヌクレオチドに相補的なポリヌクレオチドと、(5) 前記(1) (4)のRNA等価物とで構成される群から選択される。

別法では、標的ポリヌクレオチドは、その標的ポリヌクレオチドの断片である。

【0046】

(本発明の記載について)

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

【0047】

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

【0048】

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

【0049】

(定義)

用語「RMEP」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RMEPのアミノ酸配列を指す。

【0050】

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

【0051】

用語「アレル変異配列」は、RMEPをコードする遺伝子の別の形を指す。アレル変異配列は、核酸配列における少なくとも1つの変異によって生じ、変異mRNA若しくは変異ポリペプチドになり、これらの構造や機能は変わる場合もあれば変わらない場合もある。ある遺伝子は、天然型のアレル変異配列が存在しないもの、1つ或いは多数存在するものがある。一般にアレル変異配列を生じる変異は、ヌクレオチドの自然な欠失、付加、或いは置換による。これらの各変異は、単独或いは他の変異と同時に起こり、所定の配列内で一回或いはそれ以上生じる。

【0052】

RMEPをコードする「変異」核酸配列は、様々なヌクレオチドの欠失、挿入、或いは置換が起こっても、RMEPと同じポリペプチド或いはRMEPの機能特性の少なくとも1つを備えるポリペプチドを指す。この定義には、RMEPをコードするポリヌクレオチド配列の正常な染色体の遺伝子座ではない位置でのアレル変異配列との不適当或いは予期しないハイブリダイゼーション、並びにRMEPをコードするポリヌクレオチドの特定のオリゴヌクレオチドプローブを用いて容易に検出可能な或いは検出困難な多形性を含む。コードされたタンパク質も変異され得り、サイレント変化を生じRMEPと機能的に等価となるアミノ酸残基の欠失、挿入、或いは置換を含み得る。意図的なアミノ酸置換は、生物学的或いは免疫学的にRMEPの活性が保持される範囲で、残基の極性、電荷、溶解度、疎水性、親水性、及び/または両親媒性についての類似性に基づいて成され得る。例えば

10

20

30

40

50

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

【0053】

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

10

【0054】

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

【0055】

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

【0056】

用語「抗体」は、抗原決定基と結合可能なFab及びF(ab')₂、及びそれらの断片、Fv断片などの無傷の分子を指す。RMEPポリペプチドと結合する抗体は、抗体を免疫する小ペプチドを含む無傷の分子またはその断片を用いて作製可能である。動物（例えば、マウス、ラット、若しくはウサギ）を免疫化するのに使用されるポリペプチド或いはオリゴペプチドは、RNAの翻訳から、或いは化学的に合成可能であり、必要に応じて担体タンパク質と結合させることも可能である。ペプチドと化学的に結合した一般に用いられる担体は、ウシ血清アルブミン、チログロブリン、及びキーホールリンペットヘモニアン（KLH）を含む。次ぎに、この結合したペプチドを用いて動物を免疫化する。

20

【0057】

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

30

【0058】

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

40

【0059】

用語「生物学的に活性」は、天然分子の構造的、調節的、或いは生化学的な機能を有するタンパク質を指す。同様に、用語「免疫学的に活性」または「免疫原性」は、天然或いは

50

組換え体の R M E P、合成の R M E P 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。

【 0 0 6 0 】

用語「相補的」は、塩基対合によってアニールする 2 つの一本鎖核酸配列間の関係を指す。例えば、配列「5' A G T 3'」が相補的な配列「3' T C A 5'」と対をなす。

【 0 0 6 1 】

「所定のポリヌクレオチド配列を含む組成物」または「所定のアミノ酸配列を含む組成物」は広い意味で、所定のヌクレオチド配列若しくはアミノ酸配列を含む任意の組成物を指す。この組成物は、乾燥した製剤或いは水溶液を含み得る。R M E P 若しくは R M E P の断片をコードするポリヌクレオチド配列を含む組成物は、ハイブリダイゼーションプローブとして使用され得る。このプローブは、凍結乾燥状態で保存可能であり、糖質などの安定化剤と結合させることが可能である。ハイブリダイゼーションにおいて、プローブは、塩（例えば、NaCl）及び界面活性剤（例えば、SDS：ドデシル硫酸ナトリウム）、その他の物質（例えば、デンハート液、乾燥ミルク、サケ精子DNAなど）を含む水溶液に展開され得る。

10

【 0 0 6 2 】

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

20

【 0 0 6 3 】

用語「保存的なアミノ酸置換」は、元のタンパク質の特性を殆ど変えない置換を指す。即ち、置換によってそのタンパク質の構造や機能が大きくは変わらず、そのタンパク質の構造、特にその機能が保存される。以下に、あるタンパク質の元のアミノ酸が別のアミノ酸に置換される保存的なアミノ酸置換を示す。

元の残基

保存的な置換

30

A l a	G l y , S e t
A r g	H i s , L y s
A s n	A s p , G l n , H i s
A s p	A s n , G l u
C y s	A l a , S e r
G l n	A s n , G l u , H i s
G l u	A s p , G l n , H i s
G l y	A l a
H i s	A s n , A r g , G l n , G l u
I l e	L e u , V a l
L e u	I l e , V a l
L y s	A r g , G l n , G l u
M e t	L e u , I l e
P h e	H i s , M e t , L e u , T r p , T y r
S e r	C y s , T h r
T h r	S e r , V a l
T r p	P h e , T y r
T y r	H i s , P h e , T r p
V a l	I l e , L e u , T h r

40

一般に、保存されたアミノ酸置換の場合は、a) 置換された領域のポリペプチドの骨格構

50

造、例えば、シートやヘリックス高次構造、b)置換された部位の分子の電荷または疎水性、及び/または、c)側鎖の大半が維持される。

【0064】

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

【0065】

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

10

【0066】

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

【0067】

「異なる発現」とは、少なくとも2つの異なったサンプルの比較により決められる遺伝子またはタンパク質の発現の上昇即ちアップレギュレーションや、減少即ちダウンレギュレーション、または発現しないことを指す。このような比較は、例えば、処置したサンプルと未処置のサンプルとの間、または病変サンプルと正常なサンプルとの間で行う。

20

【0068】

用語「断片」は、RMEPまたはRMEPをコードするポリヌクレオチドの固有の部分であって、その親配列(parent sequence)と同一であるがその配列より長さが短いものを指す。「断片」の最大の長さは、親配列から1つのヌクレオチド/アミノ酸残基を差し引いた長さである。例えば、ある断片は、5~1000個の連続するヌクレオチド或いはアミノ酸残基を含む。プローブ、プライマー、抗原、治療用分子、またはその他の目的に用いる断片は、少なくとも5、10、15、16、20、25、30、40、50、60、75、100、150、250若しくは500個の連続するヌクレオチド或いはアミノ酸残基の長さである。断片は、優先的に分子の特定の領域から選択される場合もある。例えば、ポリペプチド断片は、所定の配列に示された最初の250若しくは500のアミノ酸(或いは、ポリペプチドの最初の25%または50%)から選択された連続するアミノ酸の所定の長さを含み得る。これらの長さは一例であり、配列表及び表、図面を含む明細書に記載の任意の長さが、本発明の実施例に含まれ得る。

30

【0069】

SEQ ID NO: 48-94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: 48-94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: 48-94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: 48-94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。ある断片と一致するSEQ ID NO: 48-94の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に測定できる。

40

【0070】

SEQ ID NO: 1-47のある断片は、SEQ ID NO: 48-94のある断片によってコードされる。SEQ ID NO: 1-47のある断片は、SEQ ID NO: 1-47を特異的に同定する固有のアミノ酸配列の領域を含む。例えば、SEQ ID NO: 1-47のある断片は、SEQ ID NO: 1-47を特異的に認識する抗体の開発における免疫原性ペプチドとして有用である。ある断片と一致するSEQ ID NO: 1-47の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に決定できる。

50

【0071】

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

【0072】

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

【0073】

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

【0074】

ポリヌクレオチド配列間の同一性のパーセントは、MEGALIGN version 3.12e 配列アラインメントプログラムに組込まれるCLUSTAL Vアルゴリズムのデフォルトパラメータを用いて決定可能である。このプログラムはLASERGENEソフトウェアパッケージの一部であり、分子生物学分析プログラム一式(DNASTAR, Madison WI)である。このCLUSTAL Vは、Higgins, D.G. 及び P.M. Sharp (1989) CABIOS 5:151-153、Higgins, D.G. 他 (1992) CABIOS 8:189-191に記載されている。ポリヌクレオチド配列の対のアライメントの場合、デフォルトパラメータは、Ktuple = 2、gap penalty = 5、window = 4、「diagonals saved」= 4と設定する。「重み付けされた」残基重み付け表が、デフォルトとして選択された。同一性のパーセントは、アラインメントされたポリヌクレオチド配列間の「類似性のパーセント」としてCLUSTAL Vによって報告される。

【0075】

別法では、National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410) が提供する、広く用いられている無料の配列比較アルゴリズム一式が、NCBI (Bethesda, MD) を含む幾つかのソース及びインターネット (<http://WWW.ncbi.nlm.nih.gov/BLAST/>) で入手可能である。このBLASTソフトウェア一式には、既知のポリヌクレオチド配列と様々なデータベースの別のポリヌクレオチド配列とのアラインメントに用いられる「blastn」を含む、様々な配列分析プログラムが含まれる。「BLAST 2 Sequences」と呼ばれるツールが入手可能であり、2つのヌクレオチド配列の対を直接比較するために用いられる。「BLAST 2 Sequences」は、<http://WWW.ncbi.nlm.nih.gov/gorf/b12.html> にアクセスして、対話形式で利用できる。「BLAST 2 Sequences」ツールは、blastn 及び blastp (以下に記載) の両方に用いることができる。BLASTプログラムは、一般的には、デフォルトを設定するギャップ及び他のパラメータと共に用いられる。例えば、2つのヌクレオチド配列を比較する場合、ある者は、デフォルトパラメータに設定された「BLAST 2 Sequences」ツール Version 2.0.12 (April - 21 - 2000) でblastnを使用するであろう。そのようなデフォルトパラメータは、例えば、以下のようにする。

【0076】

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

10

20

30

40

50

Open Gap : 5 及び Extension Gap : 2 penalties
 Gap x drop-off : 50
 Expect : 10
 Word Size : 11
 Filter : on

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

10

【0077】

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

【0078】

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

20

【0079】

ポリペプチド配列間の同一性のパーセントは、MEGALIGN バージョン3.12e 配列アラインメントプログラム（上記）に組込まれるCLUSTAL Vアルゴリズムのデフォルトパラメータを用いて決定可能である。CLUSTAL Vを用いる対方式のポリペプチド配列のアライメントの場合、デフォルトパラメータは、Ktuple = 1、gap penalty = 3、window = 5、及び「diagonals saved」= 5と設定する。PAM250マトリクスが、デフォルトの残基重み付け表として選択される。ポリヌクレオチドアラインメントと同様に、アラインメントされたポリペプチド配列の対の同一性のパーセントは、「類似性のパーセント」としてCLUSTAL Vによって報告される。

30

【0080】

別法では、NCBI BLASTソフトウェア式が用いられる。例えば、2つのポリペプチド配列を対で比較をする場合、ある者は、デフォルトパラメータで設定された「BLAST 2 Sequences」ツールVersion 2.0.12 (Apr - 21 - 2000)でblastpを使用するであろう。そのようなデフォルトパラメータは、例えば、以下のようにする。

【0081】

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

40

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

50

片を用いて、同一性のパーセントが測定される長さを示すことができる。

【0082】

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

【0083】

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。

【0084】

「ハイブリダイゼーション」とは、所定のハイブリダイゼーション条件下で、ある一本鎖ポリヌクレオチドがある相補的な一本鎖と塩基対を形成するアニーリングのプロセスである。特異的なハイブリダイゼーションとは、2つの核酸配列が高い相同性を有することを意味する。アニーリングが許容される条件下で、特異的なハイブリダイゼーション複合体が形成され、洗浄過程の後にもハイブリダイズしたままである。洗浄過程は、ハイブリダイゼーションプロセスの厳密性即ちストリンジェンシー（stringency）の決定において特に重要であり、よりストリンジェントな条件では、非特異的な結合、即ち完全には一致しない核酸鎖間の対の結合が減少する。核酸配列間のアニーリングが許容される条件は、当業者によって日常的に決定され、ハイブリダイゼーションの間は一定であるが、洗浄過程は、目的のストリンジェンシーにするためにその最中に条件の変更が可能であり、ハイブリダイゼーション特異性が得られる。アニーリングが許容される条件は、例えば、温度が68℃で、約6×SSC、約1%（w/v）のSDS、並びに約100 µg/mlのせん断して変性したサケ精子DNAが含まれる。

【0085】

一般に、ハイブリダイゼーションのストリンジェンシーは、洗浄過程を行う際の温度によっても左右される。この洗浄温度は通常、所定のイオン強度とpHにおける特定の配列の熱融点（ T_m ）より約5～20℃低く選択される。この T_m は、（所定のイオン強度とpHの下）標的の配列の50%が完全に一致するプローブとハイブリダイズする温度である。 T_m を計算する式及び核酸のハイブリダイゼーションの条件は、周知であり、Sambrook, J. 他による、1989, Molecular Cloning: A Laboratory Manual, 第2版の1-3巻, Cold Spring Harbor Press, Plainview NY; 特に2巻の9章に記載されている。

【0086】

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

【0087】

用語「ハイブリダイゼーション複合体」は、相補的な塩基間の水素結合によって、形成された2つの核酸配列の複合体を指す。ハイブリダイゼーション複合体は溶液中（例えば、C₀tまたはR₀t分析）で形成されるか、或いは溶液中の1つの核酸配列と固体の支持物（例えば、紙、膜、フィルタ、チップ、ピン、或いはスライドガラス、または細胞及びその核酸を固定する任意の適当な基板）に固定されたもう一つの核酸配列とで形成され得

る。

【0088】

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。

【0089】

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

【0090】

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RMEP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RMEPのポリペプチド断片またはオリゴペプチド断片を含む。

10

【0091】

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

【0092】

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。

【0093】

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

20

【0094】

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

【0095】

「機能的に結合した」は、第1の核酸配列と第2の核酸配列が機能的な関係にある状態を指す。例えば、プロモーターがコード配列の転写または発現に影響を与える場合、そのプロモーターはそのコード配列に機能的に結合している。一般に、機能的に結合したDNA配列は、同じ読み枠内で2つのタンパク質をコードする領域が結合する必要がある場合は、非常に近接或いは連続する。

30

【0096】

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

40

【0097】

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

【0098】

「プローブ」とは、同一配列或いはアレル核酸配列、関連する核酸配列の検出に用いる、RMEPやそれらの相補配列、またはそれらの断片をコードする核酸配列のことである。プローブは、検出可能な標識またはレポーター分子が結合され単離されたオリゴヌクレオチドやポリヌクレオチドである。典型的な標識には、放射性アイソトープ及びリガンド、

50

化学発光試薬、酵素がある。「プライマー」とは、相補的な塩基対を形成して標的のポリヌクレオチドにアニーリング可能な、通常はDNAオリゴヌクレオチドである短い核酸である。プライマーがポリヌクレオチドにアニーリングした後、あるDNAポリメラーゼ酵素によって、標的のDNA一本鎖に沿って伸長される。プライマーの組は、例えば、PCR法における核酸配列の増幅（及び同定）に用いることができる。

【0099】

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

10

【0100】

プローブ及びプライマーの準備及び使用方法については、例えば、Sambrook, J. 他による、1989年、名称「Molecular Cloning: A Laboratory Manual」、第2版の1-3巻(Cold Spring Harbor Press, Plainview NY)、またはAusubel, F.M. 他による、1987年、名称「Current Protocols in Molecular Biology」(Greene Publi. Assoc. & Wiley-Intersciences, New York NY)、並びにInnis他による、1990年、名称「PCR Protocols, A Guide to Methods and Applications」(Academic Press, San Diego CA.)を参照されたい。PCR用のプライマーの組は、例えば、Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA)などのそのような目的のためのコンピュータプログラムを用いて、ある既知の配列から引き出すことができる。

20

【0101】

プライマーとして用いるオリゴヌクレオチドは、当分野で周知のプライマー選択用のコンピュータプログラムで選択される。例えば、OLIGO 4.06ソフトウェアは、それぞれが最大100ヌクレオチドまでのPCR用のプライマーの対の選択、及び32,000塩基までの入力ポリヌクレオチド配列から最大5,000ヌクレオチドまでの大きなポリヌクレオチド及びオリゴヌクレオチドの分析に有用である。類似のプライマー選択用プログラムには、能力を拡大する追加の機能が含まれている。例えば、PrimOUプライマー選択プログラム(Genome Center at University of Texas South West Medical Center, Dallas TXより入手可能)は、メガベース配列から特定のプライマーを選択できるため、ゲノムワイドスコープ(genome-wide scope)におけるプライマーの設計に有用である。Primer3プライマー選択プログラム(Whitehead Institute/MIT Center for Genome Research, Cambridge MA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。また、Primer3は、特にマイクロアレイのオリゴヌクレオチドの選択に有用である(後の方の2つのプライマー選択プログラムのソースコードは、それぞれのソースから得ることができ、ユーザーのニーズを満たすように変更することもできる)。PrimerGenプログラム(UK Human Genome Mapping Project Resource Centre, Cambridge UKより入手可能)は、多数の配列アラインメントに基づいてプライマーを設計するため、アラインメントされた核酸配列の最も保存された領域或いは最も保存されていない領域のどちらかとハイブリダイズするプライマーを選択することができる。従って、このプログラムは、固有及び

30

40

50

保存されたオリゴヌクレオチドやポリヌクレオチドの断片の同定に有用である。上記した任意の選択方法で同定されたオリゴヌクレオチドやポリヌクレオチドの断片は、例えば、PCR法やシーケンシングプライマー、マイクロアレイエレメント、或いはサンプルの核酸の完全或いは部分的に相補的なポリヌクレオチドを同定する特定のプローブなどの、ハイブリダイゼーション技術に有用である。オリゴヌクレオチドの選択方法は、上記した方法に制限されるものではない。

【0102】

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

10

【0103】

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

【0104】

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

20

【0105】

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

【0106】

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

30

【0107】

用語「サンプル」は、その最も広い意味で用いられている。RMEP、RMEP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。

【0108】

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

40

【0109】

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

50

除去されたものを指す。

【0110】

「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。

【0111】

用語「基板」は、任意の好適な固体或いは半固体の支持物を指し、膜及びフィルタ、チップ、スライド、ウエハ、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。この基板には、壁または塹壕、ピン、チャンネル、細孔などの様々な表面形態があり、そこにポリヌクレオチドやポリペプチドが結合する。

10

【0112】

「転写イメージ」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。

【0113】

「形質転換」とは、外来DNAが受容細胞に導入されるプロセスのことである。形質転換は、当分野で周知の種々の方法により、自然或いは人工の条件下で起こり、原核宿主細胞若しくは真核宿主細胞の中に外来核酸配列を挿入する任意の周知の方法によって行うことができる。この形質転換の方法は、形質転換される宿主細胞のタイプによって選択される。この方法には、バクテリオファージまたはウイルス感染、電気穿孔法（エレクトロポレーション）、リポフェクション、及び微粒子照射が含まれるが、これらに限定されるものではない。「形質転換された」細胞には、導入されたDNAが自律的に複製するプラスミドとして或いは宿主染色体の一部として複製可能である安定的に形質転換された細胞が含まれる。さらに、限られた時間に一時的に導入DNA若しくは導入RNAを発現する細胞も含まれる。

20

【0114】

本明細書における「遺伝子組換え生物」とは、当分野で周知の遺伝子組換え技術などを用いて、人間が生物の1つ以上の細胞に異種の核酸を導入した任意の生物であり、動物及び植物を含むが、それらに限定されるものではない。微量注入や組換えウイルスに感染させるなどの慎重な遺伝子操作によって、細胞の前駆体に直接或いは間接的に異種核酸を細胞に導入する。「遺伝子操作」とは、典型的な交雑育種や*in vitro*での受精ではなく、組換えDNA分子を導入することである。本発明に従った遺伝子組換え生物には、細菌及びラン藻類、菌類、植物、動物が含まれる。本発明の単離されたDNAは、当分野で周知の、例えば、感染、形質移入、形質転換、トランス接合（trans conjugation）などの方法によって、宿主に導入することができる。本発明のDNAをそのような生物に導入する技術は周知であり、前出のSambrook他（1989）に記載されている。

30

【0115】

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

40

50

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ヌクレオチド多型」(SNP)も含み得る。SNPの存在は、例えば、或る集団、病態、病態の性向を示唆し得る。

【0116】

特定のポリペプチド配列の「変異体」とは、デフォルトパラメータ設定の「BLAST 2 Sequences」ツールVersion 2.0.9 (May - 07 - 1999)を用いるblastpによって、あるポリペプチド配列のある長さに対する該特定のポリペプチド配列の同一性が、少なくとも40%と決定されたポリペプチド配列のことである。このようなポリペプチドの対は、ある長さにおいて、例えば、少なくとも50%または60%、70%、80%、85%、90%、95%、98%、或いはそれ以上の同一性を示し得る。 10

【0117】

(発明)

本発明は、新規のヒトRNA代謝タンパク質(RMEP)及びRMEPをコードするポリヌクレオチドの発見に基づき、これらの組成物を利用した神経系の疾患、自己免疫/炎症性の疾患、細胞増殖異常、及び発生または発達障害の診断、治療、及び予防に関する。

【0118】

表1は、本発明の完全長ポリヌクレオチド配列及びポリペプチド配列の識別番号を示す。各ポリヌクレオチドおよびそれに対応するポリペプチドは、1つのインサイトプロジェクト識別番号(Incyte Project ID)に相関する。各ポリペプチド配列は、記載されているようにポリペプチド配列識別番号(Polypeptide SEQ ID NO:)およびインサイトポリペプチド配列番号(Incyte Polypeptide ID)の両方によって示されている。各ポリヌクレオチド配列は、記載されているようにポリヌクレオチド配列識別番号(Polynucleotide SEQ ID NO:)およびインサイトポリヌクレオチドコンセンサス配列番号(Incyte Polynucleotide ID)の両方によって示されている。 20

【0119】

表2は、GenBankタンパク質(genept)データベースにおいてBLAST解析により同定された本発明のポリペプチドに相同性を有する配列を示す。列1および列2はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号(Polypeptide SEQ ID NO:)およびそれに対応するインサイトポリペプチド配列番号(Incyte Polypeptide ID)を示す。列3は、GenBankの最も近い相同体のGenBankの識別番号(Genbank ID NO:)を示す。列4は、各ポリペプチドとそのGenBank相同体との間の一致を表す確率スコアを示す。列5は、GwnBank相同体のアノテーションを示し、更に該当箇所には適当な引用文も示す。これらを引用することを以って本明細書の一部とする。 30

【0120】

表3は、本発明の各ポリペプチドの様々な構造的特徴を示す。列1および列2はそれぞれ、本発明の各ポリペプチドのポリペプチド配列識別番号(SEQ ID NO:)およびそれに対応するインサイトポリペプチド配列番号(Incyte Polypeptide 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、潜在的なリン酸化部位および潜在的なグリコシル化部位を示す。列6は、シグネチャ(signature)配列、ドメイン、およびモチーフを含むアミノ酸残基を示す。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。 40

【0121】

表2および表3は共に、本発明の各ポリペプチドの特性を一覧にしたものであって、これらの特性が請求するポリペプチドがRNA代謝タンパク質であることを立証する。SEQ 50

I D N O : 4 6 が、ナイセリアメニンギチジスの G l u - t R N A ^{G l n} アミドトランスフェラーゼ (G e n B a n k I D g 7 2 2 6 6 0 1) のサブユニット A と 2 9 % の同一性を有することが B a s i c L o c a l A l i g n m e n t S e a r c h T o o l (B L A S T) によって示された (表 2 を参照) 。 B L A S T の確率スコアは 1.3×10^{-37} である。これは、探しているポリペプチド配列アラインメントが偶然の一致により得られる確率を示す。 S E Q I D N O : 4 6 はまた、アミダーゼシグネチャ配列を含む。これは、隠れマルコフモデル (H M M) を基にした保存されたタンパク質ファミリードメインの P F A M データベースにおいて、統計的に有意な一致を検索して決定された (表 3 を参照) 。 B L I M P S 及び P R O F I L E S C A N 分析から得られたデータにより、 S E Q I D N O : 4 6 が、アミノ交換反応に関与するポリペプチドの特徴であるアミダーゼシグネチャ配列を含むことが裏付けられた。これらのデータから、 S E Q I D N O : 4 6 が、本発明の段階ではヒトには見られないが真核生物及び或る種の細胞小器官に見られる G l u - t R N A ^{G l n} アミドトランスフェラーゼに関連するという確証が得られた。また、 S E Q I D N O : 4 7 が、トウモロコシの 6 0 S 酸性リボソームタンパク質 (G e n B a n k I D g 7 9 0 5 0 8) と 9 7 % の同一性を有することが B a s i c L o c a l A l i g n m e n t S e a r c h T o o l (B L A S T) によって示された (表 2 を参照) 。 B L A S T の確率スコアは 5.4×10^{-51} である。 S E Q I D N O : 4 7 はまた、6 0 S 酸性リボソームタンパク質ドメインを含む。これは、隠れマルコフモデル (H M M) を基にした保存されたタンパク質ファミリードメインの P F A M データベースにおいて、統計的に有意な一致を検索して決定された (表 3 を参照) 。 B L I M P S で得られたデータにより、 S E Q I D N O : 4 7 が、リン酸化した (即ち酸性の可能性が高い) リボソームタンパク質であるということが裏付けられた。 S E Q I D N O : 1 - 4 5 も同様の方法で解析してアノテーションを付けた。 S E Q I D N O : 1 - 4 7 の解析のためのアルゴリズムおよびパラメータを表 7 に記載する。

10

20

30

40

50

【 0 1 2 2 】

表 4 に示されているように、本発明の完全長ポリヌクレオチド配列は、 c D N A 配列、またはゲノム D N A 由来のコード (エキソン) 配列、或いはこれらの 2 種類の配列のあらゆる組み合わせを用いて組み立てた。列 1 および列 2 はそれぞれ、本発明の各ポリヌクレオチドのポリヌクレオチド配列識別番号 (P o l y n u c l e o t i d e S E Q I D N O :) およびそれに対応するインサイトポリヌクレオチドコンセンサス配列番号 (I n c y t e P o l y n u c l e o t i d e I D) を示す。列 3 は、塩基対における各ポリヌクレオチド配列の長さを示す。列 4 は、例えば、 S E Q I D N O : 4 8 - 9 4 を同定するため、或いは S E Q I D N O : 4 8 - 9 4 と関連するポリヌクレオチド配列とを区別するためのハイブリダイゼーションまたは増幅技術に有用なポリヌクレオチド配列の断片を示す。列 5 は、 c D N A 配列、ゲノム D N A から推定されるコード配列 (エキソン) 、および / または c D N A およびゲノム D N A の両方からなる群に対応する識別番号を示す。これらの配列を用いて本発明の完全長ポリヌクレオチド配列を組み立てた。表 4 の列 6 および列 7 はそれぞれ、それぞれの完全長配列に対する列 5 の c D N A 配列の開始ヌクレオチド (5 ') 位置および終了ヌクレオチド (3 ') 位置を示す。

【 0 1 2 3 】

表 4 の列 5 に示されている識別番号は、具体的には、例えばインサイト c D N A およびそれらに対応する c D N A ライブラリの識別番号を示す。例えば、 6 4 2 0 1 7 H 1 はインサイト c D N A 配列の識別番号であり、 B R S T N O T 0 3 はそれが由来する c D N A ライブラリの識別番号である。 c D N A ライブラリが示されていないインサイト c D N A は、プールされている c D N A ライブラリ (例えば、 7 0 8 2 2 0 1 5 V 1) に由来する。または、列 5 の識別番号は、完全長ポリヌクレオチド配列の組み立てに用いた G e n B a n k の c D N A すなわち E S T (例えば、 g 1 1 3 6 8 4 1) の識別番号の場合もある。または、列 5 の識別番号は、 G e n s c a n 分析によって推定されるゲノム D N A のコード領域の場合もある。この G e n s c a n 推定コード配列は、配列を組み立てる前に編集する場合がある (実施例 4 を参照) 。または、列 5 の識別番号は、 “ e x o n - s t i t

c h i n g ” アルゴリズムによって c D N A および G e n s c a n 推定エキソンの両方からなる群の場合もある（実施例 5を参照）。または、列 5 の識別番号は、“ e x o n - s t r e t c h i n g ” アルゴリズムによって c D N A および G e n s c a n 推定エキソンの両方からなる群の場合もある（実施例 5を参照）。場合によっては、列 5 に示されている配列の範囲と重複するインサイト c D N A の範囲が得られ、最終的なコンセンサス配列が決定されるが、それに相当するインサイト c D N A の識別番号は示されていない。

【 0 1 2 4 】

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

10

【 0 1 2 5 】

本発明はまた、R M E P の変異体も含む。好適な R M E P の変異体は、R M E P の機能的或いは構造的特徴の少なくともどちらか一方を有し、かつ R M E P アミノ酸配列に対して少なくとも約 8 0 % のアミノ酸配列同一性、或いは少なくとも約 9 0 % のアミノ酸配列同一性、更には少なくとも約 9 5 % のアミノ酸配列同一性を有する。

【 0 1 2 6 】

本発明はまた、R M E P をコードするポリヌクレオチドを提供する。特定の実施例において、本発明は、R M E P をコードする S E Q I D N O : 4 8 - 9 4 からなる群から選択された配列を含むポリヌクレオチド配列を提供する。配列表に示した S E Q I D N O : 4 8 - 9 4 のポリヌクレオチド配列は、窒素系塩基のチミンがウラシルに置換され、糖鎖の背骨がデオキシリボースではなくリボースからなる等価 R N A 配列を含む。

20

【 0 1 2 7 】

本発明はまた、R M E P 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、R M E P をコードするポリヌクレオチド配列と少なくとも 7 0 % のポリヌクレオチド配列同一性、或いは少なくとも 8 5 % のポリヌクレオチド配列同一性、更には少なくとも 9 5 % ものポリヌクレオチド配列同一性を有する。本発明の特定の実施形態は、S E Q I D N O : 4 8 - 9 4 からなる群から選択された核酸配列と少なくとも 7 0 % のポリヌクレオチド配列同一性、或いは少なくとも 8 5 % のポリヌクレオチド配列同一性、更には少なくとも 9 5 % ものポリヌクレオチド配列同一性を有する S E Q I D N O : 4 8 - 9 4 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。上記したポリヌクレオチド変異配列は何れも、R M E P の機能的或いは構造的特徴の少なくとも 1 つを有するアミノ酸配列をコードする。

30

【 0 1 2 8 】

遺伝暗号の縮重により作り出され得る R M E P をコードする種々のポリヌクレオチド配列には、既知の自然発生する任意の遺伝子のポリヌクレオチド配列と最小の類似性しか有しないものも含まれることを、当業者は理解するであろう。したがって本発明には、可能なコドン選択に基づいた組み合わせの選択によって作り出され得る可能なポリヌクレオチド配列の変異の全てが含まれ得る。これらの組み合わせは、天然の R M E P のポリヌクレオチド配列に適用される標準的なトリプレット遺伝暗号を基に作られ、全ての変異が明確に開示されていると考慮する。

40

【 0 1 2 9 】

R M E P をコードするヌクレオチド配列及びその変異配列は一般に、好適に選択されたストリンジェントな条件下で、天然の R M E P のヌクレオチド配列とハイブリダイズ可能であるが、非天然のコドンを含めるなどの実質的に異なった使い方のコドンを有する R M E P 或いはその誘導体をコードするヌクレオチド配列を作るとは有利となり得る。特定のコドンが宿主によって利用される頻度に基づいてコドンを選択して、ペプチドの発現が特定の真核細胞または原核宿主に発生する割合を高めることが可能である。コードされたア

50

ミノ酸配列を変えないで、RMEP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。

【0130】

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

【0131】

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: 48 - 94及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる（例えば、Wahl, G. M. 及び S. L. Berger (1987) *Methods Enzymol.* 152: 399 - 407; and Kimmel, A. R. (1987) *Methods Enzymol.* 152: 507 - 511. を参照）。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。

【0132】

当分野で周知のDNAのシーケンシング方法を用いて、本発明の何れの実施例も実行可能である。この方法には、例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amersham, Pharmacia Biotech Piscataway NJ)、或いはELONGASE増幅システム (Life Technologies, Gaithersburg MD) にみられるような校正エキソヌクレアーゼとポリメラーゼとの組み合わせなどの酵素が用いられる。好ましくは、MICROLAB 2200液体転移システム (Hamilton, Reno, NV)、PTC 200 Thermal Cycler 200 (MJ Research, Watertown MA) 及びABI CATALYST 800 (PE Biosystems) などの装置を用いて配列の準備を自動化する。次に、ABI 373 或いは377 DNAシーケンシングシステム (PE Biosystems)、MEGABACE 1000 DNAシーケンシングシステム (Molecular Dynamics, Sunnyvale CA) または当分野で周知の他の方法を用いてシーケンシングを行う。得られた配列を当分野で周知の様々なアルゴリズムを用いて分析する（例えば、Ausubel, F. M. (1997) *Short Protocols in Molecular Biology*, John Wiley & Sons, New York NY, unit 7.7; Meyers, R. A. (1995) *Molecular Biology and Biotechnology*, Wiley VCH, New York NY, pp. 856 - 853. を参照）。

【0133】

当分野で周知のPCR法をベースにした種々の方法で、部分的なヌクレオチド配列を利用して、RMEP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。例えば制限部位PCR法を利用する1つの方法では、一般的なプライマー及び入れ子プライマー (nested primer) を用いてクローニングベクター内のゲノムDNAから未知の配列を増幅する（例えば、Sarkar, G. (1993) *PCR Methods Applic* 2: 318 - 322 を参照）。逆PCR法を用いる別法では、広範な方向に伸長して環状化した鋳型から未知の配列を増幅するプライマーを用いる。この鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限断片に由来する（例えば、Triglia, T. ら (1988) *Nucleic Acids Res* 16: 8186 を参照）。キャプチャPCR法を用いる第3の方法は、ヒト及び酵母菌人工染色体DNAの既知の配列に隣接するDNA断片のPCR増幅を含む（例

10

20

30

40

50

例えば、Lagerstrom, M. 他 (1991) PCR Methods Appl 1: 111-119 を参照)。この方法では、多数の制限酵素による消化及びライゲーションを用いて、PCRを行う前に未知の配列の領域の中に組換え二本鎖配列を挿入することが可能である。また、当分野で周知の別の方法を用いて未知の配列を得ることも可能である。(例えば、Parker, J. D. 他 (1991) Nucleic Acids Res. 19: 3055-3060 を参照)。更に、PCR、ネスト化プライマー、PROMOTERFINDERライブラリ(Clontech, Palo Alto CA)を用いれば、ゲノムDNA内の歩行が可能である。この方法ではライブラリをスクリーニングする必要がなく、イントロン/エキソン接合部を探すのに有用である。全てのPCR法をベースにした方法では、プライマーは、市販のOLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) 或いは別の好適なプログラムなどを用いて、長さが22~30ヌクレオチド、GC含量が50%以上、約68~72の温度で鋳型に対してアニーリングするよう設計される。

10

【0134】

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

20

【0135】

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

30

【0136】

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

【0137】

種々の目的でRMEPをコードする配列を変えるために、当分野で一般的に知られている方法を用いて、本発明のヌクレオチド配列を組換えることができる。この目的には、遺伝子産物のクローン化、プロセッシング及び/または発現の調節が含まれるが、これらに限定されるものではない。ランダムな断片によるDNAの混合や遺伝子断片と合成オリゴヌクレオチドのPCR再組み立てを用いて、ヌクレオチド配列の組換えが可能である。例えば、オリゴヌクレオチド媒介性定方向突然変異誘発を利用して、新しい制限部位を生成する突然変異の導入、グリコシル化パターンの変更、コドン選択の変更、スプライスバリエーションの作製等が可能である。

40

【0138】

本発明のヌクレオチドを、MOLECULAR BREEDING (Maxygen Inc., Santa Clara CA; 米国特許第5,837,458号; Chang, C.-C. 他 (1999) Nat. Biotechnol. 17: 793-79

50

7; Christians, F. C. 他 (1999) Nat. Biotechnol. 17:259-264; Cramer, A. 他 (1996) Nat. Biotechnol. 14:315-319) などのDNAシャフリング技術を用いてシャフリングして、RMEPの生物学的または酵素的な活性、或いは他の分子や化合物と結合する能力などのRMEPの生物学的特性を変更或いは改良することができる。DNAシャフリングは、PCR法による遺伝子断片の組換えで遺伝子変異体のライブラリを作製するプロセスである。次に、このライブラリを、目的の特性を有する遺伝子変異体を同定するために選択或いはスクリーニングする。これらの好ましい変異体をプールし、DNAシャフリング及び選択/スクリーニングを繰り返す。従って、人工的な育種及び急速な分子の進化によって多様な遺伝子が作られる。例えば、ランダムな位置に変異がある1つの遺伝子の断片を、目的の特性が最適化するまで、組換え及びスクリーニング、シャフリングを実施することもできる。別法では、所定の遺伝子の断片を、同じ或いは異なった種の同じ遺伝子ファミリーの相同な遺伝子の断片で組換え、それによってプロトコルに従った調節可能な方法で、多数の天然遺伝子の遺伝子多様性を最大にすることができる。

10

【0139】

別の実施例によれば、RMEP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RMEP自体またはその断片を合成することが可能である。例えば、ペプチド合成は種々の固相技術を用いて実行可能である(例えば、Creighton, T. (1984) Proteins. Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; Roberge, J. Y. ら (1995) Science 269:202-204を参照)。また、合成の自動化は例えばABI 431Aペプチドシンセサイザー(PE Biosystems)を用いて達成し得る。更にRMEPのアミノ酸配列または任意のその一部は、直接的な合成の際の変更、及び/または化学的方法を用いた他のタンパク質または任意のその一部からの配列との組み合わせにより、天然のポリペプチド配列を有するポリペプチドまたは変異体ポリペプチドを作製することが可能である。

20

30

【0140】

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

【0141】

生物学的に活性なRMEPを発現させるために、RMEPをコードするヌクレオチド配列またはその誘導体を好適な発現ベクターに挿入する。この発現ベクターは、好適な宿主に挿入されたコーディング配列の転写及び翻訳の調節に必要なエレメントを含む。これらのエレメントには、ベクター及びRMEPをコードするポリヌクレオチド配列におけるエンハンサー、構成型及び発現誘導型のプロモーター、5'及び3'の非翻訳領域などの調節配列が含まれる。このようなエレメントは、その長さ及び特異性が様々である。特定の開始シグナルによって、RMEPをコードする配列のより効果的な翻訳を達成することが可能である。このようなシグナルには、ATG開始コドン及びコザック配列などの近傍の配列が含まれる。RMEPをコードする配列及びその開始コドン、上流の調節配列が好適な発現ベクターに挿入された場合は、更なる転写調節シグナルや翻訳調節シグナルは必要なくなるであろう。しかしながら、コーディング配列或いはその断片のみが挿入された場合は、インフレームのATG開始コドンを含む外来性の翻訳調節シグナルが発現ベクターに含まれなければならない。外来性の翻訳要素及び開始コドンは、自然及び合成の様々なも

40

50

のから得ることが可能である。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である。(例えば、Scharf, D. 他 (1994) Results Probl. Cell Differ. 201-18-162. を参照)。

【0142】

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

10

【0143】

種々の発現ベクター/宿主系を利用して、RMEPをコードする配列の保持及び発現が可能である。これらには、限定するものではないが、組換えバクテリオファージ、プラスミド、またはコスミドDNA発現ベクターで形質転換された細菌などの微生物や、酵母菌発現ベクターで形質転換された酵母菌や、ウイルス発現ベクター(例えば、バキュロウイルス)に感染した昆虫細胞系や、ウイルス発現ベクター(例えば、カリフラワーモザイクウイルス、CaMV; タバコモザイクウイルス、TMV)または細菌発現ベクター(例えば、TiまたはpBR322プラスミド)で形質転換された植物細胞系や、動物細胞系などが含まれる(例えば、前出のSambrook、前出のAusubel、Van Heeke, G. および S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509、Engelhard, E.K. 他 (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227、Sandig, V. 他 (1996) Hum. Gene Ther. 7:1937-1945、Takamatsu, N. (1987) EMBOJ. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196、Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659、Harrington, J.J. 他 (1997) Nat. Genet. 15:345-355を参照)。レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ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. and N. Somia (1997) Nature 389:239-242等を参照)。本発明は使用される宿主細胞によって限定されるものではない。

20

30

40

【0144】

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、RMEP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、RMEPをコードするポリヌクレオチド配列の日常的なクローニング、サブクローニング、増殖には、PBLUESCRIPT(Stratagene, La Jolla CA)またはpSPORT1プラスミド(GIBCO BRL)などの多機能の大腸菌ベクターを用いることができる。ベクターの多数のクローニング部位にRMEPをコードする配列をライゲーションす

50

るとlacZ遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。更に、これらのベクターを用いて、クローニングされた配列の*in vitro*での転写、ジデオキシンスクリーニング、ヘルパーファージによる一本鎖の救出、入れ子状態の欠失を作り出すことが可能である（例えば、Van Heeke, G. 及び S. M. Schuster (1989) J. Biol. Chem. 264: 5503 - 5509. を参照）。例えば、抗体の産生のためなどに多量のRMEPが必要な場合は、RMEPの発現をハイレベルで誘導するベクターが使用できる。例えば、強力に発現を誘発するT5またはT7バクテリオファージプロモーターを含むベクターを使用できる。

【0145】

RMEPの発現に酵母の発現系の使用が可能である。因子やアルコールオキシダーゼやPGHプロモーターなどの構成型或いは誘導型のプロモーターを含む多種のベクターが、酵母菌サッカロミセス・セレビジエまたは*Pichia pastoris*に使用可能である。更に、このようなベクターは、発現したタンパク質の分泌か細胞内への保持のどちらかを誘導し、安定した増殖のために宿主ゲノムの中に外来配列を組み込む。（例えば、Ausubel, 1995, 前出、Bitter, G. A. ら (1987) Methods Enzymol. 153: 516 - 544、及び Scorer, C. A. ら (1994) Bio/Technology 121 - 181 - 184. を参照）。

【0146】

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

【0147】

哺乳動物細胞では、多種のウイルスベースの発現系が利用され得る。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RMEPをコードする配列を結合し得る。ウイルスのゲノムの非必須のE1またはE3領域への挿入により、感染した宿主細胞にRMEPを発現する生ウイルスを得ることが可能である（Logan, J. 及び Shenk, T. (1984) Proc. Natl. Acad. Sci. 81: 3655 - 3659を参照）。さらに、ラウス肉腫ウイルス(RSV)エンハンサーなどの転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させることが可能である。タンパク質を高レベルで発現させるために、SV40またはEBVを基にしたベクターを用いることが可能である。

【0148】

ヒト人工染色体(HAC)を用いて、プラスミドで発現しそれに含まれているものより大きなDNAの断片を供給可能である。治療のために約6kb~10MbのHACsを作製し、従来の輸送方法（リポソーム、ポリカチオンアミノポリマー、またはベシクル）で供給する。（例えば、Harrington, J. J. 他 (1997) Nat Genet. 15: 345 - 355. を参照）。

【0149】

哺乳動物系の組換えタンパク質の長期にわたる産生のためには、株化細胞におけるRMEPの安定した発現が望ましい。例えば、発現ベクターを用いて、RMEPをコードする配

10

20

30

40

50

列を株化細胞に形質転換することが可能である。このような発現ベクターは、ウイルス起源の複製及び／または内在性の発現要素や、同じ或いは別のベクターの上の選択マーカー遺伝子を含む。ベクターの導入の後、細胞を選択培地に移す前に、強化培地で約1～2日の間増殖させる。選択マーカーの目的は選択的な媒介物に対する抵抗性を与えるとともに、その存在により導入された配列を確実に発現する細胞の増殖及び回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に好適な組織培養技術を用いて増殖可能である。

【0150】

任意の数の選択系を用いて、形質転換された細胞系を回収することが可能である。選択系には、以下のものに限定はしないが、単純ヘルペスウイルスチミジンキナーゼ遺伝子及びアデニンホスホリボシルトランスフェラーゼ遺伝子が含まれ、それぞれ tk⁺ または apr⁺ 細胞において使用される。(例えば、Wigler, M. 他 (1977) Cell 11:223-232; 及び Lowy, I. 他 (1980) Cell 22:817-823を参照)。また代謝拮抗物質、抗生物質或いは除草剤への耐性を選択のベースとして用いることができる。例えば dhfr はメトトレキセートに対する耐性を与え、neo はアミノグリコシッドネオマイシン及び G-418 に対する耐性を与え、als 或いは pat はクロルスルフロン (chlorosulfuron)、ホスフィノトリシンアセチルトランスフェラーゼ (phosphinotricin acetyltransferase) に対する耐性を与える(例えば、Wigler, M. 他 (1980) Proc. Natl. Acad. Sci. 77:3567-3570; Colbere-Garapin, F. 他 (1981) J. Mol. Biol. 150:1-14を参照)。さらに選択に利用できる遺伝子、例えば、代謝のために細胞が必要なものを変える trpB 及び hisD が文献に記載されている(例えば、Hartman, S. C. 及び R. C. Mulligan (1988) Proc. Natl. Acad. Sci. 85:8047-51を参照)。アノトシアニン、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 GUS、ルシフェラーゼ及びその基質ルシフェリンなどの可視マーカーが用いられる。緑色蛍光タンパク質 (GFP) (Clontech, Palo Alto, CA) も使用できる。これらのマーカーを用いて、トランスフォーマントを特定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することが可能である(例えば、Rhodes, C. A. 他 (1995) Methods Mol. Biol. 55:121-131を参照)。

【0151】

マーカー遺伝子の発現の存在／非存在によって目的の遺伝子の存在が示されても、その遺伝子の存在及び発現の確認が必要な場合もある。例えば、RMEP 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、RMEP をコードする配列を含む形質転換された細胞は、マーカー遺伝子機能の欠落により特定可能である。または、1つのプロモーターの制御下でマーカー遺伝子が RMEP をコードする配列と一列に配置することも可能である。誘導または選択に応答したマーカー遺伝子の発現は、通常タンデム遺伝子の発現も示す。

【0152】

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

【0153】

特異的なポリクローナル抗体またはモノクローナル抗体のどちらかを用いる RMEP の発現の検出及び計測のための免疫学的方法は、当分野で周知である。このような技法には

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

10

20

30

40

50

【0154】

種々の標識技術及び結合技術が当業者には周知であり、様々な核酸アッセイおよびアミノ酸アッセイに用いられ得る。R M E P をコードするポリヌクレオチドに関連する配列を検出するための、標識されたハイブリダイゼーションプローブ或いは P C R 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 P C R 増幅が含まれる。別法として、R M E P をコードする配列、またはその任意の断片を m R N A 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でも可能である。当分野では周知であり市販されているこのようなベクターを、T 7, T 3, または S P 6 などの好適な R N A ポリメラーゼ及び標識されたヌクレオチドの追加によって、i n v i t r oでの R N A プローブの合成に用いることができる。これらの方法は、例えば、Amersham Pharmacia Biotech及び Promega (Madison WI)、U. S. Biochemical Corp (Cleveland OH) が市販する種々のキットを用いて行うことができる。容易な検出のために用い得る好適なレポーター分子或いは標識には、基質、コファクター、インヒビター、磁気粒子、及び放射性核種、酵素、蛍光剤、化学発光剤、色素産生剤などが含まれる。

【0155】

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

【0156】

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

【0157】

本発明の別の実施例では、R M E P 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結合させる。

例えば、市販の抗体によって認識できる異種部分を含むキメラ R M E P タンパク質が、R M E P の活性のインヒビターに対するペプチドライブラリのスクリーニングを促進し得る。また、異種タンパク質部分及び異種ペプチド部分が、市販の親和性基質を用いて融合タンパク質の精製を促進し得る。このような部分には、グルタチオン S トランスフェラーゼ (G S T)、マルトース結合タンパク質 (M B P)、チオレドキシン (T r x)、カルモジュリン結合ペプチド (C B P)、6 - H i s、F L A G、c - m c、赤血球凝集素 (H A) が含まれるが、これらに限定されるものではない。G S T 及び M B P、T r x、C B P、6 - H i s によって、固定されたグルタチオン、マルトース、フェニルアルシン酸化物 (p h e n y l a r s i n e o x i d e)、カルモジュリン、金属キレート樹脂のそれぞれで同族の融合タンパク質の精製が可能となる。F L A G、c - m c、及び赤血球凝集素 (H A) によって、これらのエピトープ標識を特異的に認識する市販のモノクローナル抗体及びポリクローナル抗体を用いた融合タンパク質の免疫親和性の精製ができる。また、R M E P をコードする配列と異種タンパク質配列との間にあるタンパク質分解切断部位を融合タンパク質が含むように遺伝子操作すると、R M E P が精製の後に異種部分から切断され得る。融合タンパク質の発現と精製の方法は、A u s u b e l . (1 9 9 5、前出 c h 1 0) に記載されている。市販されている様々なキットを用いて、融合タンパク質の発現及び精製を促進できる。

10

【 0 1 5 8 】

本発明の別の実施例では、T N T ウサギ網状赤血球可溶化液またはコムギ胚芽抽出系 (P r o m e g a) を用いて i n v i t r o で放射能標識した R M E P の合成が可能である。これらの系は、T 7 または T 3、S P 6 プロモーターと機能的に結合したタンパク質をコードする配列の転写と翻訳をつなげる。翻訳は、例えば、³⁵S メチオニンである放射能標識されたアミノ酸前駆体の存在の下で起こる。

20

【 0 1 5 9 】

本発明の R M E P またはその断片を用いて、R M E P 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、R M E P への特異的な結合をスクリーニングすることが可能である。試験化合物の例には、抗体、オリゴヌクレオチド、タンパク質 (例えば受容体) または小分子が挙げられる。

【 0 1 6 0 】

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 R M E P 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 (C o l l i g a n , J . E . 他 (1 9 9 1) C u r r e n t P r o t o c o l s i n I m m u n o l o g y 1 (2) の 5 章等を参照)。同様に、化合物は、R M E P が結合する天然受容体、或いは例えばリガンド結合部位などの少なくとも受容体のある断片に密接に関連し得る。何れの場合も、既知の技術を用いてこの化合物を合理的に設計することができる。一実施例では、このような化合物に対するスクリーニングには、分泌タンパク質或いは細胞膜上のタンパク質の何れか一方として R M E P を発現する好適な細胞の作製が含まれる。好適な細胞には、哺乳動物、酵母、大腸菌からの細胞が含まれる。R M E P を発現する細胞または R M E P を含有する細胞膜断片を試験化合物と接触させて、R M E P または化合物の何れかの結合、刺激または阻害を分析する。

30

40

【 0 1 6 1 】

あるアッセイは、単に試験化合物をポリペプチドに実験的に結合させ、結合を、フルオロフォア、放射性同位体、酵素抱合体またはその他の検出可能な標識により検出することができる。例えば、このアッセイは、少なくとも 1 つの試験化合物を、溶液中の或いは固体支持物に固定された R M E P と結合させるステップと、R M E P 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、細胞遊離剤、化学ライブラリまたは天然の生成混合物を用いて実施することができ、試験化合物は、溶液中で遊離させるか固体支持体に固定させる。

50

【0162】

本発明のRMEPまたはその断片を用いて、RMEPの活性を調整する化合物をスクリーニングすることが可能である。このような化合物には、アゴニスト、アンタゴニスト、或るいは部分的または逆アゴニスト等が含まれる。一実施例では、RMEPが少なくとも1つの試験化合物と結合する、RMEPの活性が許容される条件下でアッセイを実施し、試験化合物の存在下でのRMEPの活性が試験化合物非存在下でのRMEPの活性と比較する。試験化合物の存在下でのRMEPの活性の変化は、RMEPの活性を調整する化合物の存在を示唆する。別法では、試験化合物をRMEPの活性に適した条件下でRMEPを含む in vitro または細胞遊離系と結合させてアッセイを実施する。これらアッセイの何れかにおいて、RMEPの活性を調整する試験化合物は間接的に結合することが可能であり、試験化合物と直接接触する必要がない。少なくとも1つから複数の試験化合物をスクリーニングすることができる。

10

【0163】

別の実施例では、胚性幹細胞（ES細胞）における相同組換えを用いて動物モデル系内で、RMEPまたはその哺乳動物相同体をコードするポリヌクレオチドを「ノックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの作製に有用である（米国特許第5,175,383号及び第5,767,337号等を参照）。例えば129/SvJ細胞株等のマウス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した目的の遺伝子を含むベクターで形質転換する。このベクターは、相同組換えにより宿主ゲノムの対応する領域に組み込まれる。別法では、Cre-loxP系を用いて相同組換えを行い、組織特異的または発生段階特異的に目的遺伝子をノックアウトする（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、潜在的な治療薬や毒性薬剤で検査することができる。

20

30

【0164】

RMEP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）。

【0165】

RMEPをコードするポリヌクレオチドを用いて、ヒト疾患をモデルとした「ノックイン」ヒト化動物（ブタ）または遺伝子組換え動物（マウスまたはラット）を作製することが可能である。ノックイン技術を用いて、RMEPをコードするポリヌクレオチドの或る領域を動物ES細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について研究し、潜在的な医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばRMEPを乳汁内に分泌するなどRMEP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）。

40

【0166】

（治療）

RMEPのある領域とRNA代謝タンパク質のある領域との間に、例えば配列及びモチーフの文脈における化学的及び構造的類似性が存在する。更に、RMEPの発現は肺の疾患

50

や、無酸素症を引き起こす生理的病態に密接に関連するだけでなく、増殖組織、腫瘍組織、神経組織、副腎組織、脳腫瘍組織、胎児結腸組織、成人結腸組織、前立腺上皮組織、リンパ節癌組織、卵巣組織、膵臓組織、及び胎児脾臓組織に密接に関連する。従って、RMEPは、神経系の疾患、自己免疫/炎症性の疾患、細胞増殖異常、及び発生または発達障害、更に肺特異的組織を含む新生物においてある役割を果たすと考えられる。RMEP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、RMEPの発現または活性を低下させることが望ましい。また、RMEP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、RMEPの発現または活性を増大させることが望ましい。

【0167】

従って、一実施例において、RMEP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RMEPまたはその断片や誘導体を投与することが可能である。限定するものではないが、このような疾患には神経系の疾患、自己免疫/炎症性の疾患、細胞増殖異常、及び発生または発達障害が含まれ、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、ゲルストマン症候群、Gerstmann-Strausler-Scheinker症候群を含むプリオン病(prion disease)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、中枢神経系性精神薄弱及び他の発生障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィー及び他の神経筋障害、末梢神経疾患、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、気分性及び不安性精神障害、及び妄想性精神病と、季節性の感情の障害(SAD)、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、带状疱疹後神経痛、及びトゥーレット病が含まれ、自己免疫/炎症性の疾患の中には、炎症及び日光性角化症、後天性免疫不全症候群(AIDS)及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィ(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時的リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、膵炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病(MCTD)、骨髄線維症、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、子宮の癌などが含まれ、発生または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、癲癇、性腺形成異常、WAGR症候群(ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神薄弱)、スミス マジェニス症候群(Smith-Magenis syndrome)、脊髄形成異常症候群

、遺伝性粘膜上皮異形成、遺伝性角皮症、シャルコー マリー 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、Syndenham 舞蹈病 (Syndenham's chorea) 及び脳性小児麻痺などの発作障害、脊髄二分裂、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感覚神経性聴力損失が含まれる。

【0168】

別の実施例では、限定するものではないが上に列記した疾患を含む RMEP 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、RMEP またはその断片や誘導体を発現し得るベクターを患者に投与することも可能である。

【0169】

更に別の実施例では、限定するものではないが上に列記した疾患を含む RMEP 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 RMEP を含む組成物を好適な医薬用担体と共に患者に投与することも可能である。

【0170】

更に別の実施例では、限定するものではないが上に列記した疾患を含む RMEP 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、RMEP の活性を調節するアゴニストを患者に投与することも可能である。

【0171】

更なる実施例では、RMEP 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 RMEP のアンタゴニストを投与することが可能である。限定するものではないが、このような疾患の例には、上記した神経系の疾患、自己免疫 / 炎症性の疾患、細胞増殖異常、及び発生または発達障害が含まれる。一実施態様では、RMEP 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 RMEP 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運搬機構として間接的に用いられ得る。

【0172】

別の実施例では、限定するものではないが上に列記した疾患を含む RMEP 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、RMEP をコードするポリヌクレオチドの相補配列を発現するベクターを患者に投与することも可能である。

【0173】

別の実施例では、本発明の任意のタンパク質、アンタゴニスト、抗体、アゴニスト、相補的な配列、ベクターを別の好適な治療薬と組み合わせて投与することもできる。当業者は、従来の医薬原理にしたがって併用療法で用いる好適な治療薬を選択可能である。治療薬との組み合わせにより、上に列記した種々の疾患の治療または予防に相乗効果をもたらし得る。この方法を用いて少ない量の各薬剤で医薬効果をあげることが可能であり、広範囲な副作用の可能性を低減し得る。

【0174】

RMEP のアンタゴニストは、当分野で一般的な方法を用いて製造することが可能である。詳しくは、精製された RMEP を用いて抗体を作ったり、治療薬のライブラリをスクリーニングして RMEP と特異的に結合するものを同定が可能である。RMEP 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 フラグメント、及び Fab 発現ライブラリによって作られたフラグメントが含まれる。但し、これらに限定されるものではない。治療用には、中和抗体 (即ち、二量体の形成を阻害するもの) が特に好ましい。

【0175】

抗体の産生のためには、ヤギ、ウサギ、ラット、マウス、ヒト及びその他のものを含む種々の宿主が、RMEP または任意の断片、または免疫原性の特性を備えるそのオリゴペプチドの注入によって免疫化され得る。宿主の種に応じて、種々のアジュバントを用いて免疫応答を高めることもできる。このようなアジュバントにはフロイントアジュバント、水酸化アルミニウムなどのミネラルゲルアジュバント、リゾレシチン、プルロニックポリオール、ポリアニオン、ペプチド、油性乳剤、キーホールリンペットヘモシニアン、及びジニ

10

20

30

40

50

トロフェノールなどの界面活性剤が含まれるが、これらに限定されるものではない。ヒトに用いられるアジュバントの中では、BCG (b a c i l l i C a R M E P t t e - G u e r i n) 及び C o r y n e b a c t e r i u m p a r v u m が特に好ましい。

【0176】

RMEPに対する抗体を誘発するために用いられるオリゴペプチド、ペプチド、または断片は、少なくとも約5個のアミノ酸からなり、一般的には約10個以上のアミノ酸からなるものが好ましい。これらのオリゴペプチド或いはペプチド、またはそれらの断片は、天然のタンパク質のアミノ酸配列の一部と同一であることが望ましい。RMEPアミノ酸の短いストレッチは、KLHなどの別のタンパク質の配列と融合し、キメラ分子に対する抗体が産生され得る。

10

【0177】

RMEPに対するモノクローナル抗体は、培地内の連続した細胞株によって、抗体分子を産生する任意の技術を用いて作製することが可能である。これらの技術には、ハイブリドーマ技術、ヒトB細胞ハイブリドーマ技術、及びEBV-ハイブリドーマ技術が含まれるが、これらに限定されるものではない(例えば、Kohler, G. ら. (1975) Nature 256:495-497; Kozbor, D. ら. (1985) J. Immunol. Methods 81-8-42; Cote, R. J. ら. (1983) Proc. Natl. Acad. Sci. 80:2026-2030; Cole, S. P. ら. (1984) Mol. Cell Biol. 62:109-120を参照)。

20

【0178】

更に、「キメラ抗体」作製のために発達したヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの技術が、好適な抗原特異性及び生物学的活性を備える分子を得るために用いられる(例えば、Morrison, S. L. 他. (1984) Proc. Natl. Acad. Sci. 81-4851-4855; Neuberger, M. S. 他. (1984) Nature 312:604-608; Takeda, S. ら. (1985) Nature 314:452, 454を参照)。別法では、当分野で周知の方法を用いて、一本鎖抗体の産生のための記載された技術を適用して、RMEP特異性一本鎖抗体を生成する。関連する特異性を備えるが別のイディオタイプの組成の抗体は、ランダムな組み合わせの免疫グロブリンライブラリから鎖混合によって生成することもできる(例えば、Burton D. R. (1991) Proc. Natl. Acad. Sci. 88:11120-3を参照)。

30

【0179】

抗体は、リンパ球集団の中の i n v i v o 産生を誘発することによって、または免疫グロブリンライブラリのスクリーニングまたは文献に示されているような、高度に特異的な結合試薬のパネルをスクリーニングすることによって、得ることもできる(例えば、Orlandi, R. 他. (1989) Proc. Natl. Acad. Sci. 86:3833-3837; Winter, G. 他. (1991) Nature 349:293-299を参照)。

【0180】

RMEPに対する特異的な結合部位を含む抗体も得ることができる。例えば、限定するものではないが、抗体分子のペプシン消化によって生成されるF(ab')₂断片、及びF(ab')₂断片のジスルフィド架橋を減少させて生成されるFab断片が含まれる。別法では、所望の特異性を有するモノクローナルFab断片の迅速且つ容易な同定が可能となるように、Fab発現ライブラリを作製することもできる(例えば、Huse, W. D. ら. (1989) Science 254:1275-1281を参照)。

40

【0181】

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。確立された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる競合的な結合アッセイ、または免疫放射線アッセイの様々なプロトコルが、当分野で

50

は周知である。通常このようなイムノアッセイでは、RMEPとその特異的抗体との間で形成された複合体の測定が含まれる。二つの非干渉性RMEP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)。

【0182】

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

10

20

【0183】

ある下流での適用におけるこのような医薬品の品質及び適性を調べるために、ポリクローナル抗体医薬の抗体価及び結合活性を更に評価する。例えば、少なくとも $1 \sim 2 \text{ mg/ml}$ の特異的な抗体、好ましくは $5 \sim 10 \text{ mg/ml}$ の特異的な抗体を含むポリクローナル抗体医薬は一般に、RMEP抗体複合体を沈殿させなければならない処理に用いられる。様々な適用例における抗体の特異性及び抗体価、結合活性、抗体の品質や使用法の指針は一般に入手可能である。(例えば、Catty, 前出, 及びColligan 他、前出を参照)。

【0184】

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

30

【0185】

治療に用いる場合、アンチセンス配列を好適な標的細胞に導入するのに好適な任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を発現する発現プラスミドの形で細胞内に送達することができる(例えば、Slater, J. E. 他(1998) J. Allergy Clin. Immunol. 102(3): 469-475; and Scanlon, K. J. 他(1995) 9(13): 1288-1296. を参照)。また、アンチセンス配列は、例えばレトロウイルスやアデノ関連ウイルスベクター等のウイルスベクターを用いて細胞内に導入することもできる(例えば、Miller, A. D. (1990) Blood 76: 271; Ausubel, 前出; Ucker, W. and W. Walther (1994) Pharmacol. Ther. 63(3): 323-347を参照)。その他の遺伝子送達機構には、リボソーム系、人工的なウイルスエンベロープ、及び当分野で周知のその他の系が含まれる(Rossi, J. J. (1995) Br. Med. Bull. 51(1): 217-225;

40

50

Boado, R. J. 他 (1998) J. Pharm. Sci. 87(11): 1308-1315; and Morris, M. C. 他 (1997) Nucleic Acids Res. 25(14): 2730-2736. を参照)。

【0186】

本発明の別の実施例では、RMEPをコードするポリヌクレオチドを、体細胞若しくは生殖細胞の遺伝子治療に用いることが可能である。遺伝子治療は、(i) 遺伝子欠損症（例えば、X染色体連鎖遺伝 (Cavazzana-Calvo, M. 他 (2000) Science 288: 669-672) によって特徴づけられる重度の複合型免疫欠損 (SCID) - X1)、遺伝性アデノシン-デアミナーゼ (ADA) 欠損症 (Blaese, R. M. 他 (1995) Science 270: 475-480; Bordignon, C. 他 (1995) Science 270: 470-475) に関連する重度の複合型免疫欠損、嚢胞性繊維症 (Zabner, J. 他 (1993) Cell 75: 207-216; Crystal, R. G. 他 (1995) Hum. Gene Therapy 6: 643-666; Crystal, R. G. 他 (1995) Hum. Gene Therapy 6: 667-703)、サラセミア (thalassaemia)、家族性高コレステロール血症、第VIII因子若しくは第IX因子欠損による血友病 (Crystal, R. G. (1995) Science 270: 404-410; Verma, I. M. and Somia, N. (1997) Nature 389: 239-242) を治療したり、(ii) 条件的致死性遺伝子産物（例えば、細胞増殖の制御不能による癌の場合）を発現させたり、及び (iii) 細胞内の寄生虫（例えば、ヒト免疫不全ウイルス (HIV) (Baltimore, D. (1988) Nature 335: 395-396; Poeschl, E. 他 (1996) Proc. Natl. Acad. Sci. USA. 93: 11395-11399) や、B型若しくはC型肝炎ウイルス (HBV、HCV)、Candida albicans 及び Paracoccidioides brasiliensis 等の真菌寄生虫、Plasmodium falciparum 及び Trypanosoma cruzi 等の原虫寄生体) に対する防御機能を有するタンパク質を発現させて行うことができる。RMEP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RMEP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。

【0187】

本発明の更なる実施例では、RMEPの欠損による疾患や異常症は、RMEP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RMEP欠損細胞に導入することによって治療する。in vivo 或いは ex vivo の細胞に用いる機械的な導入技術には、(i) 個々の細胞内へのDNAのマイクロインジェクション、(ii) 金粒子の打ち込み、(iii) リポソーム仲介性トランスフェクション、(iv) 受容体仲介性遺伝子導入、及び (v) DNAトランスポソン (Morgan, R. A. and W. F. Anderson (1993) Annu. Rev. Biochem. 62: 191-217; Ivics, Z. (1997) Cell 91: 501-510; Boulay, J.-L. and H. Recipon (1998) Curr. Opin. Biotechnol. 9: 445-450) の使用が含まれる。

【0188】

RMEPの発現に影響を及ぼし得る発現ベクターには、限定するものではないが、PCDNA3.1、EPITAG、PRCCMV2、PREP、PVAXベクター (Invitrogen, Carlsbad CA)、PCMV-SCRIPT、PCMV-TAG、PEGSH/PERV (Stratagene, La Jolla CA)、PTET-OFF、PTET-ON、PTRE2、PTRE2-LUC、PTK-HYG (Clontech, Palo Alto CA) が含まれる。RMEPを発現させるために、(i) 恒常的に活性なプロモーター（例えば、サイトメガロウイルス (CMV)、ラウス

肉腫ウイルス(RSV)、SV40ウイルス、チミジンキナーゼ(TK)、若しくは -
アクチン遺伝子等)、(ii)誘導性プロモーター(例えば、市販されているT-REX
プラスミド(Invitrogen)に含まれている、テトラサイクリン調節性プロモ-
ーター(Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. U.S.A. 89:5547-5551; Gossen
, M. 他(1995) Science 268:1766-1769; Rossi,
F.M.V. and H.M. Blau (1998) Curr. Opin. Bio-
technol. 9:451-456))、エクジソン誘導性プロモーター(市販され
ているプラスミドPVGRXR及びPINDに含まれている:Invitrogen)、
FK506/ラパマイシン誘導性プロモーター、またはRU486/ミフェプリストーン
誘導性プロモーター(Rossi, F.M.V. and H.M. Blau, 前出)
、または(iii)正常な個体に由来するRMEP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。

10

20

30

40

50

【0189】

市販のリボソーム形質転換キット(例えば、Invitrogenが販売しているPER
FECT LIPID及び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であ
る。別法では、リン酸カルシウム法(Graham, F.L. and A.J. Eb
(1973) Virology 52:456-467)若しくは電気穿孔法(Neu
mann, B. 他(1982) EMBO J. 1:841-845)を用いて形質転
換を行う。初代細胞にDNAを導入するためには、これらの標準的な哺乳動物トランスフ
ェクションプロトコルを変更する必要がある。

【0190】

本発明の別の実施例では、RMEPの発現に関連する遺伝子欠損によって起こる疾患や異
常症は、(i)レトロウイルス末端反復配列(LTR)プロモーター若しくは独立したプ
ロモーターのコントロール下でRMEPをコードするポリヌクレオチドと、(ii)好適
なRNAパッケージングシグナルと、(iii)追加のレトロウイルス・シス作用性RN
A配列及び効率的なベクターの増殖に必要なコーディング配列を伴うRev応答性エレメ
ント(RRE)とからなるレトロウイルスベクターを作製して治療することができる。レ
トロウイルスベクター(例えば、PFB及びPFBNEO)はStratagene社か
ら入手可能であり、公表データ(Riviere, I. 他.(1995) Proc.
Natl. Acad. Sci. U.S.A. 92:6733-6737)に基づい
ている。上記データを引用することをもって本明細書の一部とする。このベクターは、V
SVg(Armentano, D. 他(1987) J. Virol. 61:164
7-1650; Bender, M.A. 他(1987) J. Virol. 61:
1639-1646; Adam, M.A. and A.D. Miller (1988
) J. Virol. 62:3802-3806; Dull, T. 他(1998)
J. Virol. 72:8463-8471; Zufferey, R. 他(199
8) J. Virol. 72:9873-9880)等の乱交雑エンベロープタンパク
質若しくは標的細胞上の受容体に対する親和性を有するエンベロープ遺伝子を発現する好
適なベクター産生細胞系(VPCL)において増殖される。RIGGに付与された米国特
許第5,910,434号(「Method for obtaining retrov
irus packaging cell lines producing high tr
ansducing efficiency retroviral supernata
nt」)において、レトロウイルスパッケージング細胞系を得るための方法が開示されて
おり、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、あ
る細胞集団(例えば、CD4⁺T細胞)の形質導入、並びに形質導入した細胞を患者に戻
す方法は、遺伝子治療の分野では周知であり、多数の文献に記載されている(Ranga
, U. 他.(1997) J. Virol. 71:7020-7029; Baue
r, 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)。

【0191】

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、RMEPの発現に関連する1
 或いは複数の遺伝子異常を有する細胞にRMEPをコードするポリヌクレオチドを送達す
 る。アデノウイルス系ベクターの作製及びパッケージングは当分野では周知である。複製
 欠損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を脾臓の無損
 傷の脾臓の中に導入するために可変性であることが証明された(Csete, M. E. 10
 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が、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ま
 た、Antinuzzi, P. A. 他 (1999) Annu. Rev. Nutr. 19:511-544; and Verma, I. M. and N. Somia (1
 997) Nature 18:389:239-242を参照し、引用することをもって
 本明細書の一部とする。

【0192】

別法では、ヘルペス系遺伝子治療の送達系を用いて、RMEPの発現に関連する1或いは 20
 複数の遺伝子異常を有する標的細胞にRMEPをコードするポリヌクレオチドを送達する
 。単純疱疹ウイルス(HSV)系のベクターは、HSV親和性の中枢神経細胞にRMEP
 を導入する際に特に重要である。ヘルペス系ベクターの作製及びパッケージングは当分野
 では周知である。複製適格性の単純疱疹ウイルス(HSV)I型系のベクターは、霊長類
 の眼にレポーター遺伝子を送達するために用いられてきた(Liu, X. 他 (199
 9) Exp. Eye Res. 169:385-395)。HSV-1ウイルスベクタ
 ーの作製は、DeLucaに付与された米国特許第5,804,413号(Herpes
 simplex virus swains for gene transfer)に記載
 されており、引用することをもって本明細書の一部とする。米国特許第5,804,41
 3号には、ヒト遺伝子治療を含む目的のために、好適なプロモーターのコントロールの下 30
 で、細胞に導入される少なくとも1つの内在性遺伝子を含むゲノムからなる組換えHSV
 d92についての記載がある。また上記特許には、ICP4、ICP27及びICP2
 2のために除去される組換えHSV株の作製及び使用方法が開示されている。HSVベク
 ターについては、Goins, W. F. 他 (1999) J. Virol. 73:5
 19-532 and Xu, H. 他 (1994) Dev. Biol. 163:15
 2-161を参照し、引用することをもって本明細書の一部とする。クローニングされた
 ヘルペスウイルス配列の操作や、巨大ヘルペスウイルスのゲノムの異なった部分を含む多
 数のプラスミドをトランスフェクトした後の組換えウイルスの継代、ヘルペスウイルスの
 成長及び増殖、並びにヘルペスウイルスの細胞への感染は当分野で周知の技術である。

【0193】

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RMEP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)の生物学的な研
 究が広範に行われ、遺伝子伝達ベクター(gene transfer vector)が
 SFVゲノムに基づいていることが分かった(Garoff, H. and K. - J.
 Li (1998) Cun. Opin. Biotech. 9:464-469)。
 ウイルスRNAの複製中に、通常はウイルスカプシドタンパク質をコードするサブゲノム
 RNAが作り出される。このサブゲノムRNAが完全長のゲノムRNAより高いレベルで
 複製されるため、酵素活性(例えばプロテアーゼ及びポリメラーゼ)を有するウイルスタン
 パク質に対してカプシドタンパク質が過剰に産生される。同様に、RMEPをコードす 50

る配列を ウイルスゲノムのカプシドをコードする領域に導入することによって、ベクター導入細胞において多数の R M E P をコードする R N A が産生され、高いレベルで R M E P が合成される。通常は ウイルス感染は 2 ~ 3 日以内の細胞溶解に関係するが、シンドビスウイルス (S I N) の変異体を有するハムスターの正常な腎細胞 (B H K - 2 1) の持続的な感染を確立する能力は、 ウイルスの溶解性の複製が遺伝子治療に適用できるように好適に変更することが可能であることを示唆している (D r y g a , S . A . 他 . (1 9 9 7) V i r o l o g y 2 2 8 : 7 4 - 8 3) 。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 R M E P を導入することができる。ある集団における細胞のサブセットの特定の形質導入には、形質導入する前に細胞のソーティングを必要とする場合がある。 ウイルスの感染性 c D N A クローンの操作、 ウイルス c D N A 及び R N A のトランスフェクション、並びに ウイルスの感染方法は当分野で周知である。

10

【 0 1 9 4 】

例えば開始部位から約 - 1 0 から約 + 1 0 までの転写開始部位に由来するオリゴヌクレオチドを用いて、遺伝子の発現を阻害することが可能である。同様に、三重らせん塩基対合法を用いて阻害することができる。三重らせん構造は、二重らせんがポリメラーゼ、転写因子、または調節分子の結合のために十分に広がるのを阻止するため有用である。三重鎖 D N A を用いる最近の治療の進歩は文献に記載されている (例えば、G e e , J . E . ら . (1 9 9 4) I n : H u b e r , B . E . 及び B . I . C a r r , M o l e c u l a r a n d I m m u n o l o g i c A p p r o a c h e s , F u t u r a P u b l i s h i n g C o . , M t . K i s c o , N Yを参照) 。相補的な配列またはアンチセンス分子もまた、転写物がリボソームに結合するのを阻止することによって m R N A の翻訳を阻止するように設計できる。

20

【 0 1 9 5 】

酵素性 R N A 分子であるリボザイムは、R N A の特異的切断を触媒するために用いることができる。リボザイム作用の機構には、相補的な標的 R N A へのリボザイム分子の配列特異性ハイブリダイゼーションが含まれ、ヌクレオチド鎖切断が続く。例えば、R M E P をコードする配列のヌクレオチド鎖切断を、特異的且つ効果的に触媒する組換え型のハンマーヘッド型リボザイム分子が含まれる。

【 0 1 9 6 】

任意の潜在的 R N A 標的内の特異的なリボザイム切断部位が、後続の配列 G U A 、 G U U 、及び G U C を含むリボザイム切断部位に対して、標的分子をスキヤニングすることによって初めに同定される。一度同定されると、切断部位を含む標的遺伝子の領域に対応する 1 5 個から 2 0 個のリボヌクレオチドの短い R N A 配列を、オリゴヌクレオチドの機能を不全にする二次的な構造の特徴について評価することが可能である。候補標的の適合性も、リボヌクレアーゼ保護アッセイを用いて、相補的なオリゴヌクレオチドとのハイブリダイゼーションの容易性をテストすることによって評価することが可能である。

30

【 0 1 9 7 】

本発明の相補的なリボ核酸分子及びリボザイムは、当分野で周知の方法を用いて、核酸分子の合成のために作製することができる。これらの方法には、固相ホスホラミダイト化合物などのオリゴヌクレオチドを化学的に合成する方法が含まれる。別法では、R N A 分子が i n v i t r o 及び i n v i v o で R M E P をコードする D N A 配列の転写によって生成され得る、このような D N A 配列は T 7 または S P 6 等の好適な R N A ポリメラーゼプロモータを用いて、種々のベクターの中に組み入れることが可能である。別法では、相補的な R N A を構成的または誘導的に合成するこれらの c D N A 作製物は、細胞株、細胞、または組織の中に導入することができる。

40

【 0 1 9 8 】

R N A 分子を修飾することによって、細胞内の安定性を高め、半減期を長くすることができる。可能な修飾には、分子の 5 ' 及び / または 3 ' 端部でのフランキング配列の追加、または分子のバックボーン内のホスホジエステル結合の代わりにホスホロチオネートまた

50

は2' Oメチルを用いる修飾が含まれるが、これらに限定されるものではない。PNAの生成に固有のこの概念は、内在性のエンドヌクレアーゼによって容易には認識されないアデニン、シチジン、グアニン、チミン、及びウリジンのアセチル-、メチル-、チオ-、及び同様の修飾形態だけでなく、イノシン、キューエオシン (queuosine)、及びワイブトシン (wybutosine) などの従来のものでない塩基を含めることによって、これらの分子の全体に拡大することができる。

【0199】

本発明の更なる実施例は、RMEPをコードするポリヌクレオチドの発現の変化に有効な化合物をスクリーニングする方法を含む。特定のポリヌクレオチドの発現の変化に有効な化合物には、限定するものではないが、特定のポリヌクレオチド配列と相互作用可能な非高分子化学物質、オリゴヌクレオチド、アンチセンスオリゴヌクレオチド、三重らせん形成オリゴヌクレオチド、転写因子やその他のポリペプチド転写調節因子が含まれる。有効な化合物は、ポリヌクレオチド発現のインヒビター或いはエンハンサーとして作用し、ポリヌクレオチドの発現を変化させ得る。従って、RMEP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、RMEP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、RMEP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、RMEP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。

【0200】

特定のポリヌクレオチドの発現の変化の有効性を調べるために、少なくとも1個から複数個の試験化合物をスクリーニングすることができる。試験化合物は、有効な化合物の化学修飾を含む当分野で周知の任意の方法で得ることができる。このような方法は、ポリヌクレオチドの発現を変化させる場合、一般に市販されている或いは専売の天然または非天然の化合物ライブラリから選択する場合、標的ポリヌクレオチドの化学的及び/または構造的な特性に基づいて化合物を合理的にデザインする場合、更に組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合に有効である。RMEPをコードするポリヌクレオチドを含むサンプルは、少なくとも1つの試験化合物に曝露して得る。サンプルには、例えば無傷細胞、透過化処理した細胞、in vitro細胞遊離系または再構成生化学系が含まれ得る。RMEPをコードするポリヌクレオチドの発現における変化は、当分野で周知の任意の方法でアッセイする。通常、RMEPをコードするポリヌクレオチドの配列に相補的なヌクレオチド配列を有するプローブを用いたハイブリダイゼーションにより、特定のヌクレオチドの発現を検出する。ハイブリダイゼーションの収量を定量し、その値が1或いは複数の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較における基準となり得る。試験化合物に曝露されるポリヌクレオチドの発現の変化が検出される場合は、ポリヌクレオチドの発現の変化に試験化合物が有効であることを示している。特定のポリヌクレオチドの発現の変化に有効な化合物を調べるために、例えば Schizosaccharomyces pombe 遺伝子発現系 (Atkins, D. 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細胞等のヒト細胞株 (Clarke, M.L. 他 (2000) *Biochem. Biophys. Res. Commun.* 268:8-13) を用いてスクリーニングする。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性を調べるための、各オリゴヌクレオチド (デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、及び修飾オリゴヌクレオチド) の組み合わせライブラリのスクリーニングを含む (Bruce, T.W. 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)。

【0201】

ベクターを細胞または組織に導入する多数の方法が利用でき、in vivo、in vitro、及び ex vivo での使用に等しく適している。ex vivo での治療の場合、患者から採取された肝細胞の中にベクターを導入して、自家移植で同じ患者に戻すため

10

20

30

40

50

にクローニング増殖される。トランスフェクション、リボソーム注入またはポリカチオンアミノポリマーによる運搬は、当分野で周知の方法を用いて実行することができる（例えば、Goldman, C. K. 他. (1997) Nature Biotechnology 15: 462 - 66: を参照）。

【0202】

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

【0203】

本発明の別の実施例は、上記した全ての治療効果のために、医学上認められる担体と共に医薬品或いは無菌組成物の投与に関連する。このような組成物は、RMEP、RMEPの抗体、擬態、アゴニスト、アンタゴニスト、またはRMEPのインヒビターなどからなる。この組成物は、単体で、或いは安定剤などの1種類以上の別の薬剤と共に、無菌の生体適合性医薬品担体に投与することができる。このような医薬品担体には、生理食塩水、緩衝食塩水、ブドウ糖、及び水などが含まれるがこれらに限定されるものではない。この組成物は、単独或いは薬物またはホルモンなどの別の薬剤と共に投与することができる。

10

【0204】

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

【0205】

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

20

【0206】

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

30

【0207】

RMEPまたはその断片を含む高分子を直接細胞内に輸送するべく、特殊な形態に組成物が調製されるのが好ましい。例えば、細胞不透過性高分子を含むリボソーム製剤は、細胞融合及び高分子の細胞内輸送を促進し得る。別法では、RMEPまたはその断片をHIV Tat-1タンパク質の陽イオンN末端部に結合することもできる。このようにして作製された融合タンパク質は、マウスモデル系の脳を含む全ての組織の細胞に形質導入されることが確認されている（Schwarze, S. R. 他 (1999) Science 285: 1569 - 1572）。

40

【0208】

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

【0209】

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RMEPまたはその断片、RMEPの抗体、RMEPのアゴニストまたはアンタゴニスト、インヒビターなどの活性処方成分の量に関連する。薬用有効度及び毒性は、たとえば、ED₅₀（服用に対して

50

集団の50%に医薬的効果がある用量)またはLD₅₀(服用に対して集団の50%に致命的である用量)統計を計算するなど、細胞培養または動物実験における標準的な薬剤手法によって決定することができる。毒性効果と治療効果との薬用量比は治療指数であり、LD₅₀/ED₅₀と示すことができる。高い治療指数を示す組成物が望ましい。細胞培養アッセイ及び動物実験から得られたデータが、ヒトへの適用のために、薬用量の範囲を調剤するのに用いられる。このような組成物が含まれる薬用量は、毒性を殆ど或いは全く含まず、ED₅₀を含む血中濃度の範囲であることが望ましい。薬用量は、用いられる投与形態及び患者の感受性、投与の経路によって、この範囲内で様々である。

【0210】

正確な薬用量は、治療が必要な患者に関する要素を考慮して、実務者によって決められるであろう。薬用量及び投与は、効果的なレベルの活性成分を与えるため或いは所望の効果を維持するために調節される。薬用量の要素として考慮されるものには、疾患の重症度、患者の一般的な健康状態、年齢、体重、及び患者の性別、投与の時間及び頻度、併用する薬剤、反応感受性、及び治療に対する応答が含まれる。作用期間が長い組成物は、三日か四日に一度、一週間に一度、二週間に一度、特定の製剤の半減期及びクリアランス率によって左右され、投与され得る。

10

【0211】

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

20

【0212】

(診断)

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

30

【0213】

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

40

【0214】

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

【0215】

50

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

【0216】

プローブはまた、関連する配列の検出に利用され、RMEP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: 48-94の配列、或いはRMEP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。

10

【0217】

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

20

【0218】

RMEPをコードするポリヌクレオチド配列を用いて、RMEPの発現に関連する疾患を診断することが可能である。限定するものではないが、このような疾患には神経系の疾患、自己免疫/炎症性の疾患、細胞増殖異常、及び発生または発達障害が含まれ、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、ゲルストマン症候群、Gerstmann-Straussler-Scheinker症候群を含むプリオン病(prion disease)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、中枢神経系性精神薄弱及び他の発生障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィー及び他の神経筋障害、末梢神経疾患、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、気分性及び不安性精神障害、及び妄想性精神病と、季節性の感情の障害(SAD)、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、带状疱疹後神経痛、及びトウレット病が含まれ、自己免疫/炎症性の疾患の中には、炎症及び日光性角化症、後天性免疫不全症候群(AIDS)及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィ(APECD)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時的リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性

30

40

50

硬化症、重症筋無力症、心筋または心膜炎症、骨関節炎、骨粗しょう症、膵炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病(MCTD)、骨髄線維症、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、子宮の癌などが含まれ、発生または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、癲癇、性腺形成異常、WAGR症候群(ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神薄弱)、スミス マジェニス症候群(Smith-Magenis syndrome)、脊髄形成異常症候群、遺伝性粘膜上皮異形成、遺伝性角皮症、シャルコー マリー 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、Syndenham 舞蹈病(Syndenham's chorea)及び脳性小児麻痺などの発作障害、脊髄二分裂、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感覚神経性聴力損失が含まれる。RMEPをコードするポリヌクレオチド配列は、サザン法やノーザン法、ドットプロット法、或いはその他の膜系の技術、PCR法、ディップスティック(dipstick)、ピン(pin)、ELISA式アッセイ、及び変異RMEPの発現を検出するために患者から採取した体液或いは組織を利用するマイクロアレイに使用することが可能である。このような質的或いは量的方法は、当分野では周知である。

10

20

30

40

50

【0219】

ある実施態様では、RMEP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。RMEPをコードするヌクレオチド配列は、標準的な方法で標識化され、ハイブリダイゼーション複合体の形成に好適な条件の下、患者から採取した体液或いは組織のサンプルに加えることができるであろう。好適な培養期間の後、サンプルを洗浄し、シグナルを定量して基準値と比較する。患者のサンプルのシグナルの量が、制御サンプルと較べて著しく変わっている場合は、サンプル内のRMEPをコードするヌクレオチド配列の変異レベルにより、関連する疾患の存在が明らかになる。このようなアッセイを用いて、動物実験、臨床試験、或いは個人の患者の治療を監視における、特定の治療効果を推定することが可能である。

【0220】

RMEPの発現に関連する疾患の診断の基準となるものを提供するために、発現の正常すなわち標準的なプロファイルが確立される。これは、ハイブリダイゼーション或いは増幅に好適な条件の下、動物或いはヒトの何れかの正常な被験者から抽出された体液或いは細胞と、RMEPをコードする配列或いはその断片とを結合させることにより達成され得る。標準的なハイブリダイゼーションは、正常な被験者から得た値と周知の量の実質的に精製されたポリヌクレオチドが用いられる実験からの値とを比較することによって定量可能である。正常なサンプルから得た標準的な値を、疾患の症状を示す被験者から得た値と比較可能である。基準値と被験者の値との偏差を用いて罹患しているかどうかを決定する。

【0221】

疾患の存在が確定され、治療プロトコルが開始されると、ハイブリダイゼーションアッセイを通常ベースで繰り返して、被験者における発現のレベルが正常な患者に示される値に近づき始めたかどうかを推定することが可能である。繰り返し行ったアッセイの結果を、数日から数ヶ月の期間の治療の効果を見るのに用いることができる。

【0222】

癌では、個体からの生体組織における異常な量の転写物が、疾患の発生の素因を示し、また実際に臨床的症状が出る前に疾患を検出する方法を提供することが可能である。この種

のより明確な診断により、医療の専門家が予防方法或いは積極的な治療法を早くから利用して、癌の発生または進行を防ぐことが可能となる。

【0223】

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

10

【0224】

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

20

30

【0225】

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

【0226】

更に別の実施例では、本明細書で記載した任意のポリヌクレオチド配列に由来するオリゴヌクレオチドまたはより長い断片を、マイクロアレイにおける標的として用いることができる。マイクロアレイを、上記したように多数の遺伝子の相対的な発現レベルを同時にモニタリングする転写イメージング技術に用いることができる。マイクロアレイはまた、遺伝子変異、突然変異及び多型の同定に用いることができる。この情報を用いて、遺伝子機能を決定し、疾患の遺伝的根拠を解明し、疾患を診断し、遺伝子発現に関連する疾病の進行/後退をモニタリングし、疾患の治療における治療薬の開発や活性のモニタリングを行うことができる。特に、患者にとって最適かつ有効な治療法を選択するために、この情報を用いて患者の薬理ゲノムプロファイルを作成することができる。例えば、患者の薬理ゲノムプロファイルに基づいて、患者に対して極めて効果的でありながら副作用を殆ど示さ

40

50

ない治療薬を選択することができる。

【0227】

別の実施例では、RMEP、RMEPの断片、RMEPに特異的な抗体をマイクロアレイ上のエレメントとして用いることができる。マイクロアレイを用いて、上記のようにタンパク質間相互作用、薬剤-標的相互作用及び遺伝子発現プロファイルをモニタリング及び測定することが可能である。

【0228】

特定の実施例は、或る組織または細胞型の転写イメージを生成する本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞型により遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所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 (Seilliamer 他、米国特許第5,840,484号の "Comparative Gene Transcript Analysis" を参照。この特許に言及することを以って本明細書の一部とする)。従って、特定の組織または細胞型の転写物または逆転写物の全てに本発明のポリヌクレオチドまたはその相補配列をハイブリダイズすることにより、転写イメージが生成され得る。或る実施例では、本発明のポリヌクレオチドまたはその相補配列がマイクロアレイ上に複数のエレメントのサブセットを構成するハイスループット型でハイブリダイゼーションさせる。結果として得られる転写イメージは、遺伝子活性のプロファイルとなり得る。

10

【0229】

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

20

【0230】

本発明のポリヌクレオチドの発現プロファイルを示す転写イメージはまた、合成化合物または天然化合物の毒性試験のみならず、in vitroモデル系及び薬剤の前臨床評価に関連して使用され得る。全ての化合物は、作用及び毒性の機構を示唆する、頻繁に分子フィンガープリント若しくは毒性シグネチャ (signature) と称されるような特徴的な遺伝子発現パターンを引き起こす (Nuwaysir, E. F. 他 (1999) Mol. Carcinog. 24:153-159、Steiner, S. and N. L. Anderson (2000) Toxicol. Lett. 112-113:467-471、また言及することを以って本明細書の一部とする)。試験化合物が、毒性を有する既知の化合物のシグネチャと同一のシグネチャを有する場合には、毒性特性を共有している可能性が高い。フィンガープリンまたはシグネチャが、より多くの遺伝子及び遺伝子ファミリーからの発現情報を含んでいれば、より有用かつ正確になる。理想としては、発現のゲノム全域にわたって測定し、最高品質のシグネチャを提供することである。任意の試験化合物によっても発現が変化しない遺伝子も同様に重要である。それは、これらの遺伝子の発現レベルを用いて残りの発現データを標準化することができるためである。標準化処理は、異なる化合物で処理した後の発現データの比較に有用である。毒性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毒性スクリーニングにおいて、全ての発現した遺伝子配列を含めることは重要でありまた望ましいことである。

30

40

【0231】

一実施例では、試験化合物の毒性は、核酸を含有する生体サンプルをその試験化合物で処

50

理して評価する。処理した生体サンプル中で発現した核酸は、本発明のポリヌクレオチドに特異的な1若しくは複数のプローブでハイブリダイズさせ、それによって本発明のポリヌクレオチドに対応する転写レベルを定量することができる。処理した生体サンプル中の転写レベルを、非処理生体サンプル中のレベルと比較する。両サンプルの転写レベルの差が、処理されたサンプル中で試験化合物が引き起こす毒性反応を示唆する。

【0232】

別の実施例は、本発明のポリペプチド配列を用いて組織または細胞型のプロテオームを分析することに関連する。「プロテオーム」という用語は、或る特定の組織または細胞型におけるタンパク質発現の包括的パターンを指す。プロテオームを構成する各タンパク質は、個々に更なる分析をすることができる。プロテオーム発現パターン即ちプロファイルは、所定の条件下で所定の時間に発現したタンパク質の数及びそれらの相対的な存在量を定量することにより分析する。従って、ある細胞のプロテオームのプロファイルは、特定の組織または細胞型のポリペプチドを分離及び分析することにより作成し得る。或る実施例では、このような分離は2次元ゲル電気泳動によって行う。この2次元ゲル電気泳動法では、まず、1次元の等電点電気泳動によりサンプルからタンパク質を分離し、次に、2次元のドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に従って分離する（前出の Steiner and Anderson）。これらのタンパク質は、通常クーマシーブルーまたはシルバーまたは蛍光染色などの染色剤を用いてゲルを染色して、分散した個別の位置にあるスポットとしてゲル中で可視化される。各タンパク質スポットの光学密度は、通常サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処理済みまたは未処理のいずれかの生体サンプルから得られる等位置にあるタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を調べる。スポット内のタンパク質は、例えば化学的または酵素的に切断した後、質量分析する標準的な方法を用いて部分的にシーケンシングする。スポット内のタンパク質の同一性は、好適には少なくとも5個の連続するアミノ酸残基であるその部分的な配列を、本発明のポリペプチド配列と比較することにより決定し得る。場合によっては、決定的なタンパク質同定のための更なる配列が得られる。

10

20

【0233】

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

30

【0234】

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

40

【0235】

別の実施例では、試験化合物の毒性は、タンパク質を含む生体サンプルをその試験化合物で処理して評価する。処理された生体サンプル中で発現したタンパク質を分離して、各タ

50

ンパク質の量が定量できるようにする。各タンパク質の量を、未処理生体サンプル中の対応するタンパク質の量と比較する。両サンプル中のタンパク質の量の差は、処理されたサンプル中の試験化合物に対する毒性反応を示唆する。個々のタンパク質は、それらのアミノ酸残基をシーケンシングし、これらの部分配列を本発明のポリペプチドと比較することで同定する。

【0236】

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

10

【0237】

当分野で周知の方法でマイクロアレイを準備して使用し、分析する。(例えば、Brennan, T. M. 他 (1995) 米国特許第5,474,796号; Schena, M. 他 (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler 他 (1995) PCT出願番号WO95/251116; Shalon, D. 他 (1995) PCT出願番号WO95/35505; Heller, R. A. 他 (1997) Proc. Natl. Acad. Sci. 94:2150-2155; 及び Heller, M. J. 他 (1997) 米国特許第5,605,662号を参照)。様々なタイプのマイクロアレイが周知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, Londonに記載されている。また、この文献を引用することを以って本明細書の一部とする。

20

【0238】

本発明の別の実施例ではまた、RMEPをコードする核酸配列を用いて、天然のゲノム配列をマッピングするのに有用なハイブリダイゼーションプローブを作製することが可能である。コーディング配列または非コーディング配列の何れかを用いることができるが、或る例では、コーディング配列より非コード配列が好ましい。例えば、多重遺伝子ファミリーのメンバー間にコーディング配列が保存されていることにより、染色体マッピング時に望ましくない交差ハイブリダイゼーションが生じる可能性がある。この配列は、特定の染色体、染色体の特定領域または人工の染色体、例えば、ヒト人工染色体(HAC)、酵母人工染色体(YAC)、細菌人工染色体(BAC)、細菌P1産物、或いは単一染色体cDNAライブラリに対してマッピングされる(Harrington, J. J. ら (1997) Nat Genet. 15:345-355、Price, C. M. (1993) Blood Rev. 7:127-134、Trask, B. J. (1991) Trends Genet. 7:149-154等を参照)。一度マッピングすると、本発明の核酸配列を用いて、例えば病状の遺伝と特定の染色体領域やまたは制限断片長多型(RFLP)の遺伝とが相関するような遺伝子連鎖地図を作成可能である(Lander, E. S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照)。

30

40

【0239】

insit 蛍光ハイブリダイゼーション(FISH)は、他の物理的及び遺伝子地図データと相関し得る(例えば、Heinz-Ulrich, 他による(1995) in Meyers, 前出, pp. 965-968を参照)。遺伝子地図データの例は、種々の科学誌あるいはOnline Mendelian Inheritance in Man(OMIM)のワールドワイドウェブのサイトで見つけることができる。物理的な染色体地図上のRMEPをコードする遺伝子の位置と特定の疾患との相関性、或いは特定の疾患に対する素因が、このような疾患と関連するDNA領域の決定に役立つため、更なる

50

位置を決定するクローニングが行われる。

【0240】

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

10

【0241】

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

20

【0242】

薬剤スクリーニングに用いる別の方法は、目的のタンパク質に対して、好適な結合親和性を有する化合物のスクリーニング処理能力を高めるために用いられる（例えば、Geyse, 他による(1984) PCT出願番号 WO 84/03564を参照）。この方法では、相当な数の異なる小さな試験用化合物が、プラスチックピン或いは他の基板の上に合成される。試験用化合物は、RMEP、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RMEPが、当分野で周知の方法で検出される。精製されたRMEPはまた、前記した薬剤をスクリーニングする技術に用いられるプレート上で直接被覆することもできる。別法では、非中和抗体を用いて、ペプチドを捕らえ、固体支持物に固定することもできる。

30

【0243】

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

【0244】

別の実施例では、発展途上の分子生物学技術にRMEP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。

【0245】

当分野の技術者であれば、更なる説明がなくても前述の説明だけで最大限に本発明を利用できるであろう。したがって、以下に記載する実施例は、例示目的であって本発明を限定するものではない。

40

【0246】

前述した及び以下に記載する全ての特許出願、特許、刊行物、特に米国特許出願第60/201,875号、同第60/200,184号、同第60/202,090号、同第60/210,232号、及び同第60/220,553号に言及することをもって本明細書の一部とする。

【0247】

(実施例)

50

1 cDNAライブラリの作製

インサイトcDNAはLIFESEQ GOLD データベース (Incyte Genomics, Palo Alto CA) に含まれているcDNAライブラリに由来し、表4の列5に示されている。ある組織をグアニジニウムイソチオシアネート溶液においてホモジナイズして溶解する一方、別の組織を、フェノールにおいて、或いはグアニジニウムイソチオシアネート及びフェノールの単相溶液であるTRIZOL (Life Technologies) などの変性剤の好適な混合液においてホモジナイズして溶解した。この溶解物を塩化セシウムにおいて遠心分離によって、或いはクロロホルムで抽出した。RNAは、イソプロパノール或いは酢酸ナトリウムとエタノールのどちらか、或いは別の一般的な方法でこの溶解物から沈殿させた。

10

【0248】

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

【0249】

ある場合には、Stratagene社にRNAを提供し、Stratagene社が対応するcDNAライブラリを作製した。そうでない場合は、UNIZAPベクターシステム (Stratagene) またはSUPERSCRIPT プラスミドシステム (Life Technologies) を用いて当分野で周知の推奨方法または類似の方法でcDNAを合成してcDNAライブラリを作製した。(例えば、Ausubel, 1997, 前出, ユニット5.1-6.6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させた。合成オリゴヌクレオチドアダプターを二本鎖cDNAに結合させてから、好適な1或いは複数の制限酵素でcDNAを消化した。殆どのライブラリでは、SEPHACRYL S1000、SEPHAROSE CL2B、SEPHAROSE CL4Bカラムクロマトグラフィー (Amersham Pharmacia Biotech)、またはアガロースゲル電気泳動法によってcDNAの大きさ(300 ~ 1000 bp)を選択した。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)、pINCYプラスミド (Incyte Pharmaceuticals, Palo Alto CA)、またはそれらの誘導体などの好適なプラスミドのポリリンカーの適合性制限酵素部位にcDNAを結合させた。この組換えプラスミドを、Stratagene社のXL1-Blue, XL1-BlueMRF、SOLR、またはLife Technologies社のDH5 またはDH10B、ELECTROMAX DH10Bを含むコンピテント大腸菌細胞に導入し形質転換した。

20

30

40

【0250】

2 cDNAクローンの単離

上記実施例1に記載したように得たプラスミドは、UNIZAPベクターシステム (Stratagene) 或いは細胞溶解を利用した*in vivo*切除によって宿主細胞から回収した。プラスミドの精製には、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 Plus、QIAWELL 8 Plus Plasmid、またはQIAWELL 8 Ultra Plasmid 精製システム、またはREAL Prep 96 プラスミドキットの内の少なくとも1つを用いた。沈殿させた後、0.1 mlの蒸留

50

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。

【0251】

別法では、ハイスルーブットの直接結合PCR法によって宿主細胞溶解物からプラスミドDNAを増幅した。(Rao, V. B. (1994) Anal. Biochem. 216: 1-14)。宿主細胞の溶解及び熱サイクリングステップは、1つの反応混合液で行った。サンプルを処理してから384-ウェルプレートに移して保管し、増幅したプラスミドDNAの濃度を、PICOGREEN色素(Molecular Probes, Eugene OR)及びFluoroskan II蛍光スキャナ(Labsystems Oy, Helsinki, Finland)を用いて蛍光定量的に測定した。

【0252】

10

3 シークエンシング及び分析

実施例2に記載したようにプラスミドから回収したインサイトcDNAを、以下に示すようにシークエンシングした。cDNAのシークエンシング反応は標準的な方法で行うか、またはHYDRAマイクロディスペンサー(Robbins Scientific)或いはMICROLAB 2200(Hamilton)液体移送装置と共にABI CATALYST 800(PE Biosystems)サーマルサイクラー或いはPTC-200 thermal cycler(MJ Research)などのハイスルーブット装置を用いて行った。cDNAのシークエンシング反応は、Amersham Pharmacia Biotech社の試薬、またはABI PRISM BIGDYE Terminator cycle sequencing ready reactionキット(PE Biosystems)などのABIシークエンシングキットに含まれる試薬を用いて行った。cDNAシークエンシングの反応物の電気泳動的による分離及び標識したポリヌクレオチドの検出は、MEGABACE 1000 DNAシークエンシングシステム(Molecular Dynamics)、標準ABIプロトコル及び塩基対呼び出しソフトウェアを用いるABI PRISM 373または377シークエンシングシステム(PE Biosystems)、または当分野で周知のその他の配列解析システムを用いて行った。cDNA配列内の読み枠は、標準的な方法(Ausubel, 1997, 前出, unit 7.7)を用いて決定した。cDNA配列の幾つかを選択して、実施例8に記載した方法で配列を伸長した。

20

【0253】

30

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

40

50

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

10

【0254】

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

20

【0255】

完全長のポリヌクレオチド配列およびポリペプチド配列の組み立て及び分析に用いる上記のプログラムは、SEQ ID NO: 48 - 94 のポリヌクレオチド配列断片の同定にも利用できる。ハイブリダイゼーション及び増幅技術に有用である約 20 ~ 約 4000 ヌクレオチドの断片を表 4 の列 4 に示した。

【0256】

4 ゲノム DNA からのコード配列の同定および編集

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

30

40

50

ヌクレオチド配列は、実施例 3 に説明した組み立て方法で G e n s c a n 推定コード配列とインサイト c D N A および / または公共の c D N A 配列を組み立てて作製した。別法では、完全長ポリヌクレオチド配列は、その全てが編集した或いは未編集の G e n s c a n 推定コード配列から作製した。

【 0 2 5 7 】

5 c D N A 配列データを用いるゲノム配列データの組み立て

ステッチ配列 (S t i c h e d S e q u e n c e)

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

【 0 2 5 8 】

ストレッチ配列 (S t r e t c h e d S e q u e n c e)

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

【 0 2 5 9 】

6 R M E P をコードするポリヌクレオチドの染色体マッピング

S E Q I D N O : 4 8 - 9 4 を組み立てるために用いた配列を、B L A S T 及び S m i t h - W a t e r m a n アルゴリズムを用いて、インサイト L I F E S E Q 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。S E Q I D N O : 4 8 - 9 4 と一致するこれらのデータベースの配列を、P h r a p (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。S t a n f o r d H u m a n G e n o m e S e C e n t e r (S H G C) 、W h i t e h e a d I n s t i t u t e f o r G e n o m e R e s e a r c h (W I G R) 及び G e n e t h o n などの公

共の情報源から入手できる放射線ハイブリッド (r a d i a t i o n h y b r i d) 及び遺伝子マッピングのデータを用いて、クラスター化した配列がすでにマッピングされているかを調べる。クラスターにマッピングされた配列が含まれている場合は、そのクラスターの全ての配列 (特定の S E Q I D N O を含む) をそのマッピング位置に割り当てた。

【 0 2 6 0 】

遺伝子地図の位置は、範囲、区間、またはヒト染色体によって表される。センチモルガンで示したマッピング位置の範囲は、染色体の短腕 (p) の末端から測定した (センチモルガン (c M) は、同一染色体上の遺伝子間の乗換え率に基づいた距離を表す単位である。平均すると、1 c M はヒトの染色体の1メガベースに概ね等しいが、組換え率の高い部分と低い部分があるため、大きく変化し得る)。距離 c M は、配列がそれぞれのクラスターに含まれている放射線ハイブリッドマーカーの境界を検出できる G e n e t h o n によってマッピングされた遺伝子マーカーに基づいている。N C B I 「 G e n e M a p 9 9 」 (<http://www.ncbi.nlm.nih.gov/genemap>) などの公衆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。

10

【 0 2 6 1 】

このようにして、S E Q I D N O : 5 3 が、第1染色体の159.6 c M ~ 164.1 c M の範囲内にマッピングされた。S E Q I D N O : 6 1 が、第8染色体の30.7 c M ~ 60.00 c M の範囲内にマッピングされた。S E Q I D N O : 6 9 が、第10染色体の158.30 c M ~ q 末端の範囲内にマッピングされた。S E Q I D N O : 7 0 が、第1染色体の63.90 c M ~ 74.80 c M の範囲内にマッピングされた。S E Q I D N O : 7 1 が、第1染色体の159.60 c M ~ 164.10 c M の範囲内にマッピングされた。S E Q I D N O : 7 3 が、第11染色体の34.30 c M ~ 37.00 c M の範囲内にマッピングされた。S E Q I D N O : 7 5 が、第2染色体の107.10 c M ~ 118.00 c M の範囲内にマッピングされた。S E Q I D N O : 7 6 が、第7染色体の7.80 c M ~ 10.60 c M の範囲内にマッピングされた。S E Q I D N O : 7 9 が、第22染色体の22.20 c M ~ 40.20 c M の範囲内にマッピングされた。S E Q I D N O : 8 1 が、第4染色体のp 末端 ~ 6.70 c M の範囲内にマッピングされた。S E Q I D N O : 8 4 が、第5染色体の156.0 c M ~ 157.6 c M の範囲内にマッピングされた。S E Q I D N O : 8 8 が、第11染色体の117.9 c M ~ 123.5 c M の範囲内にマッピングされた。S E Q I D N O : 9 1 が、第5染色体の152.3 c M ~ 155.5 c M の範囲内にマッピングされた。

20

30

【 0 2 6 2 】

7 ポリヌクレオチド発現の分析

ノーザン分析は、遺伝子の転写物の存在を検出するために用いられる実験用技術であり、特定の細胞種或いは組織からのRNAが結合されている膜への標識されたヌクレオチド配列のハイブリダイゼーションを伴う (例えば、S a m b r o o k , 前出 , 7 章 ; 及び A u s u b e l , F . M . 他、前出 , 4 章 及び 1 6 章 を参照)。

【 0 2 6 3 】

B L A S T に用いる類似のコンピュータ技術を用いて、G e n B a n k 或いは L I F E S E Q (I n c y t e P h a r m a c e u t i c a l s) のような c D N A データベース内の同一或いは関連する分子を検索する。この分析は多くの膜系ハイブリダイゼーションより非常に速度が速い。さらにコンピュータ検索の感度を変更して、任意の特定の一致が、厳密な一致或いは相同的一致の何れかとして分類されるかを確定することができる。検索の基準は、

40

【 0 2 6 4 】

【 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

として定義される積スコアである。積スコアは、0 ~ 100 の標準化された値であり、以下のように求める。BLASTスコアにヌクレオチド配列の一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除す。高スコアのセグメントの対(HSP)において一致する各塩基に+5のスコアを割り当て、各不適性塩基対に-4を割り当てることにより、BLASTスコアを計算する。2つの配列は、2以上のHSPを共有し得る(ギャップにより離隔される)。2以上のHSPがある場合には、最高BLASTスコアの塩基対を用いて積スコアを計算する。積スコアは、BLASTアラインメントの断片的重複と質とのバランスを表す。例えば積スコア100は、比較した2つの配列の短い方の長さ全体にわたって100%一致する場合にのみ得られる。積スコア70は、100%の同一性で重畳が70%であるか、或いは88%の同一性で重畳が100%であるかの何れかの場合である。積スコア50は、100%の同一性で重畳が50%であるか、或いは79%の同一性で重畳が100%であるかの何れかの場合である。

10

【0265】

或いは、RMEPをコードするポリヌクレオチド配列は、由来する組織に対して分析する。例えば、ある完全長の配列は、少なくとも部分的にインサイトcDNA配列をオーバーラップさせて組み立てられる(実施例3を参照)。各cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各ヒト組織は、心血管系、結合組織、消化系、胚構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液および免疫系、肺、筋骨格系、神経系、脾臓、呼吸器系、感覚器官、皮膚、顎口腔系、分類不能/混合、または尿管などの1つの生物/組織のカテゴリーに分類される。各カテゴリーにおけるライブラリの数をカウントし、その合計数を全カテゴリーのライブラリ数で除す。同様に、各ヒト組織は、癌、細胞系、発生、炎症、神経、外傷、心血管、プール(pool)などの1つの疾患/症状のカテゴリーに分類され、各カテゴリーにおけるライブラリの数をカウントし、その合計数を全カテゴリーのライブラリ数で除す。得られるパーセンテージは、RMEPをコードするcDNAの疾患特異的な発現を反映する。cDNA配列およびcDNAライブラリ/組織の情報は、LIFESEQ GOLD データベース(Incyte Genomics, Palo Alto CA)から得ることができる。

20

30

【0266】

8 RMEPをコードするポリヌクレオチドの伸長

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

40

【0267】

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

【0268】

当分野で既知の方法を利用したPCR法で高い忠実度で増幅した。PCRはPTC-200 thermal cycler (MJ Research, Inc.)を用いて96ウェルブロックプレートで行った。反応混合液は、鋳型DNA及び200 nmolの各プライマー、 Mg^{2+} と $(NH_4)_2SO_4$ と-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

50

)、E L O N G A S E 酵素 (L i f e T e c h n o l o g i e s)、P f u D N A ポリ
 メラーゼ (S t r a t a g e n e) を含む。プライマーの組、P C I A と P C I 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 9 4 で 3 分間
 ステップ 2 9 4 で 1 5 秒
 ステップ 3 6 0 で 1 分間
 ステップ 4 6 8 で 2 分間
 ステップ 5 ステップ 2、3、及び 4 を 2 0 回繰り返す
 ステップ 6 6 8 で 5 分間
 ステップ 7 4 で 保管

10

別法では、プライマーの組、T 7 と S K + に対して以下のパラメータで増幅を行った。

ステップ 1 9 4 で 3 分間
 ステップ 2 9 4 で 1 5 秒
 ステップ 3 5 7 で 1 分間
 ステップ 4 6 8 で 2 分間
 ステップ 5 ステップ 2、3、及び 4 を 2 0 回繰り返す
 ステップ 6 6 8 で 5 分間
 ステップ 7 4 で 保管。

【 0 2 6 9 】

各ウェルの DNA 濃度は、1 X T E 及び 0 . 5 μ l の希釈していない P C R 産物に溶解
 した 1 0 0 μ l の P I C O G R E E N 定量試薬 (0 . 2 5 % (v / v) P I C O G R E
 E N ; M o l e c u l a r P r o b e s , E u g e n e O R) を不透明な蛍光光度計
 プレート (C o m i n g C o s t a r , A c t o n M A) の各ウェルに分配して D N
 A が試薬と結合できるようにして測定する。このプレートを F l u o r o s k a n I I
 (L a b s y s t e m s O y , H e l s i n k i , F i n l a n d) でスキャンして
 、サンプルの蛍光を計測して D N A の濃度を定量化する。反応混合物の 5 ~ 1 0 μ l のア
 リコットを 1 % のアガロースミニゲル上での電気泳動によって解析し、何れの反応物が配
 列を伸長することに成功したかを決定する。

20

【 0 2 7 0 】

伸長したヌクレオチドを脱塩及び濃縮してから 3 8 4 ウェルプレートに移し、C v i J I
 コレラウィルスエンドヌクレアーゼ (M o l e c u l a r B i o l o g y R e s e a r
 c h , M a d i s o n W I) で消化し、p U C 1 8 ベクター (A m e r s h a m P h
 a r m a c i a B i o t e c h) に再連結する前に音波処理またはせん断を行った。シ
 ョットガンシーケンシングのために、消化したヌクレオチドを低濃度 (0 . 6 ~ 0 . 8
 %) のアガロースゲル上に分離して断片を切断し、寒天を A g a r A C E (P r o m e
 g a) で消化した。T 4 リガーゼ (N e w E n g l a n d B i o l a b s , B e v e
 r l y M A) を用いて伸長したクローンを p U C 1 8 ベクター (A m e r s h a m P
 h a r m a c i a B i o t e c h) に再連結し、P f u D N A ポリメラーゼ (S t r a
 t a g e n e) で制限部位の伸び出しを処理してコンピテント大腸菌細胞に形質移入した
 。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りと
 って L B / 2 X カルベニシリン培養液の 3 8 4 ウェルプレートに 3 7 で一晩培養した。

30

40

【 0 2 7 1 】

細胞を溶解して、T a q D N A ポリメラーゼ (A m e r s h a m P h a r m a c i a
 B i o t e c h) 及び P f u D N A ポリメラーゼ (S t r a t a g e n e) を用いて以
 下の手順で D N A を P C R 増幅した。

ステップ 1 9 4 で 3 分間
 ステップ 2 9 4 で 1 5 秒
 ステップ 3 6 0 で 1 分間
 ステップ 4 7 2 で 2 分間
 ステップ 5 ステップ 2、3、及び 4 を 2 9 回繰り返す

50

ステップ6 72 で5分間

ステップ7 4 で保管。

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

【0272】

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

【0273】

9 個々のハイブリダイゼーションプローブの標識化及び使用法

SEQ ID NO: 48 - 94から導き出されたハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記すが、より大きなcDNAフラグメントの場合でも基本的に同じ手順を用いる。オリゴヌクレオチドを、OLIGO4.06ソフトウェア(National Bioscience)のような最新式のソフトウェアを用いてデザインし、50 pmolの各オリゴマーと、250 μ Ciの[³²P]アデノシン三リン酸(Amersham, Chicago, IL)及びT4ポリヌクレオチドキナーゼ(DuPont NEN, Boston MA)とを組み合わせることでより標識する。標識されたオリゴヌクレオチドを、SEPHADEX G-25超精細排除デキストランビードカラム(Amersham Pharmacia Biotech)を用いて実質的に精製する。毎分10⁷カウントの標識されたプローブを含むアリコットを、次のエンドヌクレアーゼ、Ase I、Bgl II、Eco RI、Pst I、Xba I或いはPvu II(DuPont NEN)の1つを用いて切断したヒトゲノムDNAの典型的な膜ベースのハイブリダイゼーション解析において用いる。

【0274】

各切断物からのDNAを、0.7%アガロースゲル上で分画して、ナイロン製メンブラン(Nytran Plus, Schleicher & Schuell, Durham NH)に転写する。ハイブリダイゼーションは40で16時間かけて行う。非特異的シグナルを取り除くため、例えば、最大0.1xクエン酸ナトリウム食塩水及び0.5%ドデシル硫酸ナトリウムの条件の下、プロットを順次室温にて洗浄する。ハイブリダイゼーションパターンをオートラジオグラフィー或いは別のイメージ化手段で視覚化して比較する。

【0275】

10 マイクロアレイ

マイクロアレイ上のアレイエレメントの連結または合成は、フォトリソグラフィ、ピエゾプリント(インクジェットプリンター、前出のBaldeschweiler等を参照)、機械的マイクロスポッティング技術及びこれらから派生したものをを用いて達成することが可能である。上記各技術において基板は、均一な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。別法では、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、熱や紫外線、または化学的或いは機械的な結合手段で基板の表面にエレメントを配置して結合させることができる。通常のアレイは利用可能な方法や機械を用いて作製でき、任意の適正な数のエレメントを含めることができる(Schena, M. 他(1995) Science 270: 467-470, Shalton, D. 他(1996) Genome Res. 6: 639-645, M

10

20

30

40

50

arshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31. を参照)。

【0276】

完全長cDNA、発現遺伝子配列断片(EST)、或いはそれらの断片やオリゴマーが、マイクロアレイのエLEMENTとなり得る。ハイブリダイゼーションに好適な断片やオリゴマーを、LASERGENEソフトウェア(DNA STAR)などの当分野で周知のソフトウェアを用いて選択することが可能である。このアレイELEMENTを、生体サンプル中のポリヌクレオチドとハイブリダイズさせる。生体サンプル中のポリヌクレオチドは、検出を容易にするために蛍光標識またはその他の分子タグに結合する。ハイブリダイゼーションの後、生体サンプルからハイブリダイズしなかったヌクレオチドを除去し、蛍光スキャナを用いて各アレイELEMENTにおけるハイブリダイゼーションを検出する。別法では、レーザー脱離及び質量スペクトロメトリーを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上のELEMENTにハイブリダイズする各ポリヌクレオチドの相補性の程度及び相対的存在量は、算定することができる。一実施例におけるマイクロアレイの調整及び使用について、以下に詳述する。

10

【0277】

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)セルソース法を用いてポリ(A)⁺RNAを精製する。各ポリ(A)⁺RNAサンプルは、MMLV逆転写酵素、0.05 pg/μlのオリゴ(dT)プライマー(21mer)、1×第1鎖緩衝液、0.03単位/μlのRNアーゼインヒビター、500 μM dATP、500 μM dGTP、500 μM dTTP、40 μM dCTP、40 μM dCTP-Cy3(BDS)またはdCTP-Cy5(Amersham Pharmacia Biotech)を用いて逆転写する。この逆転写反応は、GEMBRIGHTキット(Incyte)を用いて、200 ngのポリ(A)⁺RNAを含む25 ml容量で行う。特異的なコントロールポリ(A)⁺RNAは、in vitro転写により非コーディング酵母ゲノムDNAから合成する。370 で2時間インキュベートした後、各反応サンプル(一方はCy3標識、他方はCy5標識)は、2.5 mlの0.5 M 水酸化ナトリウムで処理し、850 で20分間インキュベートし、反応を停止させてRNAを変性する。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピナラム(CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA)を用いて精製する。結合後、2つの反応サンプルを、1 mlのグリコーゲン(1 mg/ml)、60 mlの酢酸ナトリウム及び300 mlの100%エタノールを用いてエタノール沈殿させる。サンプルは次に、Speed VAC(Savant Instruments Inc., Holbrook NY)を用いて乾燥して仕上げ、14 μl 5×SSC/0.2% SDS中で再懸濁する。

20

30

【0278】

マイクロアレイの準備

本発明の配列を用いて、アレイELEMENTを作製する。各アレイELEMENTは、クローニングcDNA挿入断片を含むベクターを含有する細菌性細胞から増幅する。PCR増幅は、cDNA挿入断片に隣接するベクター配列に相補的なプライマーを用いる。30サイクルのPCRによって、1~2 ngの初期量から5 μgを超える最終量までアレイELEMENTを増幅する。増幅されたアレイELEMENTは、SEPHACRYL-400(Amersham Pharmacia Biotech)を用いて精製する。

40

【0279】

精製したアレイELEMENTを、ポリマーコートされたスライドガラス上に固定する。顕微鏡スライドガラス(Corning)は、処理中及び処理後に大量の蒸留水での洗浄と、0.1%のSDS及びアセトン中で超音波による洗浄を行う。スライドガラスは、4%フッ化水素酸(VWR Scientific Products Corporation (VWR), West Chester PA)中でエッチングし、蒸留水中で広範囲に

50

わたって洗浄し、95%エタノール中の0.05%アミノプロピルシラン (Sigma) でコーティングする。コーティングしたスライドガラスは、110 °Cの天火で硬化させる。

【0280】

米国特許第5,807,522号に記載されている方法を用いて、コーティングしたガラス基板にアレイエレメントを付加する。この特許に引用することを以て本明細書の一部とする。平均濃度が100 ng/μlのアレイエレメントDNA 1 μlを高速機械装置により開放型キャピラリープリンティングエレメント (open capillary printing element) に充填する。次にこの装置が、スライド毎に約5 nLのアレイエレメントサンプルを分注する。

10

【0281】

マイクロアレイには、STRATALINKER UVクロスリンカー (Stratagene) を用いてUV架橋する。マイクロアレイは、室温において0.2% SDSで1回洗浄し、蒸留水で3回洗浄する。非特異的な結合部位は、リン酸緩衝生理食塩水 (PBS) (Tropix, Inc., Bedford MA) における0.2%カゼイン中で60 °Cで30分間マイクロアレイをインキュベートし、その後上述したように0.2% SDS及び蒸留水で洗浄することによってブロックする。

【0282】

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。サンプル混合液を、65 °Cで5分間加熱し、マイクロアレイ表面上に一定量分注してから1.8 cm² のカバーガラスで覆う。このアレイを、顕微鏡スライドより僅かに大きいキャピティを有する防水チャンバーに移す。チャンバーの角に140 μlの5×SSCを加えて、チャンバー内を湿度100%に保持する。このアレイを含むチャンバーを、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、第2洗浄緩衝液中 (0.1×SSC) において45 °Cで10分間それぞれ3回洗浄し、その後乾燥させる。

20

【0283】

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を励起するための488 nm、及びCy5を励起するための632 nmのスペクトル線を生成し得るInnova 70混合ガス10 Wレーザー (Coherent, Inc., Santa Clara CA) を備えた顕微鏡で検出する。20倍の顕微鏡対物レンズ (Nikon, Inc., Melville NY) を用いて、アレイ上に励起レーザー光を集中させる。このアレイを含むスライドを顕微鏡のコンピュータ制御X-Yステージに置き、対物レンズを通してラスタスキャンする。本実施例で用いた1.8 cm×1.8 cmのアレイは、20 μmの解像度でスキャンする。

30

【0284】

2つの異なるスキャンにおいて、混合ガスマルチラインレーザーは2つの蛍光体を連続的に励起する。放射された光は、波長に基づいて2つの蛍光体に対応する2つの光電子増倍管検出器 (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) に分割される。アレイと光電子増倍管との間に配設された好適なフィルタを用いて信号をフィルタリングする。用いる蛍光体の最大発光は、Cy3では565 nm、Cy5では650 nmである。装置は両方の蛍光体からのスペクトルを同時に記録できるが、レーザー源に好適なフィルタを用いて、蛍光体1つにつき1回スキャンし、各アレイを通常2回スキャンする。

40

【0285】

スキャンの感度は通常、既知濃度のサンプル混合体に添加されるcDNAコントロール種

50

により生成されるシグナル強度を用いて校正する。アレイ上の特定の位置には相補的 D N A 配列を含め、その位置におけるシグナルの強度がハイブリダイズする種の重量比 1 : 1 0 0 , 0 0 0 に相関するようにする。異なる試料 (例えば検査細胞及びコントロール細胞を代表する) からの 2 つのサンプルを、各々異なる蛍光体で標識し、他と異なって発現する遺伝子を同定するために単一のアレイにハイブリダイズさせる場合には、校正は 2 つの蛍光体を有する校正する c D N A のサンプルを標識して、ハイブリダイゼーション混合液に各々等量を加えて行う。

【0286】

光電子増倍管の出力は、I B M コンパチブル P C コンピュータにインストールされた 1 2 ビット R T I - 8 3 5 H アナログ - デジタル (A I D) 変換ボード (A n a l o g D e v i c e s , I n c . , N o r w o o d M A) を用いてデジタル化される。デジタル化されたデータは、リニア 2 0 色変換を用いてシグナル強度が青色 (低シグナル) から赤色 (高シグナル) までの擬似カラー範囲にマッピングされるイメージとして表示される。データはまた、定量的に分析される。2 つの異なる蛍光体を同時に励起して測定する場合には、各蛍光体の発光スペクトルを用いて、先ずデータは蛍光体間の光学的漏話 (重複発光スペクトルに起因する) に対して補正される。

10

【0287】

グリッドを蛍光シグナルイメージ上に重畳して、各スポットからのシグナルがグリッドの各エレメントに中央に位置するようにする。各エレメント内の蛍光シグナルを統合し、シグナルの平均強度に対応する数値を得る。シグナル分析に用いるソフトウェアは、G E M T O O L S 遺伝子発現分析プログラム (I n c y t e) である。

20

【0288】

1 1 相補的ポリヌクレオチド

R M E P をコードする配列或いはその任意の一部に対して相補的な配列は、天然の R M E P の発現を低下させるため即ち阻害するために用いられる。約 1 5 ~ 約 3 0 個の塩基対を含むオリゴヌクレオチドの使用について記すが、より小さな或いはより大きな配列の断片の場合でも本質的に同じ方法を用いることができる。O l i g o 4 . 0 6 ソフトウェア (N a t i o n a l B i o s c i e n c e s) 及び R M E P のコーディング配列を用いて、適切なオリゴヌクレオチドを設計する。転写を阻害するためには、最も独特な 5 配列から相補的なオリゴヌクレオチドを設計し、これを用いてプロモーターがコーディング配列に結合するのを阻害する。翻訳を阻害するためには、相補的なオリゴヌクレオチドを設計して、リボソームが R M E P をコードする転写物に結合するのを阻害する。

30

【0289】

1 2 R M E P の発現

R M E P の発現及び精製は、細菌若しくはウイルスを基にした発現系を用いて行うことができる。細菌で R M E P が発現するために、抗生物質耐性及び c D N A の転写レベルを高める誘導性のプロモーターを含む好適なベクターに c D N A をサブクローニングする。このようなプロモーターには、l a c オペレーター調節エレメントに関連する T 5 または T 7 バクテリオファージプロモーター及び t r p - l a c (t a c) ハイブリッドプロモーターが含まれるが、これらに限定されるものではない。組換えベクターを、B L 2 1 (D E 3) などの好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピル - D チオガラクトピラノシド (I P T G) で誘発されると R M E P を発現する。真核細胞での R M E P の発現は、昆虫細胞株または哺乳動物細胞株に一般にバキュロウイルスとして知られている A u t o g r a p h i c a c a l i f o r n i c a 核多面性ウイルス (A c M N P V) を感染させて行う。バキュロウイルスの非必須ポリヘドリン遺伝子を、相同組換え或いは転移プラスミドの媒介を伴う細菌の媒介による遺伝子転移のどちらかによって、R M E P をコードする c D N A と置換する。ウイルスの感染力は維持され、強いポリヘドリンプロモータによって高いレベルの c D N A の転写が行われる。組換えバキュロウイルスは、多くの場合は S p o d o p t e r a f r u g i p e r d a (S f 9) 昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感

40

50

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。(例えば、Engelhard, E. K. 他 (1994) Proc. Natl. Acad. Sci. U S A 91: 3224 - 3227; Sandig, V. 他 (1996) Hum. Gene Ther. 7: 1937 - 1945. を参照)。

【0290】

殆どの発現系では、RMEP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。Schistosoma japonicumからの26キロダルトンの酵素GSTによって、タンパク質の活性及び抗原性を維持した状態で固定されたグルタチオンで融合タンパク質の精製が可能となる(Amersham Pharmacia Biotech)。精製の後、GST部分を特定の操作部位でRMEPからタンパク質的に切断できる。アミノ酸8個のペプチドであるFLAGで、市販のモノクローナル及びポリクローナル抗FLAG抗体(Eastman Kodak)を用いた免疫親和性の精製が可能となる。6個の連続するヒスチジン残基のストレッチである6-Hisによって、金属キレート樹脂(QIAGEN)で精製が可能となる。タンパク質の発現及び精製の方法は、Ausubel (1995, 前出, ch 10, 16)に記載されている。これらの方法で精製したRMEPを直接用いて以下の実施例16及び17の assaysを行うことができる。

【0291】

13 機能の assay

RMEPの機能は、哺乳動物細胞培養系において生理学的に高められたレベルでのRMEP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。このようなベクターには、pCMV SPORTTM (Life Technologies.) 及びpCR3.1 (Invitrogen, Carlsbad, CA)が含まれ、どちらもサイトメガロウイルスプロモーターを含んでいる。5~10 µgの組換えベクターを、例えば内皮由来か造血由来のヒト細胞株にリポソーム製剤或いは電気穿孔法によって一時的に形質移入する。更に、標識タンパク質をコードする配列を含む1~2 µgのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入された細胞と形質移入されていない細胞とを区別できる。また、標識タンパク質の発現によって、cDNAの組換えベクターからの発現を正確に予想できる。このような標識タンパク質には、緑色蛍光タンパク質(GFP; Clontech)、及びCD64またはCD64-GFP融合タンパク質が含まれる。レーザー光学に基づいた技術を利用した自動流動細胞計測法(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。また、FCMで、先行した或いは同時の細胞死の現象を診断する蛍光分子の取り込みを検出して計量する。これらの現象には、プロピジウムヨウ化物でのDNAの染色によって計測される核DNA内容物の変化と、プロモデオキシウリジンの取り込み量の低下によって計測されるDNA合成の下方調節と、特異的な抗体との反応性によって計測される細胞表面及び細胞内のタンパク質の発現の変化と、蛍光複合アネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変化とが含まれる。流動細胞計測法は、Ormerod, M. G. による(1994) Flow Cytometry Oxford, New York, NY. に記載されている。

【0292】

遺伝子発現におけるRMEPの影響は、RMEP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)の保存された領域と結合する。形質転換された細胞と形質転換されない細胞とは、ヒトIgGかCD64に対する抗体のどちらかで被覆された磁気ビ

ードを用いて分離することができる (DYNAL, Lake Success, NY)。mRNAは、当分野で周知の方法で細胞から精製することができる。RMEP及び目的の他の遺伝子をコードするmRNAの発現は、ノーザン分析やマイクロアレイ技術で分析することができる。

【0293】

1.4 RMEP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RMEP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。

10

【0294】

別法では、RMEPアミノ酸配列をLASERGENEソフトウェア (DNASTAR) を用いて解析して免疫原性の高い領域を決定し、対応するオリゴペプチドを合成してこれを用いて当業者に周知の方法で抗体を生産する。C末端付近の、或いは隣接する親水性領域内のエピトープなどの適切なエピトープの選択については、当分野で周知である (例えば、前出のAusubel, 1995, 11章を参照)。

【0295】

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

20

【0296】

1.5 特異的な抗体を用いる天然RMEPの精製

天然RMEP或いは組換えRMEPを、RMEPに特異的な抗体を用いるイムノアフィニティークロマトグラフィーにより実質的に精製する。イムノアフィニティークラムは、CNBr-活性化SEPHAROSE (Amersham Pharmacia Biotech) のような活性化クロマトグラフィー用レジンと抗RMEP抗体とを共有結合させることにより形成する。結合の後、そのレジンを製造者の使用説明書に従ってブロッキング処理し洗浄する。

30

【0297】

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

40

【0298】

1.6 RMEPと相互作用する分子の同定

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

50

【0299】

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

【0300】

RMEPはまた、ハイスループット型の酵母2ハイブリッドシステムを使用するPATHCALLINGプロセス (CuraGen Corp., New Haven CT) に用いて、遺伝子の2つの大きなライブラリによってコードされるタンパク質間の全ての相互作用を決定することができる (Nandabalan, K. 他 (2000) 米国特許第6,057,101号)。

10

【0301】

17 RMEPの活性の実証

RMEP活性は、ポリアクリルアミドゲルシフト法 (polyacrylamide gel mobility - shift assay) によって実証する。このアッセイを行うために、RMEP cDNAを含む真核生物発現ベクターで、COS7やHeLa、またはCHOなどの哺乳動物細胞系を形質転換して、RMEPを発現させる。形質転換した後、この形質転換細胞をRMEPの発現及び蓄積が許容される好適な条件下で48~72時間インキュベートする。当分野で周知の方法で可溶化タンパク質を含む抽出物を、RMEPを発現している細胞から調製できる。RMEPを含む抽出物の一部を [³²P] 標識したRNAに添加する。当分野で周知の方法で、放射性RNAを*in vitro*で合成することができる。この混合物を、RNAアーゼインヒビターが存在するバッファー条件下で、25℃で5~10分間インキュベートする。インキュベートした後、このサンプルをポリアクリルアミドゲル電気泳動法で分離して、オートラジオグラフィーで分析する。オートラジオグラム上のバンドの存在が、RMEPと放射性転写物との複合体の形成を示唆する。形質転換していない細胞から準備したコントロール抽出物を用いて調製したサンプルには類似の移動度のバンドが存在しないはずである。

20

【0302】

別法では、RMEPのリボソームタンパク質機能を、哺乳動物細胞培養系においてリボソームタンパク質をコードする配列を生理学的に上昇したレベルで発現させて評価する。cDNAを、cDNAを高いレベルで発現させる強力なプロモーターを含む哺乳動物発現ベクターにサブクロニングする。このようなベクターには、それぞれサイトメガロウイルスプロモーター (PCMV) を含むPCMV SPORT (Life Technologies) 及びPCR3.1 (Invitrogen Corporation) が含まれる。リボソーム法または電気穿孔法の何れかを用いて、5~10 µgの組換えベクターを好ましくは内皮若しくは造血系由来のヒト細胞系にトランスフェクトする。マーカータンパク質をコードする配列を含む追加のプラスミド1~2 µgを同時にトランスフェクトする。

30

【0303】

マーカータンパク質の一過性の発現は、トランスフェクト細胞と非トランスフェクト細胞とを区別するための手段となり、組換えベクターからのcDNA発現の信頼できる予測因子である。このようなマーカータンパク質には、例えば、緑色蛍光タンパク質 (GFP: Clontech)、CD64、またはCD64 - GFP融合タンパク質が含まれる。自動化されたレーザー光学系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を発現するトランスフェクト細胞を特定し、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。

40

【0304】

FCMにより、細胞死の前の現象や細胞死時の現象を診断する蛍光分子の取り込みを検出し定量する。このような現象には、ヨウ化プロピジウムでDNA染色して測定できる核D

50

N A 内容物の変化、前方光散乱及び 90 度側方光散乱により測定できる細胞の大きさ及び粒状度の変化、プロモデオキシウリジンの取り込みの減少により測定できる D N A 合成のダウンレギュレーション、特異的な抗体との相互作用により測定できる細胞表面及び細胞内タンパク質の発現の変化、並びにフルオレセイン共役アネキシン V タンパク質の細胞表面への結合により測定できる細胞膜成分の変化がある。フローサイトメトリーの方法については、Ormerod M G (1994) Flow Cytometry, Oxford University Press, New York NY に記載されている。

【0305】

遺伝子発現におけるリボソームタンパク質の影響は、リボソームタンパク質をコードする配列でトランスフェクトされた高度に精製された細胞集団と C D 6 4 か或いは C D 6 4 - G F P の何れかを用いて評価することができる。C D 6 4 及び C D 6 4 - G F P がトランスフェクト細胞の表面で発現し、ヒト免疫グロブリン G (I g G) の保存された領域に結合する。トランスフェクト細胞を、ヒト I g G 或いは C D 6 4 の抗体の何れかでコーティングした磁気ビーズで非トランスフェクト細胞から分離する (D Y N A L , I n c . , L a k e S u c c e s s N Y) 。当分野で周知の方法を用いて m R N A を細胞から精製することができる。リボソームタンパク質をコードする m R N A 及び他の目的の遺伝子の発現を、ノーザン分析またはマクロアレイ技術で分析することができる。

10

【0306】

別法では、R M E P 活性は、 $[^{14}\text{C}]$ システインの存在下での基質 t R N A のアミノアシル化として測定する。R M E P を、バッファーにおいて t R N A ^{C y s} 及び $[^{14}\text{C}]$ システイン (若しくは好適な t R N A 及びアミノ酸基質) と共にインキュベートする。 $[^{14}\text{C}]$ 標識産物を、クロマトグラフィーにより遊離 $[^{14}\text{C}]$ アミノ酸から分離し、取り込まれた $[^{14}\text{C}]$ をシンチレーションカウンターで定量する。検出した $[^{14}\text{C}]$ の量がこのアッセイにおける R M E P の活性に比例する。

20

【0307】

別法では、R M E P 活性は、R M E P を含むサンプルを溶液中 (1 m M A T P , 5 m M H e p e s - K O H (p H 7 . 0) , 2 . 5 m M K C l , 1 . 5 m M 塩化マグネシウム , 及び 0 . 5 m M D T T 、並びに誤ってアシル化した $[^{14}\text{C}]$ - G l u - t R N A G I n (例えば、1 μ M) 、類似の濃度の非標識 L - グルタミンを含む) においてインキュベートして測定する。3 M の酢酸ナトリウム (p H 5 . 0) で反応を停止させた後、混合液を等量の水飽和フェノールで抽出し、室温で 1 分間遠心分離 (1 , 5 0 0 0 \times g) して水相と有機相とに分離させる。水相を除去して、- 7 0 で 1 5 分間、3 倍量のエタノールで沈殿させる。沈殿したアミノアシル - t R N A を、4 で 1 5 分間遠心分離 (1 , 5 0 0 0 \times g) して回収する。このペレットを 2 5 m M K O H 中において再懸濁し、6 5 で 1 0 分間脱アシル化し、0 . 1 M H C L (最終 p H 6 ~ 7) で中和し、真空下で乾燥させる。乾燥ペレットを水中で再懸濁し、セルロース T L C プレート上にスポットする。このプレートを、イソプロパノール / ギ酸 / 水若しくはアンモニア / 水 / クロロホルム / メタノールの何れかで現像する。このイメージを濃度測定し、G l u と G l n の相対量を R f 値及びスポットの相対強度に基づいて算定する。R M E P 活性を、G l u が G l u - t R N A ^{G L N} としてアシル化される一方、G l u の変換により生じる G l n の量に基づいて算定する (C u r n o w , A . W . ら (1997) P r o c . N a t l . A c a d . S c i . 94 : 11819 - 26 の方法を利用) 。

30

40

【0308】

当業者であれば、本発明の範囲及び精神から逸脱することなく本発明の記載した方法及びシステムの種々の改変を行うことができるであろう。特定の好適な実施例に基づいて本発明を説明したが、本発明の範囲が、そのような特定の実施例に不当に制限されるべきではないことを理解されたい。実際に、分子生物学或いは関連する分野の専門家には明らかな、本明細書に記載の本発明の実施例の様々な改変は、特許請求の範囲に含まれる。

【0309】

(表の簡単な説明)

50

表 1 は、本発明の完全長ポリヌクレオチド配列及びポリペプチド配列に対する系統的な名称を示す。

【 0 3 1 0 】

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

【 0 3 1 1 】

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

10

【 0 3 1 2 】

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

【 0 3 1 3 】

表 5 は、本発明のポリヌクレオチドの代表的な c D N A ライブラリを示す。

【 0 3 1 4 】

表 6 は、表 5 に示す c D N A ライブラリの作製に用いた組織およびベクターを示す付録である。

【 0 3 1 5 】

表 7 は、本発明のポリヌクレオチド及びポリペプチドの分析に用いたツール、プログラム、及びアルゴリズム、並びにその説明、引用文献、閾値パラメータを示す。 20

【 表 1 】

表1-1

インサイト プロジェクトID	ポリベブチド SEQ. ID NO:	インサイト ポリベブチドID	ポリスクレオチド SEQ. ID NO:	インサイト ポリスクレオチドID
1622129	1	1622129CD1	48	1622129CB1
1820078	2	1820078CD1	49	1820078CB1
1527017	3	1527017CD1	50	1527017CB1
1647264	4	1647264CD1	51	1647264CB1
1721989	5	1721989CD1	52	1721989CB1
1730581	6	1730581CD1	53	1730581CB1
1740714	7	1740714CD1	54	1740714CB1
1850596	8	1850596CD1	55	1850596CB1
1856109	9	1856109CD1	56	1856109CB1
1921719	10	1921719CD1	57	1921719CB1
2099829	11	2099829CD1	58	2099829CB1
2416915	12	2416915CD1	59	2416915CB1
2472784	13	2472784CD1	60	2472784CB1
2598981	14	2598981CD1	61	2598981CB1
2738075	15	2738075CD1	62	2738075CB1
2279049	16	2279049CD1	63	2279049CB1
2660904	17	2660904CD1	64	2660904CB1
3179424	18	3179424CD1	65	3179424CB1
2885096	19	2885096CD1	66	2885096CB1
2901076	20	2901076CD1	67	2901076CB1
3074572	21	3074572CD1	68	3074572CB1
1437895	22	1437895CD1	69	1437895CB1
1454656	23	1454656CD1	70	1454656CB1
121130	24	121130CD1	71	121130CB1
1257715	25	1257715CD1	72	1257715CB1
1342022	26	1342022CD1	73	1342022CB1
194704	27	194704CD1	74	194704CB1
607270	28	607270CD1	75	607270CB1
758546	29	758546CD1	76	758546CB1
866043	30	866043CD1	77	866043CB1
927065	31	927065CD1	78	927065CB1
938071	32	938071CD1	79	938071CB1

表1-2

インサイト プロジェクトID	ポリベプサド SEQ ID NO:	インサイト ポリベプサドID	ポリスクレオサド SEQ ID NO:	インサイト ポリスクレオサドID
3295984	33	3295984CD1	80	3295984CB1
4545237	34	4545237CD1	81	4545237CB1
4942964	35	4942964CD1	82	4942964CB1
5702144	36	5702144CD1	83	5702144CB1
5862945	37	5862945CD1	84	5862945CB1
6319547	38	6319547CD1	85	6319547CB1
000124	39	000124CD1	86	000124CB1
1659474	40	1659474CD1	87	1659474CB1
2267892	41	2267892CD1	88	2267892CB1
2670307	42	2670307CD1	89	2670307CB1
4524210	43	4524210CD1	90	4524210CB1
5584860	44	5584860CD1	91	5584860CB1
5807892	45	5807892CD1	92	5807892CB1
3210044	46	3210044CD1	93	3210044CB1
4942454	47	4942454CD1	94	4942454CB1

表2-1

ポリペプチド SEQ ID NO:	インサート ポリペプチドID	GenBank ID NO:	確率スコア	GenBank 相同体
1	1622129CD1	g8927590	1.00E-136	[fl1] [ヒト] {AF281133} exosome 成分 Rrp41
3	1527017CD1	g4689132	8.20E-87	30S リボソームタンパク質S7相同体[ヒト]
4	1647264CD1	g6651037	2.80E-38	RNA 結合タンパク質に類似[近交系マウス]
5	1721989CD1	g868267	9.80E-20	リボソームタンパク質L14に低い類似性 (SP:RL14_CHLTR, P28533) [線虫] Wilson, R. S. (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of <i>C. elegans</i> , <i>Nature</i> 368:32-38.
6	1730581CD1	g3721340	7.70E-91	NO27 [アフリカンマゲエル]
7	1740714CD1	g2570925	6.00E-139	運動ニューロンタンパク質相互作用 タンパク質1の生存: SIP1; SMタンパク質 相互作用タンパク質1 [ヒト] Fischer, U. S. (1997) The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis, <i>Cell</i> 90:1023-1029.
8	1850596CD1	g619302	2.80E-129	RNA 結合タンパク質=Merx (選択的にスプライシングされた). マウス 骨形成細胞系 POC4. Dahl, D. M. (1994) Pleiotropic effects of the mouse lethal yellow (Ay) mutation explained by deletion of a maternally expressed gene and the simultaneous production of agouti fusion RNAs, <i>Development</i> 120:1695- 1708.
9	1856109CD1	g2688625	4.70E-05	リボスレアーゼ III (rnc) [<i>Borrelia burgdorferi</i>]. Fraser, C. M. S. (1997) Genomic sequence of a Lyme disease spirochete, <i>Borrelia burgdorferi</i> , <i>Nature</i> 390:580-586.
10	1921719CD1	g1842111	1.00E-10	decoy [シロイヌナズナ]
11	2099829CD1	g6015629	6.20E-89	筋タンパク質84 [マウス]
12	2416915CD1	g8347090	1.00E-131	[fl1] [マウス] 推定上のhフィンガータンパク質 FL121

【表 4】

10

20

30

40

表2-2

ポリペプチド SEQ ID NO:	インサイト ポリペプチドID	GenBank ID NO:	標準スコア	GenBank 相同体
13	2472784CD1	g2098575	8.60E-163	P25451_2 [ヒト] 推定上の切断及びポリアダニル化 特異的因子[シロイヌナズナ] (Lin, X.ら (1999) Nature 402: 761-768)
14	2598981CD1	g4220489	2.80E-141	[ヒト] RNA 結合タンパク質 [ヒト] [Schizosaccharomyces pombe] 推定上のミトコンドリアリボソームタンパク質: L34ファミリー APOBEC1 (Apolipoprotein B mRNA editing protein) 及びPhorbollinに類似した推定上の新規のタンパク質 [ヒト] [ヒト] [ドブネズミ] セリン-アルギニン-リッチ スプライシング調節タンパク質 SRP86
15	2738075CD1	g5531845	1.00E-12	RNAポリメラーゼ I 調節因子 (PAP53)
18	3179424CD1	g7106069	8.00E-08	[マウス] (Hanada, X.ら (1996) EMBO J. 15: 2217-2226)
19	2885096CD1	g5102832	3.50E-79	ATP 依存性 RNA ヘリカーゼ #45 [ヒト] リボソームタンパク質 S15 [Thermotoga maritima] NO27 [アフリカツメガイル] pth [結核菌] リボソームタンパク質 S14 [Cricketulus griseus] 推定上のアルギニン-アスパラギン酸-リッチ RNA 結合タンパク質 [シロイヌナズナ]
20	2901076CD1	g7158880	0	[カンジダアルビカンス] リボソームタンパク質 S9 小サブユニット前駆体
21	3074572CD1	g1381029	0	推定上の40S リボソームタンパク質 [Schizosaccharomyces pombe] スプライシング因子 hPRP17 [ヒト] リボスクレア-ゼ P タンパク質サブユニット p14 [ヒト]
22	1437895CD1	g2696613	6.60E-140	[ヒト] (BC000940)
23	1454656CD1	g4981903	5.30E-09	スプライシング因子、アルギニン/セリン-リッチ (SRP75) に類似
24	121130CD1	g3721940	8.50E-92	
25	1257715CD1	g1870014	1.60E-26	
26	1342022CD1	g304525	1.60E-76	
27	194704CD1	g1699023	1.90E-46	
28	607270CD1	g4138828	1.60E-18	
29	758546CD1	g2440181	4.00E-46	
30	866043CD1	g3283220	8.60E-33	
31	927065CD1	g4100563	2.60E-58	
32	938071CD1	g12654241	1.00E-102	

【表 5】

10

20

30

40

表2-3

ポリペプチド SEQ ID NO:	インサート ポリペプチドID	GenBank ID NO:	得率スコア	GenBank 相合体
33	3295984CD1	g673454	0	精子細胞核周囲RNA結合タンパク質 [マウス] Schumacher, J. M. & (1995) J. Cell Biol. 129:1023-1032
34	4545237CD1	g6899218	1.60E-16	リボソームタンパク質 S5 [ウレアプラズマ-ウレアリチカム]
35	4942964CD1	g10803047	1.00E-68	[fl][トウモロコシ]40Sリボソームタンパク質 S24
36	5702144CD1	g3328106	1.10E-82	転写終結因子[ヒト] Zhang, Y. and Spremulli, L.L. (1998) Biochim. Biophys. Acta 1443:245-250
37	5862945CD1	g1001933	1.30E-08	リボソームタンパク質 L22 [Thermus thermophilus]
38	6319547CD1	g8573021	1.00E-07	[5' incm] [Leishmania major] ポリA結合タンパク質 1
39	000124CD1	g3860586	1.10E-43	ポリ(A)ポリメラーゼ (pcnB) [熱帯性フスリケツチア]
40	1659474CD1	g2950473	6.70E-13	DNA 依存性 rnaポリメラーゼポリペプチド [分裂酵母]
41	2267892CD1	g3646126	5.70E-205	ADP 依存性 RNA ヘリカーゼ[ヒト]
42	2670307CD1	g3152934	6.20E-176	[マウス] Jun coactivator Jab1 Aravind, L. and Ponting, C.P. (1998) Protein Sci. 7:1250- 1254
43	4524210CD1	g3258435	1.80E-38	アミノ酸389の長さの仮定上の核小体タンパク質 [Pyrococcus]
44	5584860CD1	g3879784	1.50E-116	RNA 認識モチーフに類似 (aka RRM, RBD) [線虫]
45	5807892CD1	g1573164	5.70E-15	リボソームタンパク質 S16 (rps16)
46	3210044CD1	g7226601	1.30E-37	Gln-tRNA (Gln) アミドトランスフェラーゼ サブユニット A [熱帯性菌] Tettelin, H. & (2000) Science 287:1809-1815.
47	4942454CD1	g790508	5.40E-51	60S 酸性リボソームタンパク質 [トウモロコシ] Goddemeier, M.L. & (1996) Plant. Mol. Biol. 30:655-658.

表3-1

SEQ ID NO.	インサイト ID 番号	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列、モチーフ、 及びドメイン	分析方法及び データベース
1	1622129CD1	245	S82, S119, S174, T226		アミノ酸 tRNAリガーゼモチーフ {Aa_Trna_Ligase_Ii_lgcg_motif): Y12-D36 3'エキソリボスクレア-ゼファアミリー (Rnase_PH): R13-A220 リボスクレア-ゼフ タンパク質:BL01277C:P117-L147 NUCLEOTIDYLTRANSFERASE TRANSFERASE POLYRIBONUCLEOTIDE PROTEIN PHOSPHORYLASE POLYNUCLEOTIDE RIBONUCLEASE PH PNPASE RNABINDING: PD002075: R13-L228 NUCLEOTIDYLTRANSFERASE; POLYRIBONUCLEOTIDE; PHOSPHORYLASE; POLYNUCLEOTIDE DOMAIN: DM03520 P50849 1-615: L6-R222 RIBONUCLEOPROTEIN HETEROGENEOUS NUCLEAR U SCAFFOLD ATTACHMENT FACTOR A HNRNP PROTEIN (ROU(1)): PD024707: P34-Y100 (p = 1.3e-05) シグナル切断: M1-A54 リボソームタンパク質 S7p/S5: S4-W178 リボソームタンパク質 S7タンパク質; BL00052A: I74-A120, BL00052B: K145-R171 リボソームタンパク質 S10タンパク質: BL00361A: V107-K122 リボソームタンパク質: ribosomal_s7.prf: M1-H83 RIBOSOMAL PROTEIN S7 30S rRNA- BINDING CHLOROPLAST 40S/MITOCHONDRION S5: PD00817: S5-W178 RIBOSOMAL PROTEIN S7: DM00334 P29765 26-155: P31-R177	MOTIFS HMMER-PFAM BLIMPS-BLOCKS BLAST-PRODOM BLAST-DOMO BLAST-PRODOM SFSCAN HMMER-PFAM BLIMPS-BLOCKS BLIMPS-BLOCKS PROFILES SCAN BLAST-PRODOM BLAST-DOMO
2	1820078CD1	118	T20	N77		
3	1527017CD1	179	S4, S5, S9, S10, S21, S67, S143			

10

20

30

40

30

[illegible]

表3-3

SEQ ID NO:	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ(signature)配列、モチーフ、及びドメイン	分析方法及びデータベース
9	1856109CD1	332	S61, S121, T161, T179, S206, T224, S251, T322,	N82, N249, N315	シグナルペプチド(signal_peptide): M1-G30 二本鎖RNA結合モチーフ(dsm): P237-L304 (score = 0.1) HYPOTHETICAL 49.1 KD PROTEIN P02A9.4 IN CHROMOSOME III: PDI40911: D67-R311 PROTEIN RNA-BINDING RNA REPEAT DEAMINASE HYDROLASE ADENOSINE DOUBLE-STRANDED III NUCLEAR: PD001171:P237-L304 (p = 0.0090) シグナル切断(signal_cleavage): M1-G42	HMMER HMMER-PFAM BLAST-PRODOM BLAST-PRODOM
10	1921719CD1	279	S28, S85, T138, T183, T194, T214, T256		DECOY 60S RIBOSOMAL PROTEIN L30 MITOCHONDRIAL PRECURSOR YML30 MITOCHONDRION TRANSIT: PD037326:R136-P209 (p = 1.6e-08) リボソームタンパク質L10 (Ribosomal_L10): K18-T117 PROTEIN RIBOSOMAL SIMILAR 60S ACIDIC PO UFDACAP1 INTERGENIC REGION: PD037726: M1-G213 RAT ACIDIC RIBOSOMAL PROTEIN: DM00904[P29764]1-318:V10-S212 SUPPRESSOR OF SABLE RNA-BINDING NUCLEAR HOMOLOG PD032978:K221-K288	SPSCAN BLAST-PRODOM
11	2099829CD1	239	S4, S11, T15, T36, Y39, S80, S93, T105, T108, T120, Y124, S171, T203, S212, S225, S229, S233, S235,	N106		HMMER-PFAM BLAST-PRODOM
12	2416915CD1	291	T18 T20 T32 T37 S77 T113 S165 T235 S59 S112 T285 S289 Y76			BLAST-DMO BLAST-PRODOM

10

20

30

40

表3-4

SEQ ID NO:	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列、モチーフ、 及びドメイン	分析方法及び データベース
13	2472784CD1	451	S29 S131 S337 S338 S343 S399 S290 T389		真核生物の推定上の RNA 結合領域 RNP-1 シグネチャ K393-F400 ATP/GTP 結合部位モチーフ A (P-loop) G22-S29 RNA 認識モチーフ (RRM, RBD, または RNP ドメイン) I354-L425 真核生物 RNA 結合 BL00030A: I354-P372 BL00030B: K393-D402 F25451_2 PD057917: L144-R206 PD056050: G305-R353 RIBONUCLEOPROTEIN REPEAT DM00012 Q10355 25-106: W347-S428 DM00012 P32588 154-238: D350-L425 DM00012 P31483 87-172: A349-K424 DM00012 Q05966 1-83: D351-S428 POLYADENYLATION CLEAVAGE SPECIFICITY RNA BINDING PD005421: Y179-I370 N66	MOTIFS MOTIFS HMMER_PFAM BLAST_BLIMPS BLAST_PRODUM BLAST_DOMO BLAST_PRODUM MOTIFS SPSCAN HMMER_PFAM BLIMPS_BLOCKS
14	2598981CD1	600	T209 T254 T520 T577 S213 T272 T512 Y50 Y175		真核生物の推定上の RNA 結合領域 RNP-1 シグネチャ K51-F58 シグナル切断: M1-A64 RNA 認識モチーフ (RRM, RBD, または RNP ドメイン) RRM: V12-A83 真核生物の推定上の RNA 結合領域 RNP-1 タンパク質 BL00030A: V12-F30 BL00030B: K51-D60	MOTIFS SPSCAN HMMER_PFAM BLIMPS_BLOCKS
15	2738075CD1	217	T207 T47 S198 T22 S155 S193 T194 S210			

10

20

30

40

表3-5

SEQ ID NO.	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列、モチーフ、 及びドメイン	分析方法及び データベース
15					真核生物の増殖上の RNA 結合領域 RNP-1 シグネチャ RNP 1: pf: N23-I85 RIBONUCLEOPROTEIN REPEAT DM00012 S20940 151-238: V12-A86 DM00012 Q04836 234-321: V12-A86 DM00012 P19339 205-288: S10-N73 DM00012 P38159 3-84: P7-A86	PROFILES CAN BLIMPS_DOMO
16	2279049CD1	319	S275 T20 S234 T264 S80 T119 T268	N153	アミノアシル転移 RNA シンターゼ クラス II シグネチャ H255-E278	MOTIFS
17	2660904CD1	108			リボソームタンパク質 S15シグネチャ G30-G101	PROFILES CAN
18	3179424CD1	92	S71 T12		シグナル 切断: M1-S15 リボソームタンパク質 L34 N51-H92 リボソームタンパク質 L34 BL00784: G50-R87	SPSCAN HMMER_PFAM BLIMPS_BLOCKS
19	2885096CD1	268	T78 T162 S241 Y91 T198 S222	N109 N193	シナジン及び デオキシシナジン酸デアミナーゼ 亜結合領域シグネチャ H144-V182 シグナル 切断: M1-S21	MOTIFS SPSCAN
					シナジン及びデオキシシナジン酸デアミナーゼ 亜結合領域シグネチャ BL00903: Y169-C178 APOLIPROTEIN B MRNA EDITING PROTEIN DM04741 P51908 1-228: K129-V239 DM04741 P41238 1-235: H144-V239 DM04741 A53853 1-236: H144-V239	BLIMPS_BLOCKS BLAST_DOMO

表3-6

SEQ ID NO.	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列、モチーフ、 及びドメイン	分析方法及び データベース
20	2901076CD1	624	S329 S351 S359 S367 T380 S502 S519 S520 S558 S61 S297 S319 S324 S487 S493 S505 S508 T511 T514 S524 S543 S553 S573 S618 T159 S175 T194 S557 T619 S8 S20 T64 T130 T144 T153 S331 S364 S380 S409 S296 S385 T419 T42 S137 T174 S221 S230 T239 S265 S390 S11 S25 S70 S341 S383 S561 S671 T732 S13 S49 T325 S351 S397 T642 T658 Y247 Y293 Y602	N25 N254 N339 N485 N570 N616	RNA 認識モチーフ (RRM, RED,または RNPドメイン) rrm: V184-I253 rrm: I21-V93 ARGININE RICH SPliceosome SPLICING FACTOR PD037489:V20-G139 TYPE B REPEAT REPEAT DM05511 S26650 1-1203:P264-R539 DM05511 P18583 113-1296:P264-R539 RNA POLYMERASE I DNA DIRECTED TRANSFERASE TRANSCRIPTION NUCLEAR PROTEIN PD025048:Q21-R415	PROMER_Pfam BLAST_PRODUM BLAST_DOMO BLAST_PRODUM BLIMPS-BLOCKS
21	3074572CD1	419		N93 N161 N237	DEAD及びDEAHボックスファミリーATP依存性 ヘリカーゼタンパク質: BL00690A:G85-Q94 BL00690B:T116-E133 BL00690C:I182-S191 DEAD及びDEAHボックスファミリーATP依存性 ヘリカーゼシグネチャ (deah_atp_helicase.prf): D163-P209 ATP依存性RNAヘリカーゼ: PD01259:F401-H544 DEAHボックスサブファミリーATP依存性 ヘリカーゼ: DM00649 P53131 84-705:L55-Y677 Atp_Gcp_A: G85-S92	BLIMPS-BLOCKS PROFILES SCAN BLAST_PRODUM BLAST_DOMO MOTIFS
22	1437895CD1	743		N10 N154 N425 N473 N560 N577		

10

20

30

40

表3-7

SEQ ID NO.	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列、モチーフ、 及びドメイン	分析方法及び データベース
23	1454656CD1	284	S57 S160 T232 T244 T279 S88 S165 T200		リボソームタンパク質 S15 (Ribosomal_S15): S151-E215 リボソームタンパク質 S15シグネチャ (ribosomal_s15.ptf): P146-E215	HMMER-PFAM PROFILESSCAN
24	121130CD1	248	T14 S49 S74 S132 S18 T242		NO27タンパク質(PD173812): M1-L112 ヌクレオリン: DM02740 S32644 G30-703:R140-G204	BLAST-PRODOM BLAST-DOMO
25	1257715CD1	214	S137 T112 S181		ペプチジル-terNAヒドロラーゼ (Pept-terNA_hydro): M31-Q208 ペプチジル-terNAヒドロラーゼ: BL01195B:L88-G99 BL01195C:V116-N154 BL01195D:M156-G164 ペプチジル-terNAヒドロラーゼ (PD005324): M32-T199	HMMER-PFAM BLIMPS-BLOCKS BLAST-PRODOM BLAST-DOMO
26	1342022CD1	184	T2 S16 S19 S81 S172 T173 S69	N71 N136	リボソームタンパク質 S11 (Ribosomal_S11): G62-R180 リボソームタンパク質 S11: BL00054A:G62-S102 BL00054B:K139-R180 リボソーム_S11: G144-L184 リボソームタンパク質 S11 (PD001010): G62-R180 大腸菌リボソームタンパク質 S11: DM00861 P19950 16-148:G51-L184 リボソーム_S11: D164-D171	HMMER-PFAM BLIMPS-BLOCKS PROFILESSCAN BLAST-PRODOM BLAST-DOMO
27	194704CD1	371	S270 S280 S291 S317 T17 S120 S152 S257 T283 S321 S339 S342 S360 T52 T111 T231 S332 Y173		アスバラギン酸-アルギニン-リッヂ mENA 結合タンパク質 (PD017473): F30-K230 RNP-1: DM03434 P08621 359-483:R204-K323	MOTIFS BLAST-PRODOM BLAST-DOMO

10

20

30

40

表3-8

SEQ ID NO:	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ (signature) 配列、モチーフ、 及びドメイン	分析方法及び データベース
28	607270CD1	396	S13 T63 S64 T153 S252 T278 S290 T392 T68 T100 S214 S270 T357 Y218 Y262		シグナル ペプチド: ML-A25 リボソームタンパク質 S9/S16 (Ribosomal_S9): G274-R396 リボソームタンパク質 S9: BL00360A:K275-Q301 BL00360B:F317-L352 BL00360C:L370-R396 リボソームタンパク質 S9シグネチャ (ribosomal_s9.prf): I306-R396 リボソームタンパク質 S9 (PD001627): G274-R396 リボソームタンパク質 S9: DM00779 F38120 149-277:D263-R396 リボソーム_S9: G334-L352 S4ドメイン (S4): R109-D156 リボソームタンパク質 S4: DM00205 F32899 101-174:V101-Y173 WDドメイン, G-βリポート: WD40:G20-D58 WD40:B60-D100 WD40:M199-S239 WD40:A243-H282	HMMER HMMER-PFAM SLIMPS-BLOCKS PROFILESKAN BLAST-PRODOM BLAST-DOMO MOTIFS HMMER-PFAM BLAST-DOMO HMMER-PFAM
29	758546CD1	184	S153 T159 S163 Y48			
30	866043CD1	282	T10 S67 T76 T103 S109 S136 S153 S200 S239 T5 T57 S131			
31	927065CD1	125	T65 Y21			MOTIFS

表3-9

SEQ ID NO	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグナチャ(s)gnature)配列、モチーフ、 及びドメイン	分析方法及び データベース
32	938071CD1	365	S23 S24 S132 T172 S186 S187 S201 S202 S210 S254 S261 S116 S132 T161 S176 T177 S178 S213 S215 S217 S219 S223 S224 S257 S265 S269 T282 T300 S301 S325 S350 S24 S27 S33 S36 S71 S110 S125 S149 S193 S202 S246 Y324	NC9		MOTIFS
33	3295984CD1	672	Y137 Y297 S24 S36 T56 T58 T120 S182 S343 T345 T427 T455 S465 S470 S474 S490 S496 S3 T52 T120 T159 S182 S201 S333 T345 S435 T438 S510	N72 N180 N472 N476 N482 N485 N486 N489	L381-L402ロイシンジッパー 二本鎖RNA結合 タンパク質: G554-A567 二本鎖RNA結合モチーフ: L388-M451, G511-L574 RNA結合タンパク質 L448-M464, G411-V424, N558-A567 ZINC FINGER RNA BINDING SPERMATID PROTEIN R80-D329 SPERMATID PERINUCLEAR RNA BINDING PROTEIN P575-G648, P330-L392, G452-N513 TRANSCRIPTION; RNA; SPERMATID; PERINUCLEAR M1-P369 DOUBLE-STRANDED RNA BINDING DOMAIN N370-T461, T494-K584	MOTIFS BLIMPS_PPFAM HMMER_PPFAM BLIMPS_PRODOM BLAST_PRODOM BLAST_DOMO

10

20

30

40

表3-10

SEQ ID NO.	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(s)配列、モチーフ、 及びドメイン	分析方法及び データベース
34	4545237CD1	430	S240 T94 S161 T214 T232 S354 T118 T232 T262 S295 S386		シグナルペプチド	HMER
					M1-G20	
					リボソームタンパク質 S5	HMER_P3AM
					I222-G352	
					リボソームタンパク質 S5 プロ	BLIMPS_BLOCKS
					T220-A271	
					I303-S339	
					リボソームタンパク質 S5 シグネチャ	PROFILES CAN
					I222-E284	
					RIBOSOMAL PROTEIN S5	BLAST_PROD OM
35	4942964CD1	137	T11 T14 T19 S109 Y95 S41 T104	N39	I222-N341	BLAST_DOMO
					RIBOSOMAL PROTEIN S5	
					R226-G367	
					リボソーム S24e	MOTIFS
					F70-N84	
					F28-K111リボソームタンパク質 S24e	HMER_P3AM
					リボソームタンパク質 S24e シグネチャ	PROFILES CAN
					E47-A103	
					リボソームタンパク質 S24e	BLIMPS_BLOCKS
					V10-E54, I61-K105	
					アルギニンリグレッサー	BLIMPS_P3AM
					N20-G71	
					タンパク質 S24E	BLAST_DOMO
					V10-K146	
					RIBOSOMAL 40S S24 S24E	BLAST_PROD OM
					V29-K111	

表3-11

SEQ ID NO.	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列、モチーフ、及びドメイン	分析方法及びデータベース
36	5702144CD1	380	Y169 S247 T45 T342 S362 S371 T208 T243 S321 S377	N255 N340	ロイシン_ジッパー L57-L78 L64-L85 L71-L92 RF_Prok_I R245-V261 シグナルペプチド ML-G20 真核生物型クラス I ペプチド鎖 終結因子 RF-1 G138-R338 真核生物型クラス I ペプチド鎖 終結因子シグネチャ E123-R161, I184-Q226, D238-K284, G316-R338 PROKARYOTIC-TYPE CLASS I PEPTIDE CHAIN RELEASE FACTORS L224-S377 PEPTIDE CHAIN RELEASE FACTOR G138-R338 シグナルペプチド ML-T35 リボソームタンパク質 L22 H69-K105, V128-L172 リボソームタンパク質 L22シグネチャ R131-E192 Rop_1 R69-F76 真核生物の増殖上の RNA 結合領域 RNP-1 シグネチャ L45-K97 RNA 認識モチーフ L27-V101 真核生物 RNA 結合 L27-L45 R69-T78	MOTIFS MOTIFS HMMER HMMER_PFAM BLIMPS_BLOCKS BLAST_DOMO BLAST_PRODOM SPSCAN BLIMPS_BLOCKS PROFILESAN MOTIFS PROFILESAN HMMER_PFAM BLIMPS_BLOCKS
37	5862945CD1	206	S42 S141 Y83			
38	6319547CD1	190	S96 T140 S188 T178 T179 S20 T78 S128	N12 N136		

表3-12

SEQ ID NO:	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列、モチーフ、 及びドメイン	分析方法及び データベース
39	000124CD1	434	S18 S47 T165 S274 T307 S389 S400 T29 T112 T124 S289 T359 Y368	N272 N152 N199 N217	ポリAポリメラゼファミリー T110-E276 RNA BINDING PROCESSING POLYNUCLEOTIDE G63-G248 RNA BINDING POLYMERASE L51-G248	HMMER_PFAM BLAST_PRODOM BLAST_DOMO MOTIFS
40	1659474CD1	339	S60 T244 S329 T69 T254 T315 Y238			
41	2267892CD1	599	T18 T79 S86 S129 S213 S282 S283 S330 S344 T397 S457 T503 T569 S589 S22 T92 T585 S77 S99 S110 S381 S492 T517		Atp_Gtp_A A209-T216 ヘリカーゼ保存C末端 ドメイン Y496-T503 DEAD/DEAHボックスヘリカーゼ Q178-E389 ヘリカーゼ保存C末端 ドメイン K436-G507 DEADボックスサブファミリーATP依存性 ヘリカーゼ G184-P222, M225-T250, V312-L335, V485-G510 DEAD-BOX SUBFAMILY ATP-DEPENDENT HELICASES N179-I542 RNA BINDING NUCLEAR DNA FACTOR T422-G507, D182-S258 Mov34/WPN/PAD1ファミリー H50-A314 SUBUNIT 26S MOV34 S12 PAD1 HOMOLOG F52-R282 PAD1 関連タンパク質 I57-S254	MOTIFS BLIMPS_PFAM HMMER_PFAM HMMER_PFAM BLIMPS_BLOCKS BLAST_DOMO BLAST_PRODOM HMMER_PFAM BLAST_PRODOM BLAST_DOMO
42	2670307CD1	334	S24 S148 S231 T257 T310 S270 S307	N332		

10

20

30

40

表3-13

SEQ TID NO.	インサイト ポリペプチド ID	アミノ酸 残数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列、モチーフ、 及びドメイン	分析方法及び データベース
43	4524210CD1	448	S159 T285 S319 T355 T413 S417 Y182 T33 S159 S265		Nol1_Nop2_Sun F296-G307 NOL1/NOP2/sun フアミリー Y194-K369 NOL1/NOP2/sun フアミリー I217-I231, G239-G262, F296-G309, K342-E367 NUCLEOLAR SUN P120 PROLIFERATING CELL ANTIGEN R176-E379	MOTIFS HMMER_Pfam ELIMPS_BLOCKS ELAST_PRODUM
44	5584860CD1	420	S218 S86 S102 S126 S142 T187 T242 T244 S102 Y181	N294	NOL1/NOP2/PNU FAMILY L197-V363 RNA 凝縮モチーフ L234-V300 RNA 結合モチーフ L234-F252 RNA BINDING PROTEIN L223-E314	BLAST_DOMO HMMER_Pfam ELIMPS_BLOCKS ELAST_PRODUM
45	5807892CD1	137	T18 T27 S122 S60 T105 T125 T130		リボソームタンパク質 S16 G24-S84 リボソームタンパク質 S16 H16-L51, L68-A94	HMMER_Pfam ELIMPS_BLOCKS ELAST_PRODUM
46	3210044CD1	556	T30, T64, T153, T210, S231, S254, S344, T353, S389, S458	N165, N199	RIEOSOMAL S16 NUCLEASE G25-A91 AMIDASES; DM00646 IAS3101 91-465:V94-G338 アミダーゼタンパク質: BL00571: N219-S270 アミダーゼシグネチャ (amidases.ppf): G235-G278 膜貫通ドメイン {transmem.domain}: I33-P58 アミダーゼ: D93-P313, R443-L537 シグナル切断: M1-G56	ELAST_PRODUM BLAST_DOMO ELIMPS_BLOCKS PROFILES CAN HMMER HMMER_Pfam SPSCAN

表3-14

SEQ ID NO.	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列、モチーフ、及びドメイン	分析方法及びデータベース
47	4942454CDI	111	S19, S47, S101, S108		RAT ACIDIC RIBOSOMAL PROTEIN P1: DM00632 S54179 1-112: M1-D111 リボソームタンパク質 P2シグネチャ: PR00456E: A75-A89; PR00456F: K38-L109 RIBOSOMAL PROTEIN, ACIDIC 60S PHOSPHORYLATION P2, P1, L12 MULTIGENE FAMILY: PD001928: M1-D111 60S 酸性リボソームタンパク質 (60S_ribosomal): M1-D111	ELAST-DOMO BLIMPS-PRINTS BLAST-PRODOM HMMER-PFAM

10

20

30

40

表4-1

ポリスクレオナド SEQ ID NO:	インサイト ポリスクレオナドID	配列長	選択された断片	配列断片	5'位置	3'位置
48	1622129CB1	882	1-59	642017H1 (BRSTNOT03) 483782F1 (HNT2RAT01) 71164239V1	1 176 647	270 882 1220
49	1820078CB1	1220	1-41, 833-1220	70024116DL 1293767T6 (PGANNOT03) 6571082H1 (NCLDTXN05)	564 403 1	1220 1203 490
50	1527017CB1	2020	1-209, 1415-2020	1541283H1 (SINTTUT01) 1824377H1 (GBLATUT01) 1659084T6 (URETUT01) 1527017T1 (UCMCL5T01) SBMA03026F1	1266 1 1279 744 324	1473 265 2020 1407 858
51	1647264CB1	637	1-40	1438348F1 (PANGNOT08) 2735749H2 (OVARNOT09) 1647264H1 (PROSTUT09) 91136841	226 286 149 1	769 508 384 637
52	1721989CB1	717	1-21	1721989F6 (BLADNOT06) 079812F1 (SYNORAB01)	1 146	419 717
53	1730581CB1	2061	1-24, 1219-1514	1525770T1 (UCMCL5T01) 1457506F6 (COLNFFUT02) 915914R1 (BRSTNOT04) 1730581F6 (BRSTTUT08) 2693987T6 (LUNGNOT23) 102784H1 (ADRENOR01)	1109 486 306 1546 953 1	1566 1012 884 2061 1565 406
54	1740714CB1	1307	1-43, 1111-1142	1740714CT1 (HIPONON01) 3074962H1 (BONEUNT01) 820443R1 (KERNANOT02)	1 1 493	1307 277 1131
55	1850596CB1	1357		2018418H1 (THP1NOT01) 1989582H1 (CORPNOT02) 1855681H1 (HNT3AZT01) 3563740H1 (SKINNOT05) 3147502H1 (PENCNOT05)	813 986 1 141 282	1160 1357 270 470 643

10

20

30

40

表4-2

ポリスクレオチド SEQ ID NO:	インサイト ポリスクレオチドID	配列長	選択された断片	配列断片	5'位置	3'位置
56	1856109CB1	1749	1-29, 1201-1749	6708693HL (HEADIR01)	2411	1738
				1518037F1 (BLADTUT04)	460	979
				2367008F6 (ADREN07)	1547	1749
				875652T1 (LUNGAST01)	1073	1718
				1981031R6 (LUNGSTUT03)	563	1109
57	1921719CB1	991	1-79	1806093F6 (SINTNOT13)	1	525
				1921719T6 (BRSTTUT01)	335	973
				902176H1 (BRSTTUT03)	697	991
				1522552F1 (BLADTUT04)	1	429
58	2099829CB1	1188	1024-1188	91155846	718	1065
				1868148T6 (SKINBT01)	402	972
				1984185T6 (LUNGAST01)	248	889
				3135711H1 (SMCNOT01)	1	274
				313888H1 (LUNGNOT02)	690	1006
59	2416915CB1	1454	1-22	1569836F6 (UTRSNOT05)	37	519
				3177069T6 (UTRSNOT04)	830	1454
				2172869H1 (ENDCNOT03)	761	1026
				3873555H1 (HEARNOT06)	446	741
				3584701H1 (293TF4T01)	1	315
				1926194R6 (BRSTNOT02)	1161	1588
				2701446H1 (OVARTUT10)	775	1135
				908518R2 (COLNNOT09)	1	616
60	2472784CB1	1588	1-242, 288-726	2834469H1 (TYMNOT03)	1	274
				599142R6 (BRSTNOT02)	1229	1588
				2023191F6 (CONNNOT01)	1038	1588
				2470222F6 (THP1NOT03)	546	1113
				1727193H1 (PROSNOT14)	387	634
				1593669X16C1 (BRINNOT14)	289	954
				SEZA06347V1	936	1527
				SEZA04028V1	1236	2111
61	2598981CB1	2111	557-1153, 1-22			

10

20

30

40

表4-3

ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチドID	配列長	選択された断片	配列断片	5'位置	3'位置
61				1593669X11C1 (BRAINTOT14)	52	711
				SEZA02427V1	1159	1621
62	2738075CB1	1155	1-36, 597-623	2450029H1 (ENDANOT01)	1	227
				222536F1 (PANCNOT01)	15	622
				205557R6 (BRPINT01)	801	1155
63	2279049CB1	1673		222536R1 (PANCNOT01)	55	1155
				1597077F6 (BRAINTOT14)	445	1005
				438020T6 (THYENOT01)	938	1650
				1798393F6 (COLNNOT27)	1208	1673
				245885F6 (ENDANOT01)	1	458
64	2660904CB1	584	1-229, 519-584	3050984H1 (LUNGNOT25)	409	687
				71284514V1	387	584
				70937372V1	1	514
65	3179424CB1	978	1-189	3204102F6 (PENCNOT03)	1	632
				586088F1 (PROSNOT02)	307	978
66	2885096CB1	1055	1-278	1702519X13C1 (BLADTUT05)	127	778
				3887887H1 (UTRSNOT05)	813	1055
				187655F6 (LEUKNOT03)	552	1043
				2885096F6 (SINGNOT02)	1	454
				5635858H1 (UTRSTMR01)	1333	1585
67	2901076CB1	2220	1189-1491, 634-662	3524308H1 (ESOGTUN01)	333	644
				1285251F6 (COLNNOT16)	552	1152
				3254924X309D1 (OVARTUN01)	39	491
				1260590R1 (SYNORAT05)	892	1508
				5117929H1 (SMCBUNT01)	1	278
				1285251T1 (COLNNOT16)	792	1218
				1852576F6 (LUNGFT03)	1730	2220

10

20

30

40

表4-4

ポリスクレオチド SEQ ID NO:	インサイト ポリスクレオチドID	配列長	選択された断片	配列断片	5'位置	3'位置
67				2921502HL (SININOT04)	1457	1740
68	3074572CB1	1890	1-24, 1837-1890	4936943HL (OVARNON03)	1665	1904
				1558589FL (SPLNNOT04)	1011	1432
				SAEA01339FL	1325	1890
				157743R6 (THP1PLE02)	221	725
				SAEA01587FL	769	1343
				3173159HL (UTRSTUT04)	1	278
				SAEA01593R1	286	843
69	1437895CB1	2893	845-1749, 1-43, 2872-2893	2630813HL (COLNUTU15)	2042	2283
				4919358HL (TESTNOT11)	1384	1615
				3205068HL (PENCNOT03)	1623	1888
				3877384F6 (HEARNOT06)	305	799
				1493166R6 (PROSNON01)	2324	2884
				4840634HL (OSTENOT01)	903	1171
				3550295HL (BRONDIT01)	1	244
				6453094HL (COLNDIC01)	192	757
				1400075FL (BRAITUT08)	1046	1606
				91377484	2337	2893
				1437895T1 (PANCNOT08)	224	2863
				786771FL (PROSNOT05)	2707	2892
				1582477HL (DUODNOT01)	1592	1793
				91102494	2314	2892
				1437895FL (PANCNOT08)	1731	2278
				3628841P6 (COLNNOT38)	722	1098
				782659R1 (MYOMNOT01)	523	885
				938801R1 (CBRVNOT01)	505	876
				876916T1 (LUNGAST01)	280	867
70	1454656CB1	885	1-47	4843096HL (OSTENOT01)	23	314
				91617769	1	414

表4-5

ポリヌクレオチド SEQ ID NO.:	インサイト ポリヌクレオチドID	配列長	選択された断片	配列断片	5'位置	3'位置
71	121130CBI	1269	1-42	1457506P6 (COLNPF02)	530	1055
				915914R1 (BRSTN0T04)	352	929
				3730083H1 (SMCCNON03)	974	1269
				5217827H1 (BRSTN0T35)	258	521
				5376327H1 (BRAXN0T01)	14	254
72	1257715CBI	1066	835-1066, 1-22	91716816	1	395
				102784H1 (ADRENOR01)	63	452
				93048792	745	1066
				1680736H1 (STOMF0T01)	1	221
				1901049F6 (BLADTUT06)	320	825
73	1342022CBI	639	1-98	1731204F6 (BRSTTUT08)	17	669
				1908142T6 (CONNTUT01)	1	614
				2257149R6 (OVARTUT01)	100	639
				834691H1 (PROSNOT07)	1181	1420
				1701714F6 (BLADTUT05)	32	698
74	194704CBI	1420	822-852, 1051-1420	1822010F6 (GBLATUT01)	361	983
				2584972H1 (BRAITUT22)	1156	1420
				1395146H1 (THYRN0T03)	1	255
				2060393R6 (OVARN0T03)	807	1420
				607270H1 (BRSTTUT01)	1	265
75	607270CBI	1457	1-103, 680-892	607270X11 (BRSTTUT01)	104	745
				1558989F6 (SPLNNOT04)	448	973
				449792F1 (TLYMNOT02)	780	1457
				1488271H1 (UCMCL5T01)	429	691
				489544F1 (HNT2AGT01)	506	1184
76	758546CBI	1184	1-53	489544R1 (HNT2AGT01)	58	664
				2657989H1 (LUNG2TUT09)	1	229
				5278386H1 (MUSLNOT01)	693	891
				004388T6 (HMC1NOT01)	990	1515
				5039960H1 (COLHTUT01)	704	952
77	866043CBI	1638	1-561	866043R6 (BRAITUT03)	1266	1638

10

20

30

40

表4-6

ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチドID	配列長	選択された断片	配列断片	5'位置	3'位置
77				2295207R6 (BRSTNOT05)	171	638
				3398785H1 (UTRSNOT16)	1	226
				5421565H1 (PROSTM07)	468	715
				004388R6 (HMC1NOT01)	838	1355
78	927065CB1	701		SXAF04722V1	254	701
				SXAE03477V1	13	411
				g1545603	1	614
				2570862T6 (HIPOAZT01)	852	1407
79	938071CB1	1829	1070-1829, 513-561, 828-910, 689-768	2344188T6 (TESTTTT02)	168	744
				2061053R6 (OVARNOT03)	1077	1413
				2362859T6 (ADREN0T09)	296	882
				2371642H1 (ADREN0T07)	1210	1425
				g1141976	1335	1829
				3148824H1 (ADREN0N04)	1	274
				SCGA02064V1	287	1013
				SCGA02183V1	1784	2326
80	3295984CB1	2541	1-242, 2385-2541	261729R6 (HNT2AGT01)	1055	1668
				SCGA00610V1	1221	1778
				SCGA10196V1	1588	2118
				4526287F6 (LYMBTGT01)	1	532
				3813241H1 (TONSNOT03)	2255	2541
				SCGA05805V1	591	1162
				1561512F6 (SP1NNOT04)	219	830
				1639525F6 (UTRSNOT06)	1226	1647
				1849445F6 (PROSTUT09)	874	1378
				2013721X28C1 (TESTNOT03)	560	1154
				5882142H1 (L1VRNON08)	559	832
				5672020H1 (MMIR3DT01)	1	252
82	4942964CB1	735	704-735	4942964T6 (BFAIFEN05)	1	735

【表 2 6】

表4-7

ポリヌクレオチド Seq ID NO:	インサイト ポリヌクレオチドID	配列長	選択された断片	配列断片	5'位置	3'位置
83	5702144CB1	2614	1-93, 1719-1901	1481581F6 (CORNOT02)	724	1378
				2287840X12F1 (BRATNON01)	464	906
				1962628F6 (BRSTNOT04)	2107	2614
				2160148F6 (ENDCNOT02)	1866	2324
				2287840X14F1 (BRATNON01)	370	861
				1329248F1 (PANCNOT07)	1529	2094
				968129X11F1 (BRSTNOT05)	1093	1612
				1635830F6 (UTRSNOT06)	1	454
				280637F1 (LIVRNOT02)	106	736
				2244058F6 (PANCNOT02)	1	510
84	5862945CB1	736	1-34	2383076F6 (TSLTNOT01)	574	1046
				3014264H1 (MUSCNOT07)	509	809
				5571304F6 (TLYNNOT08)	142	782
				591563H1 (BRAVNOT02)	1	241
				000124T6 (D937NOT01)	1742	2266
				789548R6 (PROSTUT03)	1100	1600
				3519565R6 (LUNGNOT03)	1196	1782
				902790R6 (BRSTTUT03)	629	1123
				3039566F6 (BRSTNOT16)	1	611
				SBZA03954V1	482	1008
85	6319547CB1	1046	1-33	1865044F6 (PROGNOT19)	615	1041
				SAYA00579F1	352	833
				4005010F6 (ENDCNOT04)	1	533
				2290718X14F1 (BRATNON01)	726	1230
				3139431F6 (SMCCNOT02)	529	1043
				1970333F6 (UCMCL5T01)	888	1350
86	000124CB1	2266	1-32, 2085-2266			
87	1659474CB1	1041	927-954, 677-709, 833-857			
88	2267892CB1	2722	2347-2722			

10

20

30

40

表4-8

ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチドID	配列長	選択された断片	配列断片	5'位置	3'位置
88				3815067H1 (TKSNOT03)	1574	1844
				082326R1 (HUVSTB01)	1828	2448
				082326F1 (HUVSTB01)	2007	2710
				2287150X13F1 (BRAIN001)	1	546
				256073F1 (HMT2RAT01)	2581	2722
89	2670307CB1	1287	1-382	3925865H1 (KIDN0T19)	1219	1479
				1420665F1 (KIDN0T09)	1276	1828
				008243F1 (HMCIN0T01)	315	1287
				1295856F1 (PGANN0T03)	969	1287
				1304866F1 (PIACN0T02)	732	1287
90	4524210CB1	2226	509-1516	1353856F1 (LATRTUT02)	1	485
				SCIA00216V1	430	1034
				SCIA00181V1	1251	1780
				2256793X309B2 (OVARTUT01)	1645	2225
				SCIA03684V1	1161	1705
91	5584860CB1	2362	1-49, 1066-2362	2256793X318D4 (OVARTUT01)	1	517
				2058779H1 (OVARN0T03)	1999	2226
				SCIA02060V1	665	1237
				1965766R6 (BRSTN0T04)	1954	2362
				1850306T6 (LUNGFT03)	1698	2340
92	5807892CB1	731	1-68	842676R1 (PROSTUT05)	712	1279
				71054273V1 (SG0000314)	1	614
				2105411R6 (BRAITUT03)	1293	1767
				SAEB02217F1	1191	1752
				377688R6 (NEUTFT01)	541	1153
				3745266F6 (THYMN0T08)	1	495
				958760R1 (KIDNNO05)	416	731

表4-9

ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチドID	配列長	選択された断片	配列断片	5'位置	3'位置
93	3210044CE1	2088	2051-2088, 1-190, 648-1486, 1811-1843	70822015V1	557	1148
				7705613J1 (DPERETUE01)	1	541
				70821405V1	1195	1807
				70818955V1	511	1069
				6016193H1 (HNT2UNN03)	1127	1739
94	4942454CE1	660	1-23	70821946V1 (SG0000294)	1653	2088
				4942454T6 (BRA1FEN03)	3	660
				4942454F6 (BRA1FEN03)	1	589

10

20

30

40

表5-1

ポリヌクレオチド SEQ ID NO:	インサイト プロジェクトID	代表的ライブラリ
48	1622129CB1	HNT2RAT01
49	1820078CB1	LUNGNOT20
50	1527017CB1	LUNGNOT14
51	1647264CB1	LUNGNOT27
52	1721989CB1	ISLJN0T01
53	1730381CB1	LUNGNOT03
54	1740714CB1	TYMNOT02
56	1856109CB1	BLADTUT04
57	1921719CB1	BRSTTUT01
58	2099829CB1	ENDANOT01
59	2416915CB1	BLADTUT05
60	2472784CB1	BRSTNOT02
61	2598981CB1	LIVRN0N08
62	2738075CB1	BRAITUT02
63	2279049CB1	BRAINOT14
64	2660904CB1	LUNGPUT09
65	3179424CB1	KIDNNOT09
66	2885096CB1	BLADTUT04
67	2901076CB1	ENDCN0N02
68	3074572CB1	SINUFET03
69	1437895CB1	PANCNOT08
70	1454656CB1	ADREN0T03
71	121130CB1	SPLNFET01
72	1257715CB1	MENITUT03
73	1342022CB1	COLNFET02
74	194704CB1	OVARNOT03
76	758546CB1	COLNNOT11
77	866043CB1	BRAITUT03
78	927065CB1	LNODNOT03
79	938071CB1	EPIPNON05
80	3295984CB1	THP1NOT03

表5-2

ポリスクレオチド SEQ ID NO:	インサイト プロジェクトID	代表的ライブラリ
81	4545237CB1	SPLNOT04
82	4942964CB1	BRAIFEN05
83	5702144CB1	PANCNOT07
84	5862945CB1	KERANOT01
85	6319547CB1	PROSNOT16
86	000124CB1	KIDNNOT05
87	1659474CB1	KERANOT01
88	2267892CB1	UCMCL5T01
89	2670307CB1	BRSTTUT01
90	4524210CB1	OVARTUT01
91	5584860CB1	UTRSNOT05
92	5807892CB1	KIDNFTUT01
93	3210044CB1	PITUNOT01
94	4942454CB1	BRAIFEN03

10

20

30

40

表6-1

ライブラリ	ベクター	ライブラリの説明
ADREN0T03	PSPORT1	このライブラリは、脳酪素欠乏症で死亡した年齢17歳の白人男性の副腎組織から単離されたRNAを用いて作製した。
BLADTUT04	pINCY	このライブラリは、60歳の白人男性から根治膀胱切除、前立腺切除、及び精管切除術の際に採取した膀胱腫瘍組織から単離したRNAを用いて作製した。病理学的には、グレード3の左膀胱壁移行性細胞癌であった。上皮内癌がドーム及び三角で見つかった。患者には喫煙の習慣があった。家族歴には、1型糖尿病、胃の悪性腫瘍、アテローム硬化性冠状動脈疾患、及び急性心筋梗塞があった。
BLADTUT05	pINCY	このライブラリは、66歳の白人男性の根治的前立腺切除、根治膀胱切除、及び尿路変更術の際に採取された膀胱腫瘍組織から単離したRNAを用いて作製した。病理学的には、膀胱前壁におけるグレード3の移行上皮癌を示していた。患者の病歴には、肺腫瘍及び緩解期におけるタバコ濫用が含まれていた。家族歴には、悪性の乳房新生物、結核、脳血管疾患、アテローム硬化性冠状動脈疾患、及び肺癌が含まれていた。
BRAIFEN03	pINCY	この標準化された胎児脳組織ライブラリは、胎児の脳ライブラリから独立した 3.26×10^6 のクローンから作製した。開始RNAは、妊娠23週目に左心形成不全により死産で生まれた白人の男子胎児から採取した脳組織から作製した。このライブラリは、ハイブリダイゼーションのアニーリング時間を1回当たり48時間と長くした点を除き Soaresら (PNAS (1994) 91:9228) 及び Bonaldo ら (Genome Research (1996) 6:791) の方法に従って2回標準化した。
BRAIFEN05	pINCY	この標準化された胎児脳組織ライブラリは、胎児の脳ライブラリから独立した 3.26×10^6 のクローンから作製した。開始RNAは、妊娠23週目に左心形成不全により死産で生まれた白人の男子胎児から採取した脳組織から作製した。このライブラリは、ハイブリダイゼーションのアニーリング時間を1回当たり48時間と長くした点を除き Soaresら (PNAS (1994) 91:9228) 及び Bonaldo ら (Genome Research (1996) 6:791) の方法に従って2回標準化した。
BRAIN0T14	pINCY	このライブラリは、年齢40歳の白人女性の脳腫瘍病巣の切除の際に左前葉から採取した脳組織から単離したRNAを用いて作製した。病理学的には、関連腫瘍組織は、グレード4大円形細胞性星状腫瘍を示していた。

【表 3 2】

10

20

30

表6-2

ライブラリ	ベクター	ライブラリの説明
BRA1TUT02	PSPORT1	このライブラリは、58歳の白人男性の脳腫瘍の病巣切除の際に前葉から採取した脳腫瘍組織から単離したRNAを用いて作製した。病理学的には、グレート2転移性副腎腫を示していた。患者の既往症には、グレート2腎細胞癌、不眠症、及び慢性気道閉塞があった。家族歴には、腎臓の悪性新生物があった。
BRA1TUT03	PSPORT1	このライブラリは、17歳の白人男性の脳腫瘍病巣の切除の際に左前葉から採取した脳腫瘍組織から単離したRNAを用いて作製した。病理学的には、グレート4の繊維状の巨細胞及び小細胞星状細胞腫を示していた。家族歴には、良性高血圧症及び脳血管疾患が含まれていた。
BRSTNOT02	PSPORT1	このライブラリは、55歳白人女性の片側拡大単純乳房切除の際に採取した病変乳房組織から単離したRNAを用いて作製した。病理学的には、アポクリン化生、硬化腺疾患、腺癌形成、及び異型を伴わない管過形成(ductal hyperplasia)によって特徴付けられる増殖性線維囊胞変化(fibrocytic changes)であった。病理学的には、関連腫瘍組織は、浸潤性グレート4の乳房腺癌であった。既往症には、心房性頻脈及び良性腫瘍があった。家族歴には、心血管疾患及び脳血管疾患があった。
BRSTTUT01	PSPORT1	このライブラリは、年齢55歳の白人女性から片側拡大単純乳房切断の際に採取された乳房腫瘍組織から単離したRNAを用いて作製した。病変は、左側乳房に広範に及んでいる小葉・管混合型の浸潤性グレート4乳房腺癌を示していた。腫瘍は、乳腺管近傍の真皮内深くにおいて認められ、嚢外に拡大していた。7個の中心位及び低位の腺高リンパ節及び5個の高位の腺高リンパ節は腫瘍陽性であった。増殖性の線維囊胞の変化は、アポクリン化生、硬化性腺疾患、囊胞形成、及び異型を伴わない乳腺管過形成によって特徴付けられていた。患者の既往症には、心房性頻脈、血便、及び良性乳房新生物があった。家族歴としては、良性高血圧症、アテローム硬化性冠動脈疾患、脳血管疾患、及び抑うつ障害があった。
COLNFET02	pINCY	このライブラリは、妊娠20週目で死亡したの白人女子胎児の結腸組織から単離したRNAを用いて作製した。
COLNNOT11	PSPORT1	このライブラリは、60歳の白人男性の左結腸半切除中に結腸組織を採取し、その結腸組織から単離したRNAを用いて作製した。
ENDANOT01	PBLUESCRIPT	このライブラリは、ある男性から心臓移植の際に採取した外植心臓から得た大動脈内皮細胞組織から単離したRNAを用いて作製した。

【表 3 3】

表6-3

ライブラリ	ベクター	ライブラリの説明
ENDCHON02	pINCY	この標準化した冠動脈内皮細胞組織ライブラリは、内皮組織ライブラリからの444,000の独立したクローンから作製した。開始RNAは、3歳の白人男児から採取した冠動脈内皮細胞組織から単離した。このライブラリは、Soaresら (PNAS (1994) 91:9228-9232) 及びBonaldoら (Genome Research (1996) 6:791-806) の条件を用いて2回標準化した。但し、アニーリングハイブリダイゼーション時間 (48時間/回) はかなり長くした。
EPIPNO05	pINCY	この標準化した前立腺上皮細胞組織ライブラリは、前立腺上皮細胞組織ライブラリからの2.36 × 10 ⁵ の独立したクローンから作製した。開始RNAは、17歳のヒスパニック系男性から採取した未処置の前立腺上皮細胞組織から単離した。このライブラリは、Soaresら (PNAS (1994) 91:9228-9232) 及びBonaldoら (Genome Research (1996) 6:791-806) の条件を用いて2回標準化した。但し、アニーリングハイブリダイゼーション時間 (48時間/回) はかなり長くした。
HMT2RAT01	PBLUESCRIPT	このライブラリは、ニューロン前駆細胞 (committed neuronal precursor) の特性を示すヒト奇形癌に由来する hMT2 細胞系より単離した RNA を用いて、Stratagene 社で作製された (STR937231)。細胞は24時間レチノイン酸で処理された。
ISLTNOT01	pINCY	このライブラリは、膵臓島細胞のプールされたコレクションから単離した RNA を用いて作製した。
KERANOT01	PBLUESCRIPT	このライブラリは、自発的に妊娠中絶された黒人男子胎児の足の皮膚から採取した新生児ケラチノサイトから単離した RNA を用いて作製された。
KIDNNOT05	PSPORT1	このライブラリは、脳無酸素症により生後2日で死亡したヒスパニック系の女児の腎臓組織から単離した RNA を用いて作製した。家族歴には、先天性の心疾患があった。
KIDNNOT09	pINCY	このライブラリは、妊娠23週で死亡した白人の男子胎児の腎臓組織より単離した RNA を用いて作製された。
KIDNTUT01	PSPORT1	このライブラリは、8ヶ月の女児から腎尿管切除術の際に採取した腎臓腫瘍組織から単離した RNA を用いて作製した。病理学的には、病変は90%の腎実質に及ぶウィルムス腫瘍 (腎芽細胞腫) を示していた。患者の既往症には、ヘパリン抗凝血療法を受けていた。

10

20

30

【表 3 4】

表6-4

ライブラリ	ベクター	ライブラリの説明
L1VRN008	pINCY	この標準化されたライブラリは、プールされている肝臓組織ライブラリからの独立したクローン5.7×10 ⁶ 個から作製された。開始RNAは、無酸素症で死亡した4歳のヒスパニック系男児及び妊娠16週目に無脳症で死亡した白人女子胎児から採取されプールされた肝臓組織から単離した。血清学的には、4歳児はサイトメガロウイルスに対して陽性であった。4歳児の患者の病歴には喘息が含まれていた。胎児の家族歴には、母親が産前に毎日ビタミンを摂取していたのと、僧帽弁脱出が含まれていた。このライブラリは、Soares 他 (PNAS (1994) 91:9228) 及びBonaldo 他 (Genome Research 6 (1996) :791) の条件に従って2回標準化した。但し、アーニングハイブリダイゼーションの時間は著しく長くした(48時間/回)。
LN00NOT03	pINCY	このライブラリは、67歳の白人男性の部分的肺切除及び気管支鏡検査の際に採取したリンパ節組織より単離したRNAを用いて作製した。鏡検では、この組織は広範囲に亘って壊死性であり、10%の生存可能な腫瘍を含んでいた。関連腫瘍組織は、病理学的には浸潤性グレード3-4の扁平上皮癌を示していた。患者の病歴には、血管腫があった。家族歴には、アテローム硬化性冠状動脈疾患、良性高血圧、及びうつ病性心不全があった。
LUNGNOT14	pINCY	このライブラリは、年齢47歳の白人男性の肺区域切除術の際に左下肺葉から採取された肺組織から単離されたRNAを用いて作製した。病理学的には、関連する腫瘍組織はグレード4の腺癌を示しており、肺実質は石灰化肉芽腫性炎を示していた。患者の既往症には、良性高血圧症及び慢性閉塞性肺疾患があった。家族歴には、II型糖尿病及び急性心筋梗塞があった。
LUNGNOT20	pINCY	このライブラリは、61歳白人男性の肺区域切除術の際、右上葉から採取した肺組織から単離したRNAを用いて作製した。病理学的には、右前上葉及び肺尖におけるブレブ、並びに右後上葉における肺気腫を示していた。患者の既往症には、狭心症及び胃潰瘍があった。家族歴には、硬膜下出血、未確認部位の癌、良性高血圧、アテローム硬化性冠状動脈疾患、及び肺炎が含まれていた。
LUNGNOT27	pINCY	このライブラリは、17歳のヒスパニック系女性から採取した肺組織から単離したRNAを用いて作製した。
LUNGNOT03	PSPORT1	このライブラリは、69歳の白人男性から肺区域切除の際に左下葉から採取した肺腫瘍組織から単離したRNAを用いて作製した。病理学的には、残留グレード3の浸潤性扁平上皮癌を示していた。患者の病歴には、急性心筋梗塞、前立腺肥大症、悪性皮膚癌、及び喫煙の習慣が含まれていた。

【表 3 5】

表6-5

ライブラリの説明		
ライブラリ	ベクター	
LUNGUT09	pINCY	このライブラリは、68歳の白人男性の肺切除の際に抽出された肺腫瘍組織より単離したRNAを用いて作製された。病理学的には、浸潤性グレード3の扁平上皮癌及び転移性腫瘍であった。患者の既往症には、II型糖尿病、甲状腺障害、抑うつ障害、高脂血症、食道潰瘍、及び喫煙があった。
MENITUT03	pINCY	このライブラリは、35歳白人女性の脳腫瘍の病変切除の際に採取した脳腫瘍組織から単離したRNAを用いて作製した。病理学的には、脳の右小脳橋角の良性新生物であった。患者の病歴には、甲状腺機能低下症があった。家族歴には、心筋梗塞及び乳癌があった。
OVARNOT03	PSPORT1	このライブラリは、43歳の白人女性から卵管及び卵巣摘出の際に採取した卵巣腫瘍組織から単離したRNAを用いて作製した。関連腫瘍組織は、病理学的にはグレード2の粘液性嚢胞腺癌を示していた。患者の病歴には、僧帽弁疾患、肺炎、及びウイルス性肝炎があった。家族歴には、アテローム硬化性冠状動脈疾患、痔瘻、ストレス反応、脳血管疾患、乳癌、及び子宮癌が含まれる。
OVARTUT01	PSPORT1	このライブラリは、43歳の白人女性から卵管及び卵巣の摘出の際に採取した卵巣腫瘍組織から単離したRNAを用いて作製した。病理学的には、左卵巣全体に關与するグレード2粘液性嚢胞腺癌を示していた。患者の病歴には、僧帽弁疾患、肺炎、及びウイルス性肝炎が含まれる。家族歴には、アテローム性冠状動脈疾患、痔瘻、ストレス反応、脳血管疾患、乳癌、及び子宮癌が含まれる。
PANCNOT07	pINCY	このライブラリは、妊娠23週目で死亡した白人男子胎児の脾臓組織から単離したRNAを用いて作製した。
PANCNOT08	pINCY	このライブラリは、65歳の白人女性の根治亜全脾切除の際に採取した脾臓組織から単離したRNAを用いて作製した。病理学的にはこの関連腫瘍組織は、浸潤性グレード2の腺癌を示していた。患者の病歴には、II型糖尿病、変形性関節症、心血管疾患、大腸の良性心生物、及び白内障があった。
PITUNOT01	PBLUESCRIPT	このライブラリは、外傷で死亡した年齢16〜70歳の18人の白人男女のドナーから採取した正常な下垂体から単離したRNAを用いて作製した。RNA (CLON 6584-2, lot 35278) は Clontech 社から得たものである。

【表 3 6】

10

20

30

表6-6

ライブラリ	ベクター	ライブラリの説明
PROSNOT16	pINCY	このライブラリは、68歳の白人男性から根治前立腺切除の際に採取した病変前立腺組織から単離したRNAを用いて作製した。病理学的には、腺癌（Gleasonグレード3+4）を示していた。関連する腫瘍組織は病理学的には、腺癌（Gleasonグレード3+4）を示していた。患者は高値の前立腺特異抗原（PSA）を呈し、重症筋無力症と診断された。患者の病歴には、変形性関節症及びII型糖尿病が含まれていた。家族歴には、良性高血圧、急性心筋梗塞、高脂血症、及びアテローム硬化性冠状動脈疾患が含まれていた。
SINTFET03	pINCY	このライブラリは、妊娠20週目で死亡した白人女子胎児から採取した小腸組織から単離したRNAを用いて作製した。
SPLNFET01	PBLUESCRIPT	このライブラリは、胎児脾臓組織のプールから単離したRNAを用いて作製した。ベクターをパッケージングした後、長期保存のためライブラリを安定化させるべく、 2×10^6 個の一次クローンを増幅した。増幅により配列存在量が著しく歪み得る。
SPLNNOT04	pINCY	このライブラリは、大脳無酸素症で死亡した2歳のヒスパニック系男児の脾臓組織から単離したRNAを用いて作製した。過去の医療履歴及び血清学的には問題がなかった。
THP1NOT03	pINCY	このライブラリは、未処理の THP-1 細胞から単離したホリ RNA を用いて作製した。THP-1(ATCC TIB 202)は、急性単球性白血病の1歳の白人男児の末梢血に由来するヒト前単球細胞系である。
TYMNOT02	PBLUESCRIPT	このライブラリは、非粘着性末梢血の単核細胞から単離したRNAを用いて作製した。この末梢血は、血縁関係のない男女ドナーに由来する。各ドナーからの細胞をフィコール-ハイバーク法により精製し、遠心分離器にかけ、GusCNを含むバッファーで溶解し、塩化セシウムで吸収してRNAを得、ライブラリの作製に用いた。
UCMCL5T01	PBLUESCRIPT	このライブラリは、12名の臍帯血から得られた単核細胞から単離されたRNAを用いて作製した。RNAをプールされたライセートから単離する前に細胞を12日間 IL-5 と共に培養した。
UTRSNOT05	pINCY	このライブラリは、45歳の白人女性から腹式子宮全摘出及び結腸全切除の際に採取した子宮組織から単離したRNAを用いて作製した。関連する腫瘍組織は、病理学的には子宮筋層の複数の平滑筋腫及び盲腸のグレード2の盲腸の大腸腺癌を示していた。患者の病歴には、多発性硬化病及び僧帽弁障害が含まれていた。家族歴には、I型糖尿病、脳血管疾患、アテローム硬化性冠状動脈疾患、悪性皮膚腫新生物、高血圧症、及び結腸の悪性新生物が含まれていた。

10

20

30

【表 37】

表7-1

プログラム名	説明	引用文献	パラメーター閾値
ABI FACTURA	核酸配列においてベクター配列を除去して不 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 列の比較及び注釈付け (annotation) に有用 である。	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler BLAST	核酸配列を構築するプログラム。 Basic Local Alignment Search Tool は、ア ミノ酸及び核酸配列の配列類似性検索に有用 であり、blastp 及び blastn、blastx、tblastn、 tblastx の5つのファンクションがある。	Applied Biosystems, Foster City, CA. 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	Pearson 及び Lipman アルゴリズムは、同合わ せの配列と同種の配列群との類似性を検索す る。FASTA は、fasta 及び tfasta、fastx、 tfastx、ssearch の少なくとも5つのファン クションを含む。	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E 値=1.06E-6 構築された ESTs: fasta 同一性 =95%以上、一致長さ=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	BLOCKS IMPROVED Searcher は、BLOCKS 及び PRINTS、DOMO、PRODOM、PFAM データベースに おける配列に対して、ある配列の一致性を調 べ、遺伝子ファミリー及び配列相同性、構造 的フィンガープリント領域を探索する。	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.C. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; Artwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	確率値=1.0E-3 以下
HMMER	PFAM などのタンパク質ファミリーコンセン サス配列の隠れマルコフモデル(HMM)に基づ いたデータベースに対して問合せ配列を検索 するアルゴリズム。	Krogh, A. 他 (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. 他 (1998) Nucleic Acids Res. 26:320-322. Durbin, R. 他 (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	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

表7-2

プログラム名	説明	引用文献	パラメーター関連値
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における構造及び配列のモチーフを検索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4:61-66; Gribskov, 他 (1989) Methods Enzymol. 183:146-159; Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221	標準化された質のスコアを特定 の Prosite モチーフに対する GCG 指定 "HIGH" 値 通常、スコア=1.4-2.1
Phred	高い感度及び確率で自動配列決定機のトレー スを調べる塩基読出しアルゴリズムである。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的なイ ンプリメンテーションに基づく SWAT や CrossMatch を含む Phils Revised Assembly プログラムであって、配列相同性の検索及び DNA 配列の構築に有用である。	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; Green, P., University of Washington, Seattle, WA.	スコア=120 以上 一致長さ=56 以上
Consed	Phrap で構築したものの表示及び編集をする ためのグラフィックツールである。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキャンして分泌シグナル ペプチドの存在を調べる重み付けマトリクス 解析プログラムである。	Nielson, H. 他 (1997) Protein Engineering 10: 1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	スコア=3.5 以上
TMAP	重み付けマトリクスを用いたタンパク質配 列上の膜貫通セグメントの明確化および向き を決定するためのプログラム	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	隠れマルコフモデル (HMM) を用いたタンパク 質配列上の膜貫通セグメントの明確化および 向きを決定するためのプログラム。	Sonnhammer, E.L. 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, pp. 175-182.	
Motifs	Prosite で定義された配列と一致したパター ンについてアミノ酸配列を検索するプログラ ムである。	Bairoch 他、(1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

10

20

30

【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
8 November 2001 (08.11.2001)

PCT

(10) International Publication Number
WO 01/83524 A2

(51) International Patent Classification: C07K 14/00

(21) International Application Number: PCT/US01/13862

(22) International Filing Date: 27 April 2001 (27.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/200,184 28 April 2000 (28.04.2000) US
60/201,875 4 May 2000 (04.05.2000) US
60/202,090 4 May 2000 (04.05.2000) US
60/210,232 6 June 2000 (06.06.2000) US
60/220,553 25 July 2000 (25.07.2000) US

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

(72) Inventors; and

(75) Inventors/Applicants (for US only): LAI, Preent [US/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). YU, Henry [US/US]; 826 Luis Avenue, Sunnyvale, CA 94087 (US). TANG, Y. Tom [US/US]; 4230 Ramnick Court, San Jose, CA 95118 (US). LU, Danyang, Anna, M. [US/US]; 233 Clay Drive, San Jose, CA 95121 (US). AZIMZAI, Yalda [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US). AD-YOUNG, Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). HILLMAN, Jennifer, L. [US/US]; 230 Monroe Drive #17, Mountain View, CA 94040 (US). BAUGHN,

Marish, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). YAO, Monique, G. [US/US]; 111 Pinerick Court, Mountain View, CA 94043 (US). BURFORD, Neil [US/US]; 105 Wildwood Circle, Durham, CT 06422 (US). RATHA, Rajeev [US/US]; 555 E Camino Real #709, Sunnyvale, CA 94087 (US). POLICKY, Jennifer, J. [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US).

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

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

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

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

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

WO 01/83524 A2

(54) Title: RNA METABOLISM PROTEINS

(57) Abstract: The invention provides human RNA metabolism proteins (RMEP) and polynucleotides which identify and encode RMEP. 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 RMEP.

WO 01/83524

PCT/US01/13862

RNA METABOLISM PROTEINS

TECHNICAL FIELD

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

10 BACKGROUND OF THE INVENTION

Ribonucleic acid (RNA) is a linear single-stranded polymer of four nucleotides, ATP, CTP, UTP, and GTP. In most organisms, RNA is transcribed as a copy of deoxyribonucleic acid (DNA), the genetic material of the organism. In retroviruses RNA rather than DNA serves as the genetic material. RNA copies of the genetic material encode proteins or serve various structural, catalytic, or regulatory roles in organisms. RNA is classified according to its cellular localization and function. Messenger RNAs (mRNAs) encode polypeptides. Ribosomal RNAs (rRNAs) are assembled, along with ribosomal proteins, into ribosomes, which are cytoplasmic particles that translate mRNA into polypeptides. Transfer RNAs (tRNAs) are cytosolic adaptor molecules that function in mRNA translation by recognizing both an mRNA codon and the amino acid that matches that codon. Heterogeneous nuclear RNAs (hnRNAs) include mRNA precursors and other nuclear RNAs of various sizes. Small nuclear RNAs (snRNAs) are a part of the nuclear spliceosome complex that removes intervening, non-coding sequences (introns) and rejoins exons in pre-mRNAs.

Proteins are associated with RNA during its transcription from DNA, RNA processing, and translation of mRNA into protein. Proteins are also associated with RNA as it is used for structural, catalytic, and regulatory purposes.

RNA Processing

Various proteins are necessary for processing of transcribed RNAs in the nucleus. Pre-mRNA processing steps include capping at the 5' end with methylguanosine, polyadenylating the 3' end, and splicing to remove introns. The spliceosomal complex is comprised of five small nuclear ribonucleoprotein particles (snRNPs) designated U1, U2, U4, U5, and U6. Each snRNP contains a single species of snRNA and about ten proteins. The RNA components of some snRNPs recognize and base-pair with intron consensus sequences. The protein components mediate spliceosome assembly and the splicing reaction.

35 An early step in pre-mRNA cleavage involves the cleavage factor Im (CF Im). The human CF

WO 01/83524

PCT/US01/13862

- In protein aids in the recruitment and assembly of processing factors that make up the 3' end processing complex (Rucsegger, U. et al (1998) Mol. Cell. 1:243-253). The murine formin binding proteins (FBP's) FBP11 and FBP12 are components of pre-mRNA splicing complexes that facilitate the bridging of 5' and 3' ends of the intron. These proteins function through bridging interactions involving U1 and U2 snRNPs. Autoantibodies to snRNP proteins are found in the blood of patients with systemic lupus erythematosus (Stryer, L. (1995) Biochemistry W.H. Freeman and Company, New York NY, p. 863).
- Heterogeneous nuclear ribonucleoproteins (hnRNPs) have been identified that have roles in splicing, exporting of the mature RNAs to the cytoplasm, and mRNA translation (Biamonti, G. et al. (1998) Clin. Exp. Rheumatol. 16:317-326). Some examples of hnRNPs include the yeast proteins Hrp1p, involved in cleavage and polyadenylation at the 3' end of the RNA; Cbp80p, involved in capping the 5' end of the RNA; and Npl3p, a homolog of mammalian hnRNP A1, involved in export of mRNA from the nucleus (Shen, E.C. et al. (1998) Genes Dev. 12:679-691). HnRNPs have been shown to be important targets of the autoimmune response in rheumatic diseases (Biamonti, *supra*).
- Many snRNP and hnRNP proteins are characterized by an RNA recognition motif (RRM). (Reviewed in Birney, E. et al. (1993) Nucleic Acids Res. 21:5803-5816.) The RRM is about 80 amino acids in length and forms four β -strands and two α -helices arranged in an α/β sandwich. The RRM contains a core RNP-1 octapeptide motif along with surrounding conserved sequences. In addition to snRNP proteins, examples of RNA-binding proteins which contain the above motifs include heteronuclear ribonucleoproteins which stabilize nascent RNA and factors which regulate alternative splicing. Alternative splicing factors include developmentally regulated proteins, specific examples of which have been identified in lower eukaryotes such as *Drosophila melanogaster* and *Caenorhabditis elegans*. These proteins play key roles in developmental processes such as pattern formation and sex determination, respectively. (See, for example, Hodgkin, J. et al. (1994) Development 120:3681-3689.)
- RNA Stability and Degradation
- RNA helicases alter and regulate RNA conformation and secondary structure by using energy derived from ATP hydrolysis to destabilize and unwind RNA duplexes. The most well-characterized and ubiquitous family of RNA helicases is the DEAD-box family, so named for the conserved B-type ATP-binding motif which is diagnostic of proteins in this family. Over 40 DEAD-box helicases have been identified in organisms as diverse as bacteria, insects, yeast, amphibians, mammals, and plants. DEAD-box helicases function in diverse processes such as translation initiation, splicing, ribosome assembly, and RNA editing, transport, and stability. Some DEAD-box helicases play tissue- and stage-specific roles in spermatogenesis and embryogenesis. All DEAD-box helicases contain several

WO 01/83524

PCT/US01/13862

conserved sequence motifs spread out over about 420 amino acids. These motifs include an A-type ATP binding motif, the DEAD-box/B-type ATP-binding motif, a serine/arginine/lysine tripeptide of unknown function, and a C-terminal glycine-rich motif with a possible role in substrate binding and unwinding. In addition, alignment of divergent DEAD-box helicase sequences has shown that 37 amino acid residues are identical among these sequences, suggesting that conservation of these residues is important for helicase function. (Reviewed in Linder, P. et al. (1989) *Nature* 337:121-122.)

Overexpression of the DEAD-box 1 protein (DDX1) may play a role in the progression of neuroblastoma (Nb) and retinoblastoma (Rb) tumors. These observations suggest that DDX1 may promote or enhance tumor progression by altering the normal secondary structure and expression levels of RNA in cancer cells. Other DEAD-box helicases have been implicated either directly or indirectly in ultraviolet light-induced tumors, B-cell lymphoma, and myeloid malignancies. (Reviewed in Godbout, R. et al. (1998) *J. Biol. Chem.* 273:21161-21168.)

Ribonucleases (RNases) catalyze the hydrolysis of phosphodiester bonds in RNA chains, thus cleaving the RNA. For example, RNase P is a ribonucleoprotein enzyme which cleaves the 5' end of pre-tRNAs as part of their maturation process. RNase H digests the RNA strand of an RNA/DNA hybrid. Such hybrids occur in cells invaded by retroviruses, and RNase H is an important enzyme in the retroviral replication cycle. RNase H domains are often found as a domain associated with reverse transcriptases. RNase activity in serum and cell extracts is elevated in a variety of cancers and infectious diseases (Schein, C.H. (1997) *Nat. Biotechnol.* 15:529-536). Regulation of RNase activity is being investigated as a means to control tumor angiogenesis, allergic reactions, viral infection and replication, and fungal infections.

Degradation of mRNAs having premature termination or nonsense codons is accomplished through a surveillance mechanism that has been termed nonsense-mediated mRNA decay (NMD). This mechanism helps eliminate flawed mRNAs that might code for nonfunctional or deleterious polypeptides. Various NMD components are linked to both yeast and human RNA metabolism disorders (Hentze, M. and Kulozik, A. (1999) *Cell* 96:307-310).

TRANSLATION

Ribosomal RNAs (rRNAs) are assembled, along with ribosomal proteins, into ribosomes, which are cytoplasmic particles that translate messenger RNA (mRNA) into polypeptides. The eukaryotic ribosome is composed of a 60S (large) subunit and a 40S (small) subunit, which together form the 80S ribosome. In addition to the 18S, 28S, 5S, and 5.8S rRNAs, ribosomes contain from 50 to over 80 different ribosomal proteins, depending on the organism. Ribosomal proteins are classified according to which subunit they belong (i.e., L, if associated with the large 60S large subunit or S if associated with the small 40S subunit). *E. coli* ribosomes have been the most thoroughly studied and

WO 01/83524

PCT/US01/13862

contain 50 proteins, many of which are conserved in all life forms. The structures of nine ribosomal proteins have been solved to less than 3.0 Å resolution (i.e., S5, S6, S17, L1, L6, L9, L12, L14, L30), revealing common motifs, such as β - α - β protein folds in addition to acidic and basic RNA-binding motifs positioned between β -strands. Most ribosomal proteins are believed to contact rRNA directly (reviewed in Liljas, A. and Garber, M. (1995) *Curr. Opin. Struct. Biol.* 5:721-727, see also Woodson, S.A. and Leontis, N.B. (1998) *Curr. Opin. Struct. Biol.* 8:294-300; Ramakrishnan, V. and White, S.W. (1998) *Trends Biochem. Sci.* 23:208-212.)

Ribosomal proteins may undergo post-translational modifications or interact with other ribosome-associated proteins to regulate translation. For example, the highly homologous 40S ribosomal protein S6 kinases (S6K1 and S6K2) play a key role in the regulation of cell growth by controlling the biosynthesis of translational components which make up the protein synthetic apparatus (including the ribosomal proteins). In the case of S6K1, at least eight phosphorylation sites are believed to mediate kinase activation in a hierarchical fashion (Dufner and Thomas (1999) *Exp. Cell. Res.* 253:100-109). Some of the ribosomal proteins, including L1, also function as translational repressors by binding to polycistronic mRNAs encoding ribosomal proteins (reviewed in Liljas, A. *supra* and Garber, M. *supra*).

Recent evidence suggests that a number of ribosomal proteins have secondary functions independent of their involvement in protein biosynthesis. These proteins function as regulators of cell proliferation and, in some instances, as inducers of cell death. For example, the expression of human ribosomal protein L13a has been shown to induce apoptosis by arresting cell growth in the G2/M phase of the cell cycle. Inhibition of expression of L13a induces apoptosis in target cells, which suggests that this protein is necessary, in the appropriate amount, for cell survival. Similar results have been obtained in yeast where inactivation of yeast homologues of L13a, rp22 and rp23, results in severe growth retardation and death. A closely related ribosomal protein, L7, arrests cells in G1 and also induces apoptosis. Thus, it appears that a subset of ribosomal proteins may function as cell cycle checkpoints and compose a new family of cell proliferation regulators.

Mapping of individual ribosomal proteins on the surface of intact ribosomes is accomplished using 3D immunocryoelectron microscopy, whereby antibodies raised against specific ribosomal proteins are visualized. Progress has been made toward the mapping of L1, L7, and L12 while the structure of the intact ribosome has been solved to only 20-25 Å resolution and inconsistencies exist among different crude structures (Frank, J. (1997) *Curr. Opin. Struct. Biol.* 7:266-272).

Three distinct sites have been identified on the ribosome. The aminoacyl-tRNA acceptor site (A site) receives charged tRNAs (with the exception of the initiator-tRNA). The peptidyl-tRNA site (P site) binds the nascent polypeptide as the amino acid from the A site is added to the elongating chain. Deacylated tRNAs bind in the exit site (E site) prior to their release from the ribosome. The

WO 01/83524

PCT/US01/13862

structure of the ribosome is reviewed in Stryer, L. (1995) *Biochemistry* W.H. Freeman and Company, New York NY pp. 888-908; Lodish, H. et al. (1995) *Molecular Cell Biology* Scientific American Books, New York NY pp. 119-138; and Lewin, B. (1997) *Genes VI* Oxford University Press, Inc. New York, NY).

5

tRNA Charging

Correct translation of the genetic code depends upon each amino acid forming a linkage with the appropriate transfer RNA (tRNA). The aminoacyl-tRNA synthetases (aaRSs) are essential proteins found in all living organisms. The aaRSs are responsible for the activation and correct attachment of an amino acid with its cognate tRNA, as the first step in protein biosynthesis. Prokaryotic organisms have at least twenty different types of aaRSs, one for each different amino acid, while eukaryotes usually have two aaRSs, a cytosolic form and a mitochondrial form, for each different amino acid. The 20 aaRS enzymes can be divided into two structural classes. Class I enzymes add amino acids to the 2' hydroxyl at the 3' end of tRNAs while Class II enzymes add amino acids to the 3' hydroxyl at the 3' end of tRNAs. Each class is characterized by a distinctive topology of the catalytic domain. Class I enzymes contain a catalytic domain based on the nucleotide-binding 'Rossmann fold'. In particular, a consensus tetrapeptide motif is highly conserved (Prosite Document PDOC00161, Aminoacyl-transfer RNA synthetases class-I signature). Class II enzymes contain a central catalytic domain, which consists of a seven-stranded antiparallel β -sheet domain, as well as N- and C-terminal regulatory domains. Class II enzymes are separated into two groups based on the heterodimeric or homodimeric structure of the enzyme; the latter group is further subdivided by the structure of the N- and C-terminal regulatory domains (Hartlein, M. and Cusack, S. (1995) *J. Mol. Evol.* 40:519-530). The different aaRSs are believed to be the result of divergent evolution, likely following gene duplication events. Notably, amino acids such as Gln, were among the last to appear in nature and evolutionary studies suggest that Gln-RSs appeared first in eukaryotes and were later horizontally transferred to prokaryotes (Lamour, V. et al. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 91:8670-74 and Siatecka, M. et al. (1998) *Eur. J. Biochem.* 256:80-7). The importance of Gln-RS and Gln-tRNA^{Gln} are discussed below.

In addition to their function in protein synthesis, specific aminoacyl tRNA synthetases also play roles in cellular fidelity, RNA splicing, RNA trafficking, apoptosis, and transcriptional and translational regulation. For example, human tyrosyl-tRNA synthetase can be proteolytically cleaved into two fragments with distinct cytokine activities. The carboxy-terminal domain exhibits monocyte and leukocyte chemotaxis activity as well as stimulating production of myeloperoxidase, tumor necrosis factor- α , and tissue factor. The N-terminal domain binds to the interleukin-8 type A receptor and functions as an interleukin-8-like cytokine. Human tyrosyl-tRNA synthetase is secreted from

WO 01/83524

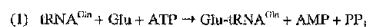
PCT/US01/13862

apoptotic tumor cells and may accelerate apoptosis (Wakasugi, K. and Schimmel, P. (1999) *Science* 284:147-151). Mitochondrial *Neurospora crassa* TyrRS and *S. cerevisiae* LeuRS are essential factors for certain group I intron splicing activities, and human mitochondrial LeuRS can substitute for the yeast LeuRS in a yeast null strain. Certain bacterial aaRSs are involved in regulating their own transcription or translation (Martinis, *supra*). Several aaRSs are able to synthesize diadenosine oligophosphates, a class of signalling molecules with roles in cell proliferation, differentiation, and apoptosis (Kisselev, L.L. et al. (1998) *FEBS Lett.* 427:157-163; Vartanian, A. et al. (1999) *FEBS Lett.* 456:175-180).

Under optimal conditions, polypeptide synthesis proceeds at a rate of approximately 40 amino acid residues per second. The rate of misincorporation during translation is on the order of 10^{-4} and is primarily the result of aminoacyl-t-RNAs being charged with the incorrect amino acid. Incorrectly charged tRNA are toxic to cells as they result in the incorporation of incorrect amino acid residues into an elongating polypeptide. The rate of translation is presumed to be a compromise between the optimal rate of elongation and the need for translational fidelity. Mathematical calculations predict that 10^{-4} is indeed the maximum acceptable error rate for protein synthesis in a biological system (reviewed in Stryer, L. *supra* and Watson, J. et al. (1987) The Benjamin/Cummings Publishing Co., Inc. Menlo Park, CA). A particularly error prone aminoacyl-tRNA charging event, the charging of tRNA^{Gln} with Gln. A mechanism exist for the correction of this mischarging event which likely has its origins in evolution. Gln was among the last of the 20 naturally occurring amino acids used polypeptide synthesis to appear in nature. Gram positive bacteria, cyanobacteria, Archaea, and eukaryotic organelles possess a noncanonical pathway for the synthesis of Gln-tRNA^{Gln} based on the transformation of Glu-tRNA^{Gln} (synthesized by Glu-tRNA synthetase, GluRS) using the enzyme Glu-tRNA^{Gln} amidotransferase (Glu-AdT). The reactions involved in the transamidation pathway are as follows (Cumow, A.W. et al. (1997) *Nucleic Acids Symposium* 36:2-4):

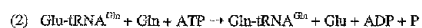
25

GluRS



Glu-AdT

30



A similar enzyme, Asp-tRNA^{Asn} amidotransferase, exists in Archaea, which transforms Asp-tRNA^{Asn} to Asn-tRNA^{Asn}. Formylase, the enzyme that transforms Met-tRNA^{Met} to fMet-tRNA^{Met} in eubacteria, is likely to be a related enzyme. A hydrolytic activity has also been identified that destroys mischarged Val-tRNA^{Val} (Schimmel, P. et al. (1998) *FASEB J.* 12:1599-1609). I likely scenario for

35

WO 01/83524

PCT/US01/13862

- the evolution of Glu-AdT in primitive life forms is the absence a specific glutamyl-tRNA synthetase (GluRS), requiring an alternative pathway for the synthesis of Gln-tRNA^{Gln}. In fact, deletion of the Glu-AdT operon in Gram positive bacteria is lethal (Cumow, A.W. et al. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94:11819-11826). The existence of GluRS activity in other organisms has been inferred by the
- 5 high degree of conservation in translation machinery in nature; however, GluRS has not been identified in all organisms, including *Homo sapiens*. Such an enzyme would be responsible for ensuring translational fidelity and reducing the synthesis of defective polypeptides.
- Autoantibodies against aminoacyl-tRNAs are generated by patients with autoimmune diseases such as rheumatic arthritis, dermatomyositis and polymyositis, and correlate strongly with
- 10 complicating interstitial lung disease (ILD) (Freist, W. et al. (1999) *Biol. Chem.* 380:623-646; Freist, W. et al. (1996) *Biol. Chem. Hoppe Seyler* 377:343-356). These antibodies appear to be generated in response to viral infection, and coxsackie virus has been used to induce experimental viral myositis in animals.
- Comparison of aaRS structures between humans and pathogens has been useful in the design
- 15 of novel antibiotics (Schimmel, *supra*). Genetically engineered aaRSs have been utilized to allow site-specific incorporation of unnatural amino acids into proteins *in vivo* (Liu, D.R. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:10092-10097).

Translation Initiation

- 20 Initiation of translation can be divided into three stages. The first stage brings an initiator transfer RNA (Met-tRNA_i) together with the 40S ribosomal subunit to form the 43S preinitiation complex. The second stage binds the 43S preinitiation complex to the mRNA, followed by migration of the complex to the correct AUG initiation codon. The third stage brings the 60S ribosomal subunit to the 40S subunit to generate an 80S ribosome at the initiation codon. Regulation of translation
- 25 primarily involves the first and second stage in the initiation process (V.M. Pain (1996) *Eur. J. Biochem.* 236:747-771).
- Several initiation factors, many of which contain multiple subunits, are involved in bringing an initiator tRNA and 40S ribosomal subunit together. One eukaryotic initiation factor (eIF) eIF5A is an 18-kD protein containing the unique amino acid residue, hypusine (N epsilon-(4-amino-2-
- 30 hydroxybutyl)lysine) (Rinaldo, M. et al. (1993) *Gene* 137:303-307). eIF2, a guanine nucleotide binding protein, recruits the initiator tRNA to the 40S ribosomal subunit. Only when eIF2 is bound to GTP does it associate with the initiator tRNA. eIF2B, a guanine nucleotide exchange protein, is responsible for converting eIF2 from the GDP-bound inactive form to the GTP-bound active form. Two other factors, eIF1A and eIF3 bind and stabilize the 40S subunit by interacting with 18S
- 35 ribosomal RNA and specific ribosomal structural proteins. eIF3 is also involved in association of the

WO 01/83524

PCT/US01/13862

40S ribosomal subunit with mRNA. The Met-tRNA_i, eIF1A, eIF3, and 40S ribosomal subunit together make up the 43S preinitiation complex (Pain, *supra*).

Additional factors are required for binding of the 43S preinitiation complex to an mRNA molecule, and the process is regulated at several levels. eIF4F is a complex consisting of three proteins: eIF4E, eIF4A, and eIF4G. eIF4E recognizes and binds to the mRNA 5'-terminal m⁷GTP cap, eIF4A is a bidirectional RNA-dependent helicase, and eIF4G is a scaffolding polypeptide. eIF4G has three binding domains. The N-terminal third of eIF4G interacts with eIF4E, the central third interacts with eIF4A, and the C-terminal third interacts with eIF3 bound to the 43S preinitiation complex. Thus, eIF4G acts as a bridge between the 40S ribosomal subunit and the mRNA (M.W. Hentze (1997) *Science* 275:500-501).

The ability of eIF4F to initiate binding of the 43S preinitiation complex is regulated by structural features of the mRNA. The mRNA molecule has an untranslated region (UTR) between the 5' cap and the AUG start codon. In some mRNAs this region forms secondary structures that impede binding of the 43S preinitiation complex. The helicase activity of eIF4A is thought to function in removing this secondary structure to facilitate binding of the 43S preinitiation complex (Pain, *supra*).

Translation Elongation

Elongation is the process whereby additional amino acids are joined to the initiator methionine to form the complete polypeptide chain. The elongation factors EF1 α , EF1 β γ , and EF2 are involved in elongating the polypeptide chain following initiation. EF1 α is a GTP-binding protein. In EF1 α 's GTP-bound form, it brings an aminoacyl-tRNA to the ribosome's A site. The amino acid attached to the newly arrived aminoacyl-tRNA forms a peptide bond with the initiator methionine. The GTP on EF1 α is hydrolyzed to GDP, and EF1 α -GDP dissociates from the ribosome. EF1 β γ binds EF1 α -GDP and induces the dissociation of GDP from EF1 α , allowing EF1 α to bind GTP and a new cycle to begin.

As subsequent aminoacyl-tRNAs are brought to the ribosome, EF-G, another GTP-binding protein, catalyzes the translocation of tRNAs from the A site to the P site and finally to the E site of the ribosome. This allows the processivity of translation.

Translation Termination

The release factor eRF carries out termination of translation. eRF recognizes stop codons in the mRNA, leading to the release of the polypeptide chain from the ribosome.

The discovery of new RNA metabolism 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 nervous system, autoimmune/inflammatory, cell proliferative, and developmental disorders,

WO 01/83524

PCT/US01/13862

and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of RNA metabolism proteins.
nervous system disorders, autoimmune/inflammatory disorders, and cell proliferative disorders including cancer

5

SUMMARY OF THE INVENTION

The invention features purified polypeptides, RNA metabolism proteins, referred to collectively as "RMEP" and individually as "RMEP-1," "RMEP-2," "RMEP-3," "RMEP-4," "RMEP-5," "RMEP-6," "RMEP-7," "RMEP-8," "RMEP-9," "RMEP-10," "RMEP-11," "RMEP-12," "RMEP-13," "RMEP-14," "RMEP-15," "RMEP-16," "RMEP-17," "RMEP-18," "RMEP-19," "RMEP-20,"
10 "RMEP-21," "RMEP-22," "RMEP-23," "RMEP-24," "RMEP-25," "RMEP-26," "RMEP-27," "RMEP-28," "RMEP-29," "RMEP-30," "RMEP-31," "RMEP-32," "RMEP-33," "RMEP-34," "RMEP-35," "RMEP-36," "RMEP-37," "RMEP-38," "RMEP-39," "RMEP-40," "RMEP-41," "RMEP-42," "RMEP-43," "RMEP-44," "RMEP-45," "RMEP-46," and "RMEP-47." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide
15 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the
20 group consisting of SEQ ID NO:1-47. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-47.

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:1-47, b) a naturally occurring polypeptide comprising an amino acid sequence
25 at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-47. In
30 another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:48-94.

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:1-47, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical

WO 01/83524

PCT/US01/13862

to an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47. In one alternative, the invention provides a cell
 5 transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90%
 10 identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a
 15 recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, b) a naturally occurring polypeptide comprising
 20 an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47.

The invention further provides an isolated polynucleotide selected from the group consisting of
 25 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:48-94, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:48-94, 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
 30 comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:48-94, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90%

WO 01/83524

PCT/US01/13862

identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:48-94, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence
 5 complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

10 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:48-94, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:48-94, c) a
 15 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.

20 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:1-47, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, c) a biologically active fragment of a polypeptide having an amino acid sequence
 25 selected from the group consisting of SEQ ID NO:1-47, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-47. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional RMEP,
 30 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:1-47, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the

WO 01/83524

PCT/US01/13862

group consisting of SEQ ID NO:1-47, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional RMEP, 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:1-47, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional RMEP, 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:1-47, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47. 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

WO 01/83524

PCT/US01/13862

acid sequence selected from the group consisting of SEQ ID NO:1-47, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, and
 5 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of
 10 the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:48-94, the method comprising a) exposing a sample
 15 comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20
 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:48-94, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:48-94, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the
 25 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:48-94, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90%
 30 identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:48-94, 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
 35 complex in the treated biological sample with the amount of hybridization complex in an untreated

WO 01/83524

PCT/US01/13862

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

- 5 Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.
- Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability score for the match between each polypeptide and its GenBank homolog is also shown.
- 10 Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.
- Table 4 lists the cDNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.
- 15 Table 5 shows the representative cDNA library for polynucleotides of the invention.
- Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.
- Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.
- 20

DESCRIPTION OF THE INVENTION

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

25

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

30

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

WO 01/83524

PCT/US01/13862

practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is

5 not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

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

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

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

20 a given sequence.

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

WO 01/83524

PCT/US01/13862

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 protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

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

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

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

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

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

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

WO 01/83524

PCT/US01/13862

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

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

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

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

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

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

WO 01/83524

PCT/US01/13862

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

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

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

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

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

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

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

WO 01/83524

PCT/US01/13862

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

A fragment of SEQ ID NO:48-94 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:48-94, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:48-94 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:48-94 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:48-94 and the region of SEQ ID NO:48-94 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-47 is encoded by a fragment of SEQ ID NO:48-94. A fragment of SEQ ID NO:1-47 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-47. For example, a fragment of SEQ ID NO:1-47 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-47. The precise length of a fragment of SEQ ID NO:1-47 and the region of SEQ ID NO:1-47 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

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

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

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

Percent identity between polynucleotide sequences may be determined using the default

WO 01/83524

PCT/US01/13862

parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191.

- 5 For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

- Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/got/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

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

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

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

WO 01/83524

PCT/US01/13862

shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

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

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

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

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

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

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

WO 01/83524

PCT/US01/13862

describe a length over which percentage identity may be measured.

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

- 5 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.

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

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

WO 01/83524

PCT/US01/13862

such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

5 The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C₀t or R₀t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or
10 their nucleic acids have been fixed).

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

"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
15 various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

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

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

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

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

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
30 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

WO 01/83524

PCT/US01/13862

sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

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

"Post-translational modification" of an RMEP 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 RMEP.

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

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

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

Oligonucleotides for use as primers are selected using software known in the art for such

WO 01/83524

PCT/US01/13862

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. 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).

WO 01/83524

PCT/US01/13862

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,

5 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
10 of deoxyribose.

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

15 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
20 A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

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

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

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

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

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

WO 01/83524

PCT/US01/13862

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

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or 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 the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to 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

WO 01/83524

PCT/US01/13862

species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

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

THE INVENTION

The invention is based on the discovery of new human RNA metabolism proteins (RMEP), the polynucleotides encoding RMEP, and the use of these compositions for the diagnosis, treatment, or prevention of nervous system, autoimmune/inflammatory, cell proliferative, and developmental disorders.

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

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

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3

WO 01/83524

PCT/US01/13862

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 RNA metabolism proteins. SEQ ID NO:46 is 29% identical to Glu-tRNA^{Gln} amidotransferase, subunit A, of *Neisseria meningitidis* (GenBank ID g7226601) as determined by the Basic Local Alignment Search Tool (BLAST, see Table 2). The BLAST probability score is 1.3e-37, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:46 also contains amidase signature sequences as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based Pfam database of conserved protein family domains (see Table 3). Data from BLIMPS and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:46 contains amidase signature sequences, features of polypeptides involved in transamination reactions. These data provide evidence that SEQ ID NO:46 is related to the Glu-tRNA^{Gln} amidotransferases found in prokaryotes and some cellular organelles but, until the instant invention, not in humans. SEQ ID NO:47 is 97% identical to the 60S acidic ribosomal protein of *Zea mays* (GenBank ID g790506) as determined by the Basic Local Alignment Search Tool (BLAST, see Table 2). The BLAST probability score is 5.4e-51. SEQ ID NO:47 also contains a 60S acidic ribosomal protein domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based Pfam database of conserved protein family domains (see Table 3). Data from BLIMPS analyses provide further corroborative evidence that SEQ ID NO:47 is a phosphorylated (hence likely to be acidic) ribosomal protein. SEQ ID NO:1-45 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-47 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:48-94 or that distinguish between SEQ ID NO:48-94 and

WO 01/83524

PCT/US01/13862

related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA sequences in column 5 relative to their respective full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 642017H1 is the identification number of an Incyte cDNA sequence, and BRSTNOT03 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 70822015V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g1136841) which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. The Genscan-predicted coding sequences may have been edited prior to assembly. (See Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon-stretching" algorithm. (See Example V.) In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

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

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

The invention also encompasses polynucleotides which encode RMEP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:48-94, which encodes RMEP. The polynucleotide sequences of SEQ ID NO:48-94, as presented in the Sequence Listing, embrace the equivalent RNA sequences,

WO 01/83524

PCT/US01/13862

wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding RMEP. 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 RMEP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:48-94 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:48-94.

10 Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of RMEP.

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

Although nucleotide sequences which encode RMEP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring RMEP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding RMEP 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 RMEP 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 RMEP and RMEP 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 RMEP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing

WO 01/83524

PCT/US01/13862

to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:48-94 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."

- 5 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
10 system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal
cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied
Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA),
15 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 RMEP may be extended utilizing a partial nucleotide
20 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
25 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
30 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

WO 01/83524

PCT/US01/13862

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.

5 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.

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

In another embodiment of the invention, polynucleotide sequences or fragments thereof which 20 encode RMEP may be cloned in recombinant DNA molecules that direct expression of RMEP, 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 RMEP.

The nucleotide sequences of the present invention can be engineered using methods generally 25 known in the art in order to alter RMEP-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, 30 alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or

WO 01/83524

PCT/US01/13862

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

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

In order to express a biologically active RMEP, the nucleotide sequences encoding RMEP 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 RMEP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding RMEP. Such signals

WO 01/83524

PCT/US01/13862

- include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding RMEP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be 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.)
- 10 Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding RMEP and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. (See, e.g., Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)
- 15 A variety of expression vector/host systems may be utilized to contain and express sequences encoding RMEP. 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

WO 01/83524

PCT/US01/13862

upon the use intended for polynucleotide sequences encoding RMEP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding RMEP can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSFORT1 plasmid (Life Technologies). Ligation of sequences encoding RMEP into the vector's multiple cloning site
 5 disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Hecke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509.) When large quantities of RMEP are needed, e.g. for the production of antibodies,
 10 vectors which direct high level expression of RMEP 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 RMEP. 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
 15 vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

Plant systems may also be used for expression of RMEP. Transcription of sequences encoding
 20 RMEP may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can
 25 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 RMEP may be ligated into an
 30 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 RMEP 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-

WO 01/83524

PCT/US01/13862

based vectors may also be used for high-level protein expression.

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

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

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk⁻* and *aprt⁻* 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 RMEP is inserted within a marker gene sequence, transformed cells containing sequences encoding RMEP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding RMEP under the control of a single

WO 01/83524

PCT/US01/13862

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 RMEP and that express RMEP 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 RMEP 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 RMEP 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) Immunocytochemical Protocols, Humana Press, Totowa NJ.)

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

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

WO 01/83524

PCT/US01/13862

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WB38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding RMEP 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 RMEP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of RMEP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunaffinity 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 RMEP encoding sequence and the heterologous protein sequence, so that RMEP 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 RMEP 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.

RMEP of the present invention or fragments thereof may be used to screen for compounds that specifically bind to RMEP. At least one and up to a plurality of test compounds may be screened for specific binding to RMEP. Examples of test compounds include antibodies, oligonucleotides,

WO 01/83524

PCT/US01/13862

proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of RMEP, 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 RMEP 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 RMEP, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing RMEP or cell membrane fractions which contain RMEP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either RMEP 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 RMEP, either in solution or affixed to a solid support, and detecting the binding of RMEP to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

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

In another embodiment, polynucleotides encoding RMEP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse T29/SvJ cell line, are derived from the early mouse embryo

WO 01/83524

PCT/US01/13862

and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains.

10 Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

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

Polynucleotides encoding RMEP 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 RMEP 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 RMEP, e.g., by secreting RMEP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

25 Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of RMEP and RNA metabolism proteins. In addition, the expression of RMEP is closely associated with diseased, proliferative, tumorous, and nervous tissues, adrenal tissue, brain tumor tissue, fetal colon tissue, adult colon tissue, prostate epithelial tissue, lymph node cancer tissue, ovarian tissue, pancreatic tissue, and fetal spleen tissue, as well as with diseases of the lung, and physiological conditions that result in anoxia. Therefore, RMEP appears to play a role in nervous system, autoimmune/inflammatory, cell proliferative, and developmental disorders, as well as neoplasms involving lung-specific tissues. In the treatment of disorders associated with increased RMEP expression or activity, it is desirable to decrease the expression or activity of RMEP. In the treatment of disorders associated with decreased RMEP expression or activity, it is desirable to

WO 01/83524

PCT/US01/13862

increase the expression or activity of RMEP.

Therefore, in one embodiment, RMEP 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 RMEP. Examples of such disorders include, but are not limited to, a nervous system disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Sträussler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebellar hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorder of the central nervous system, cerebral palsy, a neuroskeletal disorder, an autonomic nervous system disorder, a cranial nerve disorder, a spinal cord disease, muscular dystrophy and other neuromuscular disorder, a peripheral nervous system disorder, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathy; myasthenia gravis, periodic paralysis; a mental disorder including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; 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 lymphocytotoxicins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hyper eosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis,

WO 01/83524

PCT/US01/13862

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, and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucopolysaccharidosis, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss.

In another embodiment, a vector capable of expressing RMEP 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 RMEP including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified RMEP 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 RMEP including, but not limited to, those provided above.

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

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

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding RMEP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of RMEP 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

WO 01/83524

PCT/US01/13862

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.

5 An antagonist of RMEP may be produced using methods which are generally known in the art. In particular, purified RMEP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind RMEP. Antibodies to RMEP 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
10 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 RMEP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase
15 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 RMEP
20 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 RMEP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

25 Monoclonal antibodies to RMEP 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.
30 (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

WO 01/83524

PCT/US01/13862

- 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 RMEP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotype composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. 5 (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)
- Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)
- 10 Antibody fragments which contain specific binding sites for RMEP may also be generated. For example, such fragments include, but are not limited to, F(ab')₂ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. 15 (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 RMEP and its specific 20 antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering RMEP 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 RMEP. Affinity is expressed as an association 25 constant, K_a , which is defined as the molar concentration of RMEP-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 RMEP epitopes, represents the average affinity, or avidity, of the antibodies for RMEP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular RMEP epitope, 30 represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the RMEP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of RMEP, preferably in active form, from the antibody (Catty, D. (1988)

WO 01/83524

PCT/US01/13862

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

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a
 5 polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of RMEP-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*.)

10 In another embodiment of the invention, the polynucleotides encoding RMEP, 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 RMEP. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be
 15 designed from various locations along the coding or control regions of sequences encoding RMEP. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

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

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

WO 01/83524

PCT/US01/13862

cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Sommer (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in RMEP expression or regulation causes disease, the expression of RMEP 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 RMEP are treated by constructing mammalian expression vectors encoding RMEP and introducing these vectors by mechanical means into RMEP-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J.-L. and H. Récipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of RMEP include, but are not limited to, the pCDNA 3.1, pPTAG, pRCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), pCMV-SCRIPT, pCMV-TAG, pEGSH/PERV (Stratagene, La Jolla CA), and pTET-OFF, pTET-ON, pTRE2, pTRE2-LUC, pTK-HYG (Clontech, Palo Alto CA). RMEP 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 ecodysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding RMEP from a normal individual.

WO 01/83524

PCT/US01/13862

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

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to RMEP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding RMEP under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and A.D. Miller (1988) *J. Virol.* 62:3802-3806; Dull, T. et al. (1998) *J. Virol.* 72:8463-8471; Zufferey, R. et al. (1998) *J. Virol.* 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

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

WO 01/83524

PCT/US01/13862

hereby incorporated by reference. For adenoviral vectors, see also Antinovi, 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
 5 polynucleotides encoding RMEP to target cells which have one or more genetic abnormalities with respect to the expression of RMEP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing RMEP 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
 10 to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the
 15 appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different
 20 segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding RMEP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on
 25 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 RMEP into the alphavirus
 30 genome in place of the capsid-coding region results in the production of a large number of RMEP-coding RNAs and the synthesis of high levels of RMEP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application

WO 01/83524

PCT/US01/13862

- (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of RMEP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.
- Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) In Huber, B.E. and B.I. Carr, *Molecular and Immunologic Approaches*, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.
- Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding RMEP.
- 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 RMEP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.
- RNA molecules may be modified to increase intracellular stability and half-life. Possible

WO 01/83524

PCT/US01/13862

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

5 wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding RMEP. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not

10 limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased RMEP

15 expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding RMEP may be therapeutically useful, and in the treatment of disorders associated with decreased RMEP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding RMEP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in

20 altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a

25 library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding RMEP 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 RMEP are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected

30 by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding RMEP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the

35 polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide

WO 01/83524

PCT/US01/13862

can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial
 5 library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

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

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

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

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,
 25 intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting
 30 formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

WO 01/83524

PCT/US01/13862

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

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

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

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

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

WO 01/83524

PCT/US01/13862

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

DIAGNOSTICS

In another embodiment, antibodies which specifically bind RMEP may be used for the diagnosis of disorders characterized by expression of RMEP, or in assays to monitor patients being treated with RMEP or agonists, antagonists, or inhibitors of RMEP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for RMEP include methods which utilize the antibody and a label to detect RMEP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring RMEP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of RMEP expression. Normal or standard values for RMEP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to RMEP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of RMEP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding RMEP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of RMEP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of RMEP, and to monitor regulation of RMEP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding RMEP or closely related molecules may be used to identify nucleic acid sequences which encode RMEP. 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 RMEP, allelic variants, or related

WO 01/83524

PCT/US01/13862

sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the RMEP 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:48-94 or from genomic sequences including promoters, enhancers, and introns of the RMEP gene.

Means for producing specific hybridization probes for DNAs encoding RMEP include the cloning of polynucleotide sequences encoding RMEP or RMEP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding RMEP may be used for the diagnosis of disorders associated with expression of RMEP. Examples of such disorders include, but are not limited to, a nervous system disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease; prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Strausler-Scheinker syndrome; fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberculous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorder of the central nervous system, cerebral palsy, a neuroskeletal disorder, an autonomic nervous system disorder, a cranial nerve disorder, a spinal cord disease, muscular dystrophy and other neuromuscular disorder, a peripheral nervous system disorder, dermatomyositis and polymyositis; inherited, metabolic, endocrine, and toxic myopathy; myasthenia gravis, periodic paralysis; a mental disorder including mood, anxiety, and schizophrenic disorders; seasonal affective disorder (SAD); akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paraneoplastic psychoses, postherpetic neuralgia, and Tourette's disorder; 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

WO 01/83524

PCT/US01/13862

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 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, and a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucocutaneous dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss. The polynucleotide sequences encoding RMEP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered RMEP expression. Such qualitative or quantitative methods are well known in the art.

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

WO 01/83524

PCT/US01/13862

treatment of an individual patient.

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

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

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

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

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding RMEP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are
30 substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding RMEP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal

WO 01/83524

PCT/US01/13862

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 ISSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as

5 DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs

10 may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of RMEP 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; Duplax, C. et

15 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

20 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

25 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.

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

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of

WO 01/83524

PCT/US01/13862

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 Seilhammer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by

5 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.

10 Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression *in vivo*, as in the case of a tissue or biopsy sample, or *in vitro*, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with *in vitro* model systems and preclinical evaluation of

15 pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to

20 that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The

25 normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at

30 <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the

WO 01/83524

PCT/US01/13862

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.

5 Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at
10 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
15 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
20 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.

25 A proteomic profile may also be generated using antibodies specific for RMEP to quantify the levels of RMEP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lucking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendez, L.G. et al. (1999) *Biotechniques* 27:778-788). Detection may be performed by a variety of
30 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.

WO 01/83524

PCT/US01/13862

Seilhamer (1997) Electrophoresis 18:533-537), so proteomic 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.

5 In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound
10 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
15 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.
20 USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalen, 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.

25 In another embodiment of the invention, nucleic acid sequences encoding RMEP 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
30 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.

WO 01/83524

PCT/US01/13862

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.)

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

10 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

15 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

20 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, RMEP, 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

25 solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between RMEP 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

30 synthesized on a solid substrate. The test compounds are reacted with RMEP, or fragments thereof, and washed. Bound RMEP is then detected by methods well known in the art. Purified RMEP 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

WO 01/83524

PCT/US01/13862

antibodies capable of binding RMEP specifically compete with a test compound for binding RMEP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with RMEP.

In additional embodiments, the nucleotide sequences which encode RMEP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

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

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/201,875, U.S. Ser. No. 60/200,184, U.S. Ser. No. 60/202,090, U.S. Ser. No. 60/210,232, and U.S. Ser. No. 60/220,553, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

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

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPRT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units

WO 01/83524

PCT/US01/13862

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., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

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

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

III. Sequencing and Analysis

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

WO 01/83524

PCT/US01/13862

carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.) 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 and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

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

WO 01/83524

PCT/US01/13862

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

- 5 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:48-94. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

- 10 Putative RNA metabolism 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) *Curr. Opin. Struct. Biol.* 8:346-354). The program concatenates predicted exons to form an
- 15 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 RNA metabolism proteins, the encoded polypeptides were analyzed by querying against PFAM models for RNA metabolism proteins. Potential RNA metabolism proteins were also
- 20 identified by homology to Incyte cDNA sequences that had been annotated as RNA metabolism 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
- 25 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
- 30 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

WO 01/83524

PCT/US01/13862

to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence.

- 5 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
- 10 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 gbprl public
- 15 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
- 20 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
- 25 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
- 30 determine whether it contained a complete gene.

VI. Chromosomal Mapping of RMEP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:48-94 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

WO 01/83524

PCT/US01/13862

SEQ ID NO:48-94 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 Genomic Research (WIGR), and Génethon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

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

In this manner, SEQ ID NO:53 was mapped to chromosome 1 within the interval from 159.6 to 164.1 centiMorgans. SEQ ID NO:61 was mapped to chromosome 8 within the interval from 30.70 to 60.00 centiMorgans. SEQ ID NO:69 was mapped to chromosome 10 within the interval from 158.30 centiMorgans to the q terminus. SEQ ID NO:70 was mapped to chromosome 1 within the interval from 63.90 to 74.80 centiMorgans. SEQ ID NO:71 was mapped to chromosome 1 within the interval from 159.60 to 164.10 centiMorgans. SEQ ID NO:73 was mapped to chromosome 11 within the interval from 34.30 to 37.00 centiMorgans. SEQ ID NO:75 was mapped to chromosome 2 within the interval from 107.10 to 118.00 centiMorgans. SEQ ID NO:76 was mapped to chromosome 7 within the interval from 7.80 to 10.60 centiMorgans. SEQ ID NO:79 was mapped to chromosome 22 within the interval from 22.20 to 40.20 centiMorgans. SEQ ID NO:81 was mapped to chromosome 4 within the interval from the p terminus to 6.70 centiMorgans. SEQ ID NO: 84 was mapped to chromosome 5 within the interval from 156.0 to 157.6 centiMorgans. SEQ ID NO: 88 was mapped to chromosome 11 within the interval from 117.9 to 123.5 centiMorgans. SEQ ID NO:91 was mapped to chromosome 5 within the interval from 152.3 to 155.5 centiMorgans.

VII. Analysis of Polynucleotide Expression

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

WO 01/83524

PCT/US01/13862

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

5 of the search is the product score, which is defined as:

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

10 The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

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

WO 01/83524

PCT/US01/13862

cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of RMEP Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

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

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

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

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with 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

WO 01/83524

PCT/US01/13862

gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

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

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

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:48-94 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 -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10⁷ counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

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

WO 01/83524

PCT/US01/13862

Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

X. Microarrays

- The linkage or synthesis of array elements upon a microarray can be achieved utilizing
- 5 photolithography, piezoelectric printing (ink-jet printing. See, e.g., Baldeschweiler, *supra*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schemm (1999), *supra*). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a
 - 10 substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schemm, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)
 - 15 Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection.
 - 20 After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in
 - 25 detail below.

Tissue or Cell Sample Preparation

- Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 µg/µl oligo-(dT) primer (21mer), 1X first
- 30 strand buffer, 0.03 units/µl RNase inhibitor, 500 µM dATP, 500 µM dGTP, 500 µM dTTP, 40 µM dCTP, 40 µM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 µl volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37° C for 2 hr, each reaction sample (one

WO 01/83524

PCT/US01/13862

with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 µl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 µg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia 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 US Patent No. 5,807,522, incorporated herein by reference. 1 µl of the array element DNA, at an average concentration of 100 ng/µl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

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

Hybridization

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

WO 01/83524

PCT/US01/13862

140 µl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60° C. The arrays are washed for 10 min at 45° C in a first wash 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 (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

WO 01/83524

PCT/US01/13862

to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

XI. Complementary Polynucleotides

Sequences complementary to the RMEP-encoding sequences, or any parts thereof, are used to
 5 detect, decrease, or inhibit expression of naturally occurring RMEP. 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 RMEP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent
 10 promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the RMEP-encoding transcript.

XII. Expression of RMEP

Expression and purification of RMEP is achieved using bacterial or virus-based expression systems. For expression of RMEP in bacteria, cDNA is subcloned into an appropriate vector containing
 15 an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*lac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express RMEP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of
 20 RMEP 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 RMEP 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
 25 transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (SF9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, RMEP is synthesized as a fusion protein with, e.g., glutathione S-
 30 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 RMEP at

WO 01/83524

PCT/US01/13862

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. 5 10 and 16). Purified RMEP obtained by these methods can be used directly in the assays shown in Examples XVI and XVII where applicable.

XIII. Functional Assays

RMEP function is assessed by expressing the sequences encoding RMEP 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 10 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 15 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 20 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 25 expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) *Flow Cytometry*, Oxford, New York NY.

The influence of RMEP on gene expression can be assessed using highly purified populations of 30 cells transfected with sequences encoding RMEP 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

WO 01/83524

PCT/US01/13862

of mRNA encoding RMEP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIV. Production of RMEP Specific Antibodies

RMEP 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 RMEP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for anti-peptide and anti-RMEP activity by, for example, binding the peptide or RMEP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-labeled goat anti-rabbit IgG.

XV. Purification of Naturally Occurring RMEP Using Specific Antibodies

Naturally occurring or recombinant RMEP is substantially purified by immunoaffinity chromatography using antibodies specific for RMEP. An immunoaffinity column is constructed by covalently coupling anti-RMEP 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 RMEP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of RMEP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/RMEP 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 RMEP is collected.

XVI. Identification of Molecules Which Interact with RMEP

RMEP, 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 RMEP, washed, and any wells with labeled RMEP complex are assayed. Data obtained using different concentrations of

WO 01/83524

PCT/US01/13862

RMEP are used to calculate values for the number, affinity, and association of RMEP with the candidate molecules.

Alternatively, molecules interacting with RMEP 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).

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

10 XVII. Demonstration of RMEP Activity

RMEP activity is demonstrated by a polyacrylamide gel mobility-shift assay. In preparation for this assay, RMEP is expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector containing RMEP cDNA. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression and accumulation of RMEP. Extracts containing solubilized proteins can be prepared from cells expressing RMEP by methods well known in the art. Portions of the extract containing RMEP are added to [³²P]-labeled RNA. Radioactive RNA can be synthesized *in vitro* by techniques well known in the art. The mixtures are incubated at 25 °C in the presence of RNase inhibitors under buffered conditions for 5-10 minutes. After incubation, the samples are analyzed by polyacrylamide gel electrophoresis followed by autoradiography. The presence of a band on the autoradiogram indicates the formation of a complex between RMEP and the radioactive transcript. A band of similar mobility will not be present in samples prepared using control extracts prepared from untransformed cells.

In the alternative, ribosomal protein function of RMEP is assessed by expressing the sequences encoding ribosomal proteins 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 Corporation), both of which contain the cytomegalovirus promoter (P_{CMV}). Between 5-10 µg of recombinant vector are transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 µg of an additional plasmid containing sequences encoding a marker protein are cotransfected.

Transient 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

WO 01/83524

PCT/US01/13862

used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties.

FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod MG (1994) Flow Cytometry, Oxford University Press, New York NY.

The influence of ribosomal proteins on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding a ribosomal protein 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, Inc., 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 a ribosomal protein and other genes of interest can be analyzed by northern analysis or microarray techniques.

In the alternative, RMEP activity is measured as the aminoacylation of a substrate tRNA in the presence of [¹⁴C]cysteine. RMEP is incubated with tRNA^{Pro} and [¹⁴C]cysteine (or appropriate tRNA and amino acid substrates) in a buffered solution. [¹⁴C]-labeled product is separated from free [¹⁴C]-amino acid by chromatography, and the incorporated [¹⁴C] is quantified by scintillation counter. The amount of [¹⁴C] detected is proportional to the activity of RMEP in this assay.

In the alternative, RMEP activity is measured by incubating a sample containing RMEP in a solution containing 1 mM ATP, 5 mM Hepes-KOH (pH 7.0), 2.5 mM KCl, 1.5 mM magnesium chloride, and 0.5 mM DTT along with misacylated [¹⁴C]-Glu-tRNA^{Gln} (e.g., 1 μM) and a similar concentration of unlabeled L-glutamine. Following the quenching of the reaction with 3 M sodium acetate (pH 5.0), the mixture is extracted with an equal volume of water-saturated phenol, and the aqueous and organic phases are separated by centrifugation at 15,000 × g at room temperature for 1 min. The aqueous phase is removed and precipitated with 3 volumes of ethanol at -70°C for 15 min. The precipitated aminoacyl-tRNAs are recovered by centrifugation at 15,000 × g at 4°C for 15 min. The pellet is resuspended in 0.25 M KOH, deacylated at 65°C for 10 min., neutralized with 0.1 M HCl (to final pH 6-7), and dried under vacuum. The dried pellet is resuspended in water and spotted onto a cellulose TLC plate. The plate is developed in either isopropanol/formic acid/water or

WO 01/83524

PCT/US01/13862

ammonia/water/chloroform/methanol. The image is subjected to densitometric analysis and the relative amounts of Glu and Gln are calculated based on the Rf values and relative intensities of the spots. RMEP activity is calculated based on the amount of Gln resulting from the transformation of Glu while acylated as Glu-tRNA^{Gln} (adapted from Curnow, A.W. et al. (1997) Proc. Natl. Acad. Sci. 5 94:11819-26).

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

WO 01/83524

PCT/US01/13862

Table 1

Incyte Project ID	Polypeptide Seq ID NO.	Incyte Polypeptide ID	Polynucleotide Seq ID NO.	Incyte Polynucleotide ID
1622129	1	1622129CD1	48	1522129CB1
1820078	2	1820078CD1	49	1820078CB1
1527017	3	1527017CD1	50	1527017CB1
1647264	4	1647264CD1	51	1547264CB1
1721989	5	1721989CD1	52	1721989CB1
1740581	6	1740581CD1	53	1740581CB1
1740714	7	1740714CD1	54	1740714CB1
1850556	8	1850556CD1	55	1850556CB1
1851178	9	1851178CD1	56	1851178CB1
1851338	10	1851338CD1	57	1851338CB1
2039819	11	2039819CD1	58	2039819CB1
2416915	12	2416915CD1	59	2416915CB1
2472784	13	2472784CD1	60	2472784CB1
2538981	14	2538981CD1	61	2538981CB1
2738075	15	2738075CD1	62	2738075CB1
2739049	16	2739049CD1	63	2739049CB1
2860904	17	2860904CD1	64	2860904CB1
3179424	18	3179424CD1	65	3179424CB1
2885096	19	2885096CD1	66	2885096CB1
2901076	20	2901076CD1	67	2901076CB1
3074572	21	3074572CD1	68	3074572CB1
1437895	22	1437895CD1	69	1437895CB1
1454655	23	1454655CD1	70	1454655CB1
1211130	24	1211130CD1	71	1211130CB1
1257715	25	1257715CD1	72	1257715CB1
1342022	26	1342022CD1	73	1342022CB1
2347104	27	2347104CD1	74	2347104CB1
2527070	28	2527070CD1	75	2527070CB1
2685446	29	2685446CD1	76	2685446CB1
366043	30	366043CD1	77	366043CB1
327063	31	327063CD1	78	327063CB1
9380771	32	9380771CD1	79	9380771CB1

Table 1 (cont.)

Incyte Project ID	Polypeptide Seq ID NO.	Incyte Polypeptide ID	Polynucleotide Seq ID NO.	Incyte Polynucleotide Seq ID NO.
325584	30	325584CDB1	30	325584CDB1
454227	31	454227CDB1	31	454227CDB1
454228	32	454228CDB1	32	454228CDB1
454229	33	454229CDB1	33	454229CDB1
454230	34	454230CDB1	34	454230CDB1
454231	35	454231CDB1	35	454231CDB1
454232	36	454232CDB1	36	454232CDB1
454233	37	454233CDB1	37	454233CDB1
454234	38	454234CDB1	38	454234CDB1
454235	39	454235CDB1	39	454235CDB1
454236	40	454236CDB1	40	454236CDB1
454237	41	454237CDB1	41	454237CDB1
454238	42	454238CDB1	42	454238CDB1
454239	43	454239CDB1	43	454239CDB1
454240	44	454240CDB1	44	454240CDB1
454241	45	454241CDB1	45	454241CDB1
454242	46	454242CDB1	46	454242CDB1
454243	47	454243CDB1	47	454243CDB1

Table 2

Poly-peptide SEQ ID NO:	Incyte Poly-peptide ID	GenBank ID	Probability Score	GenBank Homology
1	162215CD1	g9527590	1.00E-156	[Fl] (Homo sapiens) (AF281133) exosome
2	1527017CD1	g4584132	8.20E-87	30S ribosomal protein S7 homolog (Homo sapiens)
4	1647264CD1	g5551037	2.80E-38	similar to RNA binding protein [Mus musculus domesticus]
5	1721389CD1	g868267	9.80E-02	Weak similarity to ribosomal protein L14 (SP: R14L4.CHUTP, P28353) (Caenorhabditis elegans).
6	1730581CD1	g3721940	7.70E-91	Wilson, R. et al. (1994) 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans. Nature 353:32-38.
7	1740714CD1	g2570925	6.00E-139	R037 (Xenopus laevis)
8	1850598CD1	g619302	2.80E-129	Survival of motor neuron protein interacting protein 1; SM1; SMN protein interacting protein 1 (Homo sapiens). The SMN-SP1 complex is essential for the survival of spinal cord motoneurons. Cell 90:1023-1029.
9	1856109CD1	g2684625	4.70E-05	R93-binding protein-Nerc (alternatively spliced), murine ceratocarboxima cell line, PC04. Duhi, D.M. et al. (1994) Pleiotropic effects of the mouse lethal yellow (Ay) mutation explained by deletion of a maternally expressed gene and the simultaneous production of agouti fusion RNAs. Development 120:1695-1708.
10	1921719CD1	g1843111	1.00E-10	Ribonuclease III (rnc) (Borrelia burgdorferi).
11	2099829CD1	g5015629	5.20E-89	Phosphatase 44 (1997) gene sequence of a parasite. Phosphatase, Borrelia burgdorferi. Nature 390:580-586.
12	2416915CD1	g9347090	1.00E-131	Sarcos (Arabidopsis thaliana) muscle protein 634 (Mus musculus) [Fl] (Mus musculus) putative zinc finger protein FL121

WO 01/83524

PCT/US01/13862

Table 2 (cont.)

Polypeptide Seq ID NO.	Incyte polypeptide ID	GenBank ID NO.	Probability Score	GenBank Homolog
13	2472734CD1	U058575	8.50E-163	F7551_2 [Homo sapiens]
14	2595991CD1	54220483	2.80E-141	putative cleavage and polyadenylation specificity factor 1 (Homo sapiens) { Enr.A. et al. (1999) Nucleic Acids 402: 761-768}
15	2738075CD1	29531845	5.00E-12	[Homo sapiens] RNA-binding protein
16	3179424CD1	27106069	8.00E-08	[Homo sapiens] RNA-binding protein
17	2885096CD1	25102832	3.50E-79	putative mitochondrial ribosomal protein 134 family
18	3179424CD1	27106069	8.00E-08	putative mitochondrial ribosomal protein 134 family
19	2885096CD1	25102832	3.50E-79	putative mitochondrial ribosomal protein 134 family
20	2901076CD1	57158880	0	AF0821 (Apolipoprotein B mRNA editing protein) and Phorbollin [Homo sapiens]
21	3074572CD1	91381029	0	[Homo sapiens] serine-arginine-rich RNA polymerase I associated factor (PAF53) [Mus musculus] (Hirokawa et al. (1996) EMBO J. 15: 2260-2269)
22	1437395CD1	22866613	6.60E-140	putative 40S ribosomal protein
23	1451556CD1	2881503	5.30E-09	ribosomal protein S15 [Thermotoga maritima]
24	1211310CD1	3321540	8.50E-92	ribosomal protein S15 [Thermotoga maritima]
25	1257715CD1	21870014	3.60E-26	ribosomal protein S14 [Cricetulus cricetus]
26	1345092CD1	21045525	1.60E-76	putative arginine-aspartate-rich RNA binding protein (Arabidopsis thaliana)
27	194704CD1	21699023	1.90E-46	putative arginine-aspartate-rich RNA binding protein (Arabidopsis thaliana)
28	607276CD1	24138828	1.60E-18	[Candida albicans] ribosomal protein S9 small subunit precursor
29	758546CD1	22440181	4.00E-46	putative 40S ribosomal protein
30	866043CD1	23283220	8.60E-33	[Schizosaccharomyces pombe]
31	927085CD1	24100563	2.60E-58	splicing factor HRP23 [Homo sapiens]
32	938071CD1	212654241	1.00E-102	ribonuclease P protein subunit p14 [Homo sapiens] (BC000940) Similar to splicing factor, arginine/serine-rich 4 (SRP75)

WO 01/83524

PCT/US01/13862

Table 2 (cont.)

Polypeptide SEQ ID NO:	Incyte polypeptide ID	GenBank ID NO:	Probability Score	GenBank Homolog
33	329594CD1	g672454	0	Spermatid perinuclear RNA binding protein [Mus musculus] Schacter, J. M. et al. (1995) J. Cell Biol. 129:1023-1032
34	4545237CD1	g6899218	1.60E-16	ribosomal protein S5 [Ureaplasma urealyticum]
35	4542946CD1	g10803047	1.00E-68	[Zell] [Zea mays] 40S ribosomal protein S24
36	5702144CD1	g3328106	1.10E-82	Translational release factor 1 [Homo sapiens] Zhang, Y. and Spremulli, L. L. (1998) Biochim. Biophys. Acta 1443:245-250
37	5862945CD1	g1001933	1.30E-08	ribosomal protein L22 [Thermus thermophilus]
38	6319547CD1	g857302	1.00E-07	[5' income] [Leishmania major] PolyA Binding Protein 1
39	000124CD1	g3860586	1.10E-43	PGY(A) POLYMERASE (ncmB) [Rickettsia prowazekii]
40	1659674CD1	g2950473	6.70E-13	RNA-dependent rna polymerase polypeptide 1 [Homo sapiens]
41	2267892CD1	g3646126	5.70E-205	ribosomal RNA helicase [Homo sapiens]
42	2670307CD1	g3152934	6.20E-176	[Mus musculus] 50S coactivator Jab1 Aravind, L. and Ponting, C. P. (1998) Protein Sci. 7:1250-1254
43	4524210CD1	g3258435	1.80E-38	389aa long hypothetical nucleolar protein [Pyrococcus]
44	5584860CD1	g3879784	1.50E-116	Similar to RNA recognition motif (aka RRM, RBD) [C. elegans]
45	5807892CD1	g1573164	5.70E-15	ribosomal protein S16 (rps16) [Haemophilus influenzae]
46	3210644CD1	g7226601	1.30E-37	Glu-tRNA(Gln) amidotransferase, subunit A [Neisseria meningitidis] Schmidt, H. (2000) Science 287:1809-1815
47	4542454CD1	g790508	5.40E-51	60S acidic ribosomal protein [Zea mays] Goddemeier, M. I. et al. (1996) Plant. Mol. Biol. 30:655-658.

WO 01/83524

PCT/US01/13862

Table 3

SEQ ID NO: ID	Protein ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	1622139CD1	245	S82, S119, S174, Y226		Amino acid tRNA ligase motif (As tRNA ligase 11.19c motif): Y12-D35 3' exoribonuclease family Ribonuclease H1 Ribonuclease H Proteins: B0127C; P117-1147 NUCLEOTIDYLTRANSFERASE TRANSFERASE POLYRIBONUCLEOTIDE PROTEIN PHOSPHORYLASE POLYNUCLEOTIDE PHOSPHORYLASE PH PNPASE PHBINDING: P002075; R13-1228 NUCLEOTIDYLTRANSFERASE; NUCLEOTIDYLTRANSFERASE; POLYNUCLEOTIDE DOMAIN; DN035201P50849 [1-615; L6-R222 RIBONUCLEOPROTEIN HETEROGENEOUS NUCLEAR U SCAFFOLD ATTACHMENT FACTOR A HNRNP PROTEIN (HOU1)); P004707; P34-Y100 [P = 1.3e-05] Signal cleavage: M1-A54 Ribosomal protein S7L/S5; S4-W178 Ribosomal protein S7 protein; BL000523 Y74-R120, BL00052B; K145-R171 Ribosomal protein S10 protein; BL00161A; Y107-K122 Ribosomal protein; Ribosomal S7; S5f; K1-H83 RIBOSOMAL PROTEIN S7 30S RNA- BINDING CHLOROPLAST CHLOROPLAST RIBOSOMAL S5; P004001; S4-W178 RIBOSOMAL PROTEIN S7; BL000324; P229765 [20-155; P31-R177 DN00324	EMBL-FAST EMBL-FAST BLIPS-BLOCKS BLIPS-BLOCKS PROFILESCAN BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM
2	1922078CD1	118	T20	N77		
3	1527017CD1	179	S4, S5, S9, S143, S67, S143			SPSCAN EMBL-FAST BLIPS-BLOCKS BLIPS-BLOCKS

Table 3 (cont.)

SEQ ID NO.	Incrte	Amino Acid Residues	Potential phosphorylation sites	Potential glycosylation sites	Signature Sequences, Domains and Motifs	Analytical Methods and Database
4	I64723GCD1	101	S66, T36, T47	N19	Transmembrane domain (transmem.domain): L37-Y45_S64-Y86	HMMER
5	I73199GCD1	145	T97, S126, Y134		RNA binding protein homolog (R07E5.12): PD068568; P15-D101 Signal cleavage (signal_cleavage): M1-S19	BIAST-PRODOM SPSCQN
6	I73059LCD1	249	M14, S18, S49, S75, S133, T243		Ribosomal protein L14: A31-V145 Ribosomal protein L14: BL00049C:P84-K129 SIMILARITY TO RIBOSOMAL PROTEIN L14 RIBOSOMAL PROTEIN: PD080736: V32-V145 NO27 PROTEIN (predicted nuclear protein): PD139112: M1-H113	HMMER-PFAM BLIPS-BLOCKS BIAST-PRODOM BIAST-PRODOM BIAST-PRDMMO BIAST-DOMO
7	I740714CD1	265	S19, T48, S81, T112, S136, S176, S227, S235	N83, M113	RNF-E3RING EGG-BOX DOMAIN: DM077401[S32644]S30-703-R141-S205 RNF-E3RING EGG-BOX DOMAIN: DM04007[S48131]21-239-G103-P202 PROTEIN SURVIVAL OF MOTOR NEURON INTERACTING: PD035299: M25-D262	BIAST-DOMO BIAST-DOMO BIAST-PRODOM
8	I950559GCD1	306	S2, M1, S36, S177, S199, S206, S252, T262, T295, T286, S288, T298,	N9, N761	Rap i motif: K55-V62 RNA recognition motif. (a.k.a. RRM and RBD): ZFM.V23-187 RIBONUCLEOPROTEIN NUCLEAR PROTEIN BINDING ENERGY HETEROGENEOUS C1 NUCLEOPROTEIN C PHOSPHORYLATION: PDJ35584: E88-D220 RRM: EUKARYOTIC CYTOSOLIC DM06681[S47318]5-293_G-A304 Yeast-like RNA-binding domain: BL000308: M55-N64	MOTIFS HMMER-PFAM BIAST-PRODOM BIAST-DOMO BLIPS-BLOCKS

Table 3 (cont.)

SEQ ID	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Post-translational Modifications	Signature Sequences, Domains and Motifs	Analytical Methods and References
9	1056103CD1	332	S65, S131, T161, M129, S206, T034, S251, T322,	S83, K249, K315	Signal peptide (signal_peptide): M1-G30 Double-stranded RNA binding motif (dsrm): P237-L304 (score = 0.1) HYPOTHETICAL 43.1 KD PROTEIN F02A9.4 IN CHROMOSOME III: P140911; D67-R311 PROTEIN REBINDING RNA REPEAT (PROTEIN REBINDING RNA REPEAT) DOUBLESTRANDED RNA-BINDING: P0001471; P237-L304 (p = 0.0090)	BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM
10	1931719CD1	279	S28, S85, T138, T183, K184, T214, T256		Signal cleavage (signal_cleavage): M1-Q42 DECOY 60S RIBOSOMAL PROTEIN L30 MITOCHONDRIAL PRECURSOR YML30 MITOCHONDRIAL TRANSIT: P0037326; R136-P205 (p = 1.6e-08)	SPSCAN BLAST-PRODOM
11	2098629CD1	239	S4, S11, T15, T16, Y39, S80, S93, T105, T108, T150, Y124, S171, K203, S243, S245, S249, S233, S235,	M106	Acidic protein L10 Ribosomal L101: K16-T117 PROTEIN RIBOSOMAL SIMILAR 60S ACIDIC PO UPDACA1 INTERGENIC REGION: P0037265: M1-G213 RAT ACIDIC RIBOSOMAL PROTEIN: L10M09041; P257641-318; V10-S212 SUPPRESSOR OF SABLE RNA-BINDING NUCLEIC HOMOLOG P0032578; K225-K288	HOMER-PTAM BLAST-PRODOM BLAST-PRODOM BLAST-PRODOM
12	2416915CD1	291	T18 T20 T32 T37 S77, T113 S165 T238 S269 S76			

SEQ ID	Incyte ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and motifs	Analytical Methods and Accession Motifs
NO: 13	247278CDL	431	S29 S131 S37 S38 S43 C399 S230 T399		Eukaryotic putative RNA-binding region RNP-1 Signature X393-F400 ATP/GTP-binding site motif A (P-loop) G22-S29 RNA recognition motif. (RRM, RBD, or NRP domain) T354-L425	MOTIFS
					Eukaryotic RNA-binding RcO0010A:L354-P372 F354-G399;E353-D402 T354-T361 D057917.L1444-P206 TD056050.G305-E353 EBANCCLEAPROTEIN REPEAT	BIAST_BIMPS
					Dm00012.Q10355I25-I06.N347-G498 Dm00012.P32568.I54-238.D350-L425 Dm00012.T31483I87-I72.A349-K424 Dm00012.O05566I1-83.D351-S478	BIAST_PRODOM
14	259489ICDI	600	T209 T254 T520 Y577 S213 T572 T522 Y55 I315		POLYADENYLATION CLEAVAGE SPECIFICITY RNA BINDING SD005421.F1479-N370 eukaryotic putative RNA-binding region RNP-1 Signature K51-P56	BIAST_PRODOM
15	273807CDL	317	K242 E296 S232 S144 S259 T194 S210	N66	Signal cleavage: H1-A64 RNA recognition motif. (RRM, RBD, or NRP domain) Ttm-VIP-R83	MOTIFS
					eukaryotic RNA-binding region RNP-1 Proteins SLC0010A:V171-P500 SLC0010A:V171-P500	BIMAPS_BLOCKS

Table 3 (cont.)

SEQ ID	Incyte ID	Amino acid Residues	Potential phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Potential Proteomic Annotations	Analytical Methods and Platforms	
15					<p>Biotinylated putative RNA-binding region K21-1 signature K21-1: R23-135 RIBONUCLEOPROTEIN REPEAT DK0012: 920940 151-238-V12-A86 DK0012: Q04335 234-341-V12-A86 DK0012: P13339 205-288-S10-W73 DK0012: P31159 13-86-57-A86</p>	PROTEOMSCAN	BLIMPS_DOMO
16	2279049CD1	319	S275 W20 S234 S280 T119 T258	NI55	<p>Adenine-1-transfer RNA synthetases class II S148-188 4255-4278 Ribosomal protein S15 signature G30-G101</p>	PROTEOMSCAN	NOTIF9S
17	2650904CD1	168			<p>signal_cleavage: M1-S15 Ribosomal protein L34 N51-H92</p>	PROTEOMSCAN	SFSCAN
18	3179424CD1	92	S71 T12		<p>Ribosomal protein L34 E100784:650-687 Cytidine and deoxycytidylate deaminases zinc-binding region H144-V182 signal_cleavage: M1-S21</p>	PROTEOMSCAN	BLIMPS_BLOCKS
19	2855096CD1	268	T73 T162 S241 I91 I159 S222	NI09 N193	<p>Cytidine and deoxycytidylate deaminases zinc-binding region H144-V182 signal_cleavage: M1-S21</p>	PROTEOMSCAN	NOTIF9S
					<p>Cytidine and deoxycytidylate deaminases zinc-binding region H144-V182 E100903: Y169-CL78 APOLIPROTEIN B RNA EDITING PROTEIN DK0012: 951908 1-236-K126-V236 DK0012: P41236 1-235-H144-V236 DK0012: P41236 1-235-H144-V236 DK0012: P41236 1-235-H144-V236</p>	PROTEOMSCAN	BLIMPS_BLOCKS
					<p>APOLIPROTEIN B RNA EDITING PROTEIN DK0012: 951908 1-236-K126-V236 DK0012: P41236 1-235-H144-V236 DK0012: P41236 1-235-H144-V236 DK0012: P41236 1-235-H144-V236</p>	PROTEOMSCAN	BLAST_DOMO

WO 01/83524

PCT/US01/13862

Table 3 (cont.)

SEQ ID NO: 1D	Inverte PolyPeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
20	2301076CD1	524	S329 S351 S359 S367 S380 S392 S401 S406 S418 S41 S427 S438 S434 S487 S493 S505 S508 T511 T514 S524 S543 S553 S573 S618 T159 S175 T194 S557 T519	N25 N254 N339 N495 N570 N616	RNA recognition motif. (RRM, RBD, or RNP domain) rmc1, rmc2, rmc3 RBD ARGININE RICH SOLICISSOME SPLICING FACTOR PD037489:V20-G139 TYPE B REPETAT REPEAT DM05511 [S26650] 1-1203:P264-P539 DM05511 [P18583] 113-1236:P264-P539	BLAST_F7AM
21	3074572CD1	419	S8 S20 T64 T130 T144 S143 S151 S164 S166 S168 S296 S385 T419 T43 S137 T174 S231 S230 T239 S265 S390	N93 N161 N237 N254 N255 N256 N257 N258 N259 N260 N261 N262 N263 N264 N265 N266 N267 N268 N269 N270 N271 N272 N273 N274 N275 N276 N277 N278 N279 N280 N281 N282 N283 N284 N285 N286 N287 N288 N289 N290 N291 N292 N293 N294 N295 N296 N297 N298 N299 N300 N301 N302 N303 N304 N305 N306 N307 N308 N309 N310 N311 N312 N313 N314 N315 N316 N317 N318 N319 N320 N321 N322 N323 N324 N325 N326 N327 N328 N329 N330 N331 N332 N333 N334 N335 N336 N337 N338 N339 N340 N341 N342 N343 N344 N345 N346 N347 N348 N349 N350 N351 N352 N353 N354 N355 N356 N357 N358 N359 N360 N361 N362 N363 N364 N365 N366 N367 N368 N369 N370 N371 N372 N373 N374 N375 N376 N377 N378 N379 N380 N381 N382 N383 N384 N385 N386 N387 N388 N389 N390 N391 N392 N393 N394 N395 N396 N397 N398 N399 N400 N401 N402 N403 N404 N405 N406 N407 N408 N409 N410 N411 N412 N413 N414 N415 N416 N417 N418 N419 N420 N421 N422 N423 N424 N425 N426 N427 N428 N429 N430 N431 N432 N433 N434 N435 N436 N437 N438 N439 N440 N441 N442 N443 N444 N445 N446 N447 N448 N449 N450 N451 N452 N453 N454 N455 N456 N457 N458 N459 N460 N461 N462 N463 N464 N465 N466 N467 N468 N469 N470 N471 N472 N473 N474 N475 N476 N477 N478 N479 N480 N481 N482 N483 N484 N485 N486 N487 N488 N489 N490 N491 N492 N493 N494 N495 N496 N497 N498 N499 N500 N501 N502 N503 N504 N505 N506 N507 N508 N509 N510 N511 N512 N513 N514 N515 N516 N517 N518 N519 N520 N521 N522 N523 N524 N525 N526 N527 N528 N529 N530 N531 N532 N533 N534 N535 N536 N537 N538 N539 N540 N541 N542 N543 N544 N545 N546 N547 N548 N549 N550 N551 N552 N553 N554 N555 N556 N557 N558 N559 N560 N561 N562 N563 N564 N565 N566 N567 N568 N569 N570 N571 N572 N573 N574 N575 N576 N577 N578 N579 N580 N581 N582 N583 N584 N585 N586 N587 N588 N589 N590 N591 N592 N593 N594 N595 N596 N597 N598 N599 N600 N601 N602 N603 N604 N605 N606 N607 N608 N609 N610 N611 N612 N613 N614 N615 N616 N617 N618 N619 N620 N621 N622 N623 N624 N625 N626 N627 N628 N629 N630 N631 N632 N633 N634 N635 N636 N637 N638 N639 N640 N641 N642 N643 N644 N645 N646 N647 N648 N649 N650 N651 N652 N653 N654 N655 N656 N657 N658 N659 N660 N661 N662 N663 N664 N665 N666 N667 N668 N669 N670 N671 N672 N673 N674 N675 N676 N677 N678 N679 N680 N681 N682 N683 N684 N685 N686 N687 N688 N689 N690 N691 N692 N693 N694 N695 N696 N697 N698 N699 N700 N701 N702 N703 N704 N705 N706 N707 N708 N709 N710 N711 N712 N713 N714 N715 N716 N717 N718 N719 N720 N721 N722 N723 N724 N725 N726 N727 N728 N729 N730 N731 N732 N733 N734 N735 N736 N737 N738 N739 N740 N741 N742 N743 N744 N745 N746 N747 N748 N749 N750 N751 N752 N753 N754 N755 N756 N757 N758 N759 N760 N761 N762 N763 N764 N765 N766 N767 N768 N769 N770 N771 N772 N773 N774 N775 N776 N777 N778 N779 N780 N781 N782 N783 N784 N785 N786 N787 N788 N789 N790 N791 N792 N793 N794 N795 N796 N797 N798 N799 N800 N801 N802 N803 N804 N805 N806 N807 N808 N809 N810 N811 N812 N813 N814 N815 N816 N817 N818 N819 N820 N821 N822 N823 N824 N825 N826 N827 N828 N829 N830 N831 N832 N833 N834 N835 N836 N837 N838 N839 N840 N841 N842 N843 N844 N845 N846 N847 N848 N849 N850 N851 N852 N853 N854 N855 N856 N857 N858 N859 N860 N861 N862 N863 N864 N865 N866 N867 N868 N869 N870 N871 N872 N873 N874 N875 N876 N877 N878 N879 N880 N881 N882 N883 N884 N885 N886 N887 N888 N889 N890 N891 N892 N893 N894 N895 N896 N897 N898 N899 N900 N901 N902 N903 N904 N905 N906 N907 N908 N909 N910 N911 N912 N913 N914 N915 N916 N917 N918 N919 N920 N921 N922 N923 N924 N925 N926 N927 N928 N929 N930 N931 N932 N933 N934 N935 N936 N937 N938 N939 N940 N941 N942 N943 N944 N945 N946 N947 N948 N949 N950 N951 N952 N953 N954 N955 N956 N957 N958 N959 N960 N961 N962 N963 N964 N965 N966 N967 N968 N969 N970 N971 N972 N973 N974 N975 N976 N977 N978 N979 N980 N981 N982 N983 N984 N985 N986 N987 N988 N989 N990 N991 N992 N993 N994 N995 N996 N997 N998 N999 N1000 N1001 N1002 N1003 N1004 N1005 N1006 N1007 N1008 N1009 N1010 N1011 N1012 N1013 N1014 N1015 N1016 N1017 N1018 N1019 N1020 N1021 N1022 N1023 N1024 N1025 N1026 N1027 N1028 N1029 N1030 N1031 N1032 N1033 N1034 N1035 N1036 N1037 N1038 N1039 N1040 N1041 N1042 N1043 N1044 N1045 N1046 N1047 N1048 N1049 N1050 N1051 N1052 N1053 N1054 N1055 N1056 N1057 N1058 N1059 N1060 N1061 N1062 N1063 N1064 N1065 N1066 N1067 N1068 N1069 N1070 N1071 N1072 N1073 N1074 N1075 N1076 N1077 N1078 N1079 N1080 N1081 N1082 N1083 N1084 N1085 N1086 N1087 N1088 N1089 N1090 N1091 N1092 N1093 N1094 N1095 N1096 N1097 N1098 N1099 N1100 N1101 N1102 N1103 N1104 N1105 N1106 N1107 N1108 N1109 N1110 N1111 N1112 N1113 N1114 N1115 N1116 N1117 N1118 N1119 N1120 N1121 N1122 N1123 N1124 N1125 N1126 N1127 N1128 N1129 N1130 N1131 N1132 N1133 N1134 N1135 N1136 N1137 N1138 N1139 N1140 N1141 N1142 N1143 N1144 N1145 N1146 N1147 N1148 N1149 N1150 N1151 N1152 N1153 N1154 N1155 N1156 N1157 N1158 N1159 N1160 N1161 N1162 N1163 N1164 N1165 N1166 N1167 N1168 N1169 N1170 N1171 N1172 N1173 N1174 N1175 N1176 N1177 N1178 N1179 N1180 N1181 N1182 N1183 N1184 N1185 N1186 N1187 N1188 N1189 N1190 N1191 N1192 N1193 N1194 N1195 N1196 N1197 N1198 N1199 N1200 N1201 N1202 N1203 N1204 N1205 N1206 N1207 N1208 N1209 N1210 N1211 N1212 N1213 N1214 N1215 N1216 N1217 N1218 N1219 N1220 N1221 N1222 N1223 N1224 N1225 N1226 N1227 N1228 N1229 N1230 N1231 N1232 N1233 N1234 N1235 N1236 N1237 N1238 N1239 N1240 N1241 N1242 N1243 N1244 N1245 N1246 N1247 N1248 N1249 N1250 N1251 N1252 N1253 N1254 N1255 N1256 N1257 N1258 N1259 N1260 N1261 N1262 N1263 N1264 N1265 N1266 N1267 N1268 N1269 N1270 N1271 N1272 N1273 N1274 N1275 N1276 N1277 N1278 N1279 N1280 N1281 N1282 N1283 N1284 N1285 N1286 N1287 N1288 N1289 N1290 N1291 N1292 N1293 N1294 N1295 N1296 N1297 N1298 N1299 N1300 N1301 N1302 N1303 N1304 N1305 N1306 N1307 N1308 N1309 N1310 N1311 N1312 N1313 N1314 N1315 N1316 N1317 N1318 N1319 N1320 N1321 N1322 N1323 N1324 N1325 N1326 N1327 N1328 N1329 N1330 N1331 N1332 N1333 N1334 N1335 N1336 N1337 N1338 N1339 N1340 N1341 N1342 N1343 N1344 N1345 N1346 N1347 N1348 N1349 N1350 N1351 N1352 N1353 N1354 N1355 N1356 N1357 N1358 N1359 N1360 N1361 N1362 N1363 N1364 N1365 N1366 N1367 N1368 N1369 N1370 N1371 N1372 N1373 N1374 N1375 N1376 N1377 N1378 N1379 N1380 N1381 N1382 N1383 N1384 N1385 N1386 N1387 N1388 N1389 N1390 N1391 N1392 N1393 N1394 N1395 N1396 N1397 N1398 N1399 N1400 N1401 N1402 N1403 N1404 N1405 N1406 N1407 N1408 N1409 N1410 N1411 N1412 N1413 N1414 N1415 N1416 N1417 N1418 N1419 N1420 N1421 N1422 N1423 N1424 N1425 N1426 N1427 N1428 N1429 N1430 N1431 N1432 N1433 N1434 N1435 N1436 N1437 N1438 N1439 N1440 N1441 N1442 N1443 N1444 N1445 N1446 N1447 N1448 N1449 N1450 N1451 N1452 N1453 N1454 N1455 N1456 N1457 N1458 N1459 N1460 N1461 N1462 N1463 N1464 N1465 N1466 N1467 N1468 N1469 N1470 N1471 N1472 N1473 N1474 N1475 N1476 N1477 N1478 N1479 N1480 N1481 N1482 N1483 N1484 N1485 N1486 N1487 N1488 N1489 N1490 N1491 N1492 N1493 N1494 N1495 N1496 N1497 N1498 N1499 N1500 N1501 N1502 N1503 N1504 N1505 N1506 N1507 N1508 N1509 N1510 N1511 N1512 N1513 N1514 N1515 N1516 N1517 N1518 N1519 N1520 N1521 N1522 N1523 N1524 N1525 N1526 N1527 N1528 N1529 N1530 N1531 N1532 N1533 N1534 N1535 N1536 N1537 N1538 N1539 N1540 N1541 N1542 N1543 N1544 N1545 N1546 N1547 N1548 N1549 N1550 N1551 N1552 N1553 N1554 N1555 N1556 N1557 N1558 N1559 N1560 N1561 N1562 N1563 N1564 N1565 N1566 N1567 N1568 N1569 N1570 N1571 N1572 N1573 N1574 N1575 N1576 N1577 N1578 N1579 N1580 N1581 N1582 N1583 N1584 N1585 N1586 N1587 N1588 N1589 N1590 N1591 N1592 N1593 N1594 N1595 N1596 N1597 N1598 N1599 N1600 N1601 N1602 N1603 N1604 N1605 N1606 N1607 N1608 N1609 N1610 N1611 N1612 N1613 N1614 N1615 N1616 N1617 N1618 N1619 N1620 N1621 N1622 N1623 N1624 N1625 N1626 N1627 N1628 N1629 N1630 N1631 N1632 N1633 N1634 N1635 N1636 N1637 N1638 N1639 N1640 N1641 N1642 N1643 N1644 N1645 N1646 N1647 N1648 N1649 N1650 N1651 N1652 N1653 N1654 N1655 N1656 N1657 N1658 N1659 N1660 N1661 N1662 N1663 N1664 N1665 N1666 N1667 N1668 N1669 N1670 N1671 N1672 N1673 N1674 N1675 N1676 N1677 N1678 N1679 N1680 N1681 N1682 N1683 N1684 N1685 N1686 N1687 N1688 N1689 N1690 N1691 N1692 N1693 N1694 N1695 N1696 N1697 N1698 N1699 N1700 N1701 N1702 N1703 N1704 N1705 N1706 N1707 N1708 N1709 N1710 N1711 N1712 N1713 N1714 N1715 N1716 N1717 N1718 N1719 N1720 N1721 N1722 N1723 N1724 N1725 N1726 N1727 N1728 N1729 N1730 N1731 N1732 N1733 N1734 N1735 N1736 N1737 N1738 N1739 N1740 N1741 N1742 N1743 N1744 N1745 N1746 N1747 N1748 N1749 N1750 N1751 N1752 N1753 N1754 N1755 N1756 N1757 N1758 N1759 N1760 N1761 N1762 N1763 N1764 N1765 N1766 N1767 N1768 N1769 N1770 N1771 N1772 N1773 N1774 N1775 N1776 N1777 N1778 N1779 N1780 N1781 N1782 N1783 N1784 N1785 N1786 N1787 N1788 N1789 N1790 N1791 N1792 N1793 N1794 N1795 N1796 N1797 N1798 N1799 N1800 N1801 N1802 N1803 N1804 N1805 N1806 N1807 N1808 N1809 N1810 N1811 N1812 N1813 N1814 N1815 N1816 N1817 N1818 N1819 N1820 N1821 N1822 N1823 N1824 N1825 N1826 N1827 N1828 N1829 N1830 N1831 N1832 N1833 N1834 N1835 N1836 N1837 N1838 N1839 N1840 N1841 N1842 N1843 N1844 N1845 N1846 N1847 N1848 N1849 N1850 N1851 N1852 N1853 N1854 N1855 N1856 N1857 N1858 N1859 N1860 N1861 N1862 N1863 N1864 N1865 N1866 N1867 N1868 N1869 N1870 N1871 N1872 N1873 N1874 N1875 N1876 N1877 N1878 N1879 N1880 N1881 N1882 N1883 N1884 N1885 N1886 N1887 N1888 N1889 N1890 N1891 N1892 N1893 N1894 N1895 N1896 N1897 N1898 N1899 N1900 N1901 N1902 N1903 N1904 N1905 N1906 N1907 N1908 N1909 N1910 N1911 N1912 N1913 N1914 N1915 N1916 N1917 N1918 N1919 N1920 N1921 N1922 N1923 N1924 N1925 N1926 N1927 N1928 N1929 N1930 N1931 N1932 N1933 N1934 N1935 N1936 N1937 N1938 N1939 N1940 N1941 N1942 N1943 N1944 N1945 N1946 N1947 N1948 N1949 N1950 N1951 N1952 N1953 N1954 N1955 N1956 N1957 N1958 N1959 N1960 N1961 N1962 N1963 N1964 N1965 N1966 N1967 N1968 N1969 N1970 N1971 N1972 N1973 N1974 N1975 N1976 N1977 N1978 N1979 N1980 N1981 N1982 N1983 N1984 N1985 N1986 N1987 N1988 N1989 N1990 N1991 N1992 N1993 N1994 N1995 N1996 N1997 N1998 N1999 N2000 N2001 N2002 N2003 N2004 N2005 N2006 N2007 N2008 N2009 N2010 N2011 N2012 N2013 N2014 N2015 N2016 N2017 N2018 N2019 N2020 N2021 N2022 N2023 N2024 N2025 N2026 N2027 N2028 N2029 N2030 N2031 N2032 N2033 N2034 N2035 N2036 N2037 N2038 N2039 N2040 N2041 N2042 N2043 N2044 N2045 N2046 N2047 N2048 N2049 N2050 N2051 N2052 N2053 N2054 N2055 N2056 N2057 N2058 N2059 N2060 N2061 N2062 N2063 N2064 N2065 N2066 N2067 N2068 N2069 N2070 N2071 N2072 N2073 N2074 N2075 N2076 N2077 N2078 N2079 N2080 N2081 N2082 N2083 N2084 N2085 N2086 N2087 N2088 N2089 N2090 N2091 N2092 N2093 N2094 N2095 N2096 N2097 N2098 N2099 N2100 N2101 N2102 N2103 N2104 N2105 N2106 N2107 N2108 N2109 N2110 N2111 N2112 N2113 N2114 N2115 N2116 N2117 N2118 N2119 N2120 N2121 N2122 N2123 N2124 N2125 N2126 N2127 N2128 N2129 N2130 N2131 N2132 N2133 N2134 N2135 N2136 N2137 N2138 N2139 N2140 N2141 N2142 N2143 N2144 N2145 N2146 N2147 N2148 N2149 N2150 N2151 N2152 N2153 N2154 N2155 N2156 N2157 N2158 N2159 N2160 N2161 N2162 N2163 N2164 N2165 N2166 N2167 N2168 N2169 N2170 N2171 N2172 N2173 N2174 N2175 N2176 N2177 N2178 N2179 N2180 N2181 N2182 N2183 N2184 N2185 N2186 N2187 N2188 N2189 N2190 N2191 N2192 N2193 N2194 N2195 N2196 N2197 N2198 N2199 N2200 N2201 N2202 N2203 N2204 N2205 N2206 N2207 N2208 N2209 N2210 N2211 N2212 N2213 N2214 N2215 N2216 N2217 N2218 N2219 N2220 N2221 N2222 N2223 N2224 N2225 N2226 N2227 N2228 N2229 N2230 N2231 N2232 N2233 N2234 N2235 N2236 N2237 N2238 N2239 N2240 N2241 N2242 N2243 N2244 N2245 N2246 N2247 N2248 N2249 N2250 N2251 N2252 N2253 N2254 N2255 N2256 N2257 N2258 N2259 N2260 N2261 N2262 N2263 N2264 N2265 N2266 N2267 N2268 N2269 N2270 N2271 N2272 N2273 N2274 N2275 N2276 N2277 N2278 N2279 N2280 N2281 N2282 N2283 N2284 N2285 N2286 N2287 N2288 N2289 N2290 N2291 N2292 N2293 N2294 N2295 N2296 N2297 N2298 N2299 N2300 N2301 N2302 N2303 N2304 N2305 N23		

WO 01/83524

PCT/US01/13862

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
23	145456CD1	284	S57 S160 T232 T244 T279 S48 S165 T200		Ribosomal protein S15 (ribosomal_S15): S51-S215 Ribosomal protein S15 signature (ribosomal_S15_Psf): T146-W215	HMMS-PFAM PROSITE-SCAN
24	121130CD1	248	T14 S49 S74 S132 S18 T242		Nucleolin: DN02740 S2244 G30-703:R140-G204	ELAST-PRODOM ELAST-DOMO
25	1257715CD1	214	S137 T112 S181		Peptidyl-tRNA hydrolase (Pept-tRNA hydrol): W1-Q208 E. coli ribosomal protein S11: E01132C:148-195 E. coli ribosomal protein S11: E01132D:116-164 Peptidyl-tRNA hydrolase (PD005324): M32-T199	HMMS-PFAM ELINPS-BLOCKS ELAST-PRODOM
26	1342022CD1	184	T2 S16 S19 S81 S172 W17 S65	N71 N136	Peptidyl-tRNA hydrolase: DN02080 P4462 1-193:M32-T199 Ribosomal protein S11: G62-R180 Ribosomal protein S11: H000544:G62-S102 Ribosomal protein S11: H000545:G62-S102 Ribosomal S11: G144-L184 Ribosomal protein S11 (S0001010): G62-R180	HMMS-PFAM ELINPS-BLOCKS PROSITE-SCAN ELAST-PRODOM
27	194704CD1	371	S270 S280 S295 S117 S17 S120 S152 S257 T283 S360 S53 T111 T231 S332 Y173		E. coli ribosomal protein S11: DN00661 P19950 16-148:G51-L184 Ribosomal S11: D564-D171 Aspartate-arginine-rich RNA binding protein (PD017473): F50-K230 E. coli ribosomal protein S11: DN03434 P06621 359-483:R204-K323	MOTIFS ELAST-PRODOM ELAST-DOMO

WO 01/83524

PCT/US01/13862

Table 3 (cont.)

SEQ ID NO.	Protein	Amino Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
28	607270CD1	396	T153 S252 G278 S290 T392 T69 T100 S244 S270 T157 Y218 Y262		Signal peptide: M1-A25 Ribosomal protein S9/S16 G274-R396 Ribosomal protein S9: E100360A:K275-Q301 E100360B:F317-L332 E100360C:L278-R306 Ribosomal protein S9 signature (ribosomal_S9_psf): I306-R396 Ribosomal protein S9 (P001627): G274-R396 Ribosomal protein S9: P003779 P38120 149-277:D263-R396 Ribosomal_S9: G334-L352 S109-D156 Ribosomal protein S4: D000205 P32899 101-174:Y101-Y173 WD domain, G-beta repeat: WD40:G20-D58 WD40:E60-D100 WD40:M199-S239 WD40:A243-H282	HMMER-PFAM HMMER-PFAM BLIIPS-BLOCKS PROFILER PROFILER BLAST-PRODOM BLAST-DOCO MOTIFS HMMER-PFAM BLAST-DOCO HMMER-PFAM MOTIFS
29	758546CD1	184	S153 W159 S153 Y49			
30	866043CD1	282	T10 S57 T76 T103 S109 S136 S153 S200 S239 T5 Y57 S131			
31	927065CD1	125	T35 Y21			MOTIFS

WO 01/83524

PCT/US01/13862

Table 3 (cont.)

SEQ ID NO: ID	Incyto Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
32	938071CD1	365	S23 S24 S12 T172 S186 S187 S201 S202 S203 S254 S261 S116 S132 T161 S176 T177 S178 S213 S215 S217 S219 S223 S224 S257 S265 S269 T282 T300 S301 S325 S330 S24 S27 S331 S269 S110 S125 S149 S193 S202 S246 Y324	N72 N180 N472 N476 N482 N485 N486 N489		MOTIFS
33	3295984CD1	672	Y137 Y297 S24 S36 T56 T58 T120 S182 S343 T345 T427 T435 T448 T449 S274 S490 S496 S3 T62 T120 T159 S182 S201 S333 T345 S435 T438 S510		L331-L402 Leucine zipper Double-stranded RNA binding Protein: G554-S557 Double-stranded RNA binding motif; DEAD_FFAM RNA-binding protein RNA-binding protein 1448-N456, 6411-V424, N558-A567 ZINC FINGER RNA BINDING SPERMATID PROTEIN RHQ-D329 SPERMATID PERINUCLEAR RNA BINDING PROTEIN S278-G584, P330-L392, G452-N513 PERINUCLEAR RNA; SPERMATID; RHQ-D329 DOUBLE-STRANDED RNA BINDING DOMAIN N370-L461, T454-K584	MOTIFS ELIMFS_PPFAM DEAD_FFAM ELIMFS_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_DORO BLAST_DORO

WO 01/83524

PCT/US01/13862

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
34	454537CD1	430	S240 T94 S161 W114 T232 S254 W116 T232 I264 S293 S386		signal peptide H1-G26 Ribosomal protein S5 T222-G35 Ribosomal protein S5 pro W230-A271 T303-S339 Ribosomal protein S5 signature T222-F284 RIBOSOMAL PROTEIN S5 T222-N241 RIBOSOMAL PROTEIN S5 R226-G37 R226-S37 F70-S84_S24e F28-K111 Ribosomal protein S24e E47-A103 Ribosomal protein S24e signature V10-E54, I61-K105 Ribosomal protein S24e Arginine repressor N26-G71 Protein S24e T303-S339 RIBOSOMAL 40S S24 S24e V29-K111	HMMER HMMER_P7FM BLIMPS_BLOCKS PROFILES CAN BLAST_PRODOM BLAST_DOMO NOTES HMMER_P7FM PROFILES CAN BLIMPS_BLOCKS BLIMPS_P7FM BLAST_DOMO BLAST_PRODOM
35	494256CD1	137	T11 T14 W19 S109 Y95 E41 T104	N39		

WO 01/83524

PCT/US01/13862

Table 3 (cont.)

SEQ ID NO	Incyte ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signatures Sequences, Domains and Motifs	Analytical Methods and Databases
36	5702144CD1	380	T169 S247 T45 T342 S362 S371 S376 T343 S321 S377	W255 N340	Leucine Zipper L57-L78 L64-L85 L71-L92 RNF-1 R245-V251 Signal peptide M1-G20 Prokaryotic-type class I peptide chain release factor RF-1 G138-R138 Prokaryotic-type class I peptide chain release factors signature R123-R181, L184-Q226, D238-R284, Q285-R338 Prokaryotic-type class I peptide chain release factors L224-S377 PEPTIDE CHAIN RELEASE FACTOR G138-R138 Signal peptide VI-T35 Ribosomal protein L22 R69-K105, V128-S172 Ribosomal protein L23 signature R245-E192	MOTIFS MOTIFS MEMBER MEMBER PFAM BLIMPS BLOCKS BLAST_DOMO BLAST_PRODOM SPSCAN BLIMPS BLOCKS PROFILES CAN
37	5862945CD1	206	S42 S141 Y33			
38	6319547CD1	190	S96 T140 S188 T178 T179 S29 T73 S128	M12 M136	RNF-1 R69-F76 Eukaryotic putative RNA-binding region RNP-1 signature L45-R37 RNA recognition motif L27-V107 Eukaryotic RNA-binding L27-L45 R69-F76	MOTIFS PROFILES CAN MEMBER PFAM BLIMPS BLOCKS

Protein ID	Incyte Lysopside ID	Amino Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
333	C00124CD1	434	S18 847 T165 S274 T107 S389 S400 T59 T112 S428 S285 T356 Y368	N272 N152 N199 N217	Polya A polynucleotide family T110-S276 RNA BINDING PROCESSING POLYNUCLEOTIDE G53-G248 RNA BINDING POLYMERASE L51-G248	SHMER PFAM BLAST_PRODOM BLAST_DOMO
40	1659474CD1	339	S60 T244 S325 T65 T254 T315 T236		MOTIFS	
41	2267892CD1	599	T18 T19 S86 S129 S213 S282 S233 S320 S344 T397 S457 T503 T569 S569 S22 T92 T585 S77 S99 S110 S381 S422 T517		ATP GTP A A309-T216 Helicase conserved C-terminal domain Y496-T503 DEAD/DEAH box helicase Q178-N389 Helicases conserved C-terminal domain X426-Q507 DEAD-box subfamily ATP-dependent RNA helicase G18-T222, K225-T256, V312-L335, Y465-G510 DEAD-BOX SUBFAMILY ATP-DEPENDENT HELICASES N179-T542 RNA BINDING NUCLEAR DNA FACTOR I422-Q507, D182-S258	BLASTS_PFAM BLASTS_PRODOM BLASTS_DOMO
42	2570307CD1	334	S24 S148 S231 T257 T310 S270 S307	N332	MOV34/MOV/PADI family H53-A314 SUBUNIT 26S MOV14 S12 PADI HOMOLOG S135-K282 S137-S282 S137-S282	SHMER PFAM BLAST_PRODOM BLAST_DOMO

Table 3 (cont.)

SEQ ID	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs and Sites	Analytical Methods and Software
43	4324210CD1	440	T159 T285 S319 T355 T413 S417 Y182 T33 S559 S265		NOG1_W002_sun F236-G307 NOG1_W002_sun family Y194-K369 NOG1_W002_sun family I217-I231, G339-G262, F296-G309, K342-L367	BLAST-DOGO BLAST-PPM BLAST_BLOCKS BLAST_PRODOM
44	5584560CD1	420	G215 S36 S502 T246 S202 S587 T242 T244 S502 Y181	N294	NUCLEAR SUN F120 PROLIFERATING CELL ANTIGEN R176-E379 NOG1_W002_PPM FAMILY L197-V163 Nucleation motif L123-V300 SRP-binding motif L214-F252 SRP BINDING PROTEIN L223-E314	BLAST-DOGO BLAST-PPM BLAST_BLOCKS BLAST_PRODOM
45	5807892CD1	137	T118 T27 S123 S60 T105 T125 T130		Ribosomal protein S16 G24-S84 Ribosomal protein S16 H16-F51, L68-A94	BLAST-DOGO BLAST-PPM BLAST_BLOCKS BLAST_PRODOM
46	3210044CD1	556	T20, T64, T153, T10, S244, S344, S344, S353, S349, S456	N165, N199	AMIDASES: DMU046[S51201]91-465;Y94-G338 PROTEINase: R105Y1, R214-S270 Amidase signature (amidase.pr1): G235-G278 Transmembrane domain (transmem.domain): T33-P58 Amidase: D93-P113, R443-L537 Signal cleavage: M1-G56	BLAST-DOGO BLAST-PPM BLAST_BLOCKS BLAST_PRODOM

WO 01/83524

PCT/US01/13862

Table 3 (cont.)

SEQ ID NO.	Isotype ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
47	4942454CD1	111	S19, S47, S101, S108		RAT ACIDIC RIBOSOMAL PROTEIN P1: EM00632 S54179 1-112: M1-D111 RIBOSOMAL protein P2 Signature: PR004569: A75-A89; PR004569: K95-L109 RIBOSOMAL PROTEIN, ACIDIC 80S PHOSPHORYLATION P2, P1, D12 RIBOSOMAL PROTEIN P1: PD001928: M1-D111 50S acidic ribosomal protein (50s_ribosomal): M1-D111	BLAST-DOMO BLAST-PRIMS BLAST-PRODOM BLAST-PRODOM HMMER-PFAM

WO 01/83524

PCT/US01/13862

Table 4

Polynucleotide Seq ID No.	Incyte Polynucleotide ID	Sequence Length	Selected Segments	Sequence Fragments	5' Position	3' Position
48	1622129CBL	882	-59	642017H1 (BSSTWOT03)	1	270
				433782H1 (HNTPEAT03)	176	882
49	1820078CBL	1220	1-41, 833-1220	71164239V1	647	1220
				70034115D1	564	1220
				12937577G (PGANNOT03)	403	1203
50	1527017CBL	2020	1-209, 1415-2020	6571082H1 (HCLMVN05)	1	490
				1541283H1 (SINTVUT01)	1255	1473
				1824377H1 (GHLAVVU01)	1	255
				1550847G (URETVUT01)	1279	2020
				1527017H1 (UCMCLST01)	744	1407
				SM003056F1	524	858
51	1647266CBL	637	1-40	1439348F1 (PANNOT08)	226	769
				2735749H1 (OAGNNT03)	286	508
				5947264H1 (PAGSTVU05)	149	284
52	1721985CBL	717	-21	911688H1 (EALDNV05)	1	419
				0795372H1 (SYMORAB03)	146	717
53	1730581CBL	2061	1-24, 1219-1514	1525770F1 (UCMCLST01)	1109	1566
				14575087G (COLNFT03)	486	1012
				915914H1 (BSSTWOT04)	306	884
				17305817G (BSSTWUT08)	1546	2061
				25959877G (LUNGNVU03)	953	1565
				102784H1 (AKENOR01)	1	406
54	1740714CBL	1307	1-43, 1111-1142	1740714CT1 (HIDKOR01)	1	1307
				3074952H1 (ZONNUT01)	1	277
55	1850566CBL	1357		820453H1 (KEBANK02)	493	1131
				2018418H1 (TTPINOT01)	813	1160
				1389582H1 (CORENNT03)	966	1357
				1855581H1 (HNTPEAT01)	1	270
				3563740H1 (SCLNNT05)	141	470
				3147502H1 (PENNNV05)	282	643

WO 01/83524

PCT/US01/13862

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
55	1856109CBL	1749	1-59, 1201-1749	6708693HL (HEADROT1)	1411	3738
				1518037FL (BLADTUT04)	460	979
				2867008FS (ADRENUT07)	1547	1743
				873524TR (LUNGAST07)	1073	1743
				184031RS (LUNGAST03)	263	1103
57	1921719CBL	991	1-79	184031RS (LUNGAST03)	263	329
				1921719FS (BREASTUT01)	335	973
				1921719FS (BREASTUT01)	335	973
				1921719FS (BREASTUT01)	697	991
				1921719FS (BREASTUT01)	697	991
58	2099329CBL	1133	1024-1188	1911558FS	713	1065
				1959148FS (SKINUT01)	402	972
				1984189FS (LUNGAST01)	248	889
				3135711HL (SMCKNOT01)	1	274
				313888HL (LUNGNOT02)	690	1006
59	2416915CBL	1454	1-22	1569336FS (UTRKNOT05)	37	519
				3177069FS (UTRKNOT04)	830	1454
				2172869HL (ENDCKNOT03)	761	1026
				3973555HL (HEADROT06)	446	741
				3584701HL (2537FAT01)	1	315
60	2472784CBL	1588	1-242, 286-726	1926194RS (BREASTNOT02)	1151	1588
				2701446HL (OVAFUT010)	775	1115
				906518R2 (COLONNOT03)	1	916
				2654465HL (TUBINOT03)	1	726
				2654465HL (TUBINOT03)	1239	1588
61	2598981CBL	2111	557-1153, 1-22	272142RS (LUNGNOT01)	333	1588
				272142RS (LUNGNOT01)	333	1588
				272142RS (LUNGNOT01)	346	1113
				1272143HL (LUNGNOT04)	367	934
				1593665X16C1 (BREASTNOT14)	269	954
				SB2A06347V1	936	1527
				SB2A04026V1	1236	2111

WO 01/83524

PCT/US01/13862

Table 4 (cont.)

Polynucleotide Seq ID NO.	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
61				233555X11C1 (PANCNOT01)	52	711
62	2738075CBL	1155	1-36, 597-623	233555X11C1 (PANCNOT01)	159	1631
63	2279049CBL	1673		233555X11C1 (PANCNOT01)	1	227
64	2660904CBL	584	1-229, 519-584	233555X11C1 (PANCNOT01)	15	622
65	3179424CBL	978	1-189	233555X11C1 (PANCNOT01)	801	1155
66	2885096CBL	1055	1-278	233555X11C1 (PANCNOT01)	55	1155
67	2901076CBL	2220	1189-1491, 634-662	233555X11C1 (PANCNOT01)	445	1005
				4380307F6 (THRNOT01)	938	1650
				1798333F6 (COLANOT07)	1208	1673
				2458335F6 (PANCNOT01)	1	458
				3050944H1 (LUNGNOT03)	409	687
				71286514V1	387	584
				7093737V1	1	514
				3204102F6 (PANCNOT03)	1	632
				3305851F6 (PANCNOT02)	107	932
				3305851F6 (PANCNOT01)	427	778
				3305851F6 (PANCNOT03)	552	1043
				3305851F6 (PANCNOT05)	813	1055
				3305851F6 (PANCNOT02)	1	454
				3305851F6 (PANCNOT01)	1333	1585
				3305851F6 (PANCNOT01)	333	644
				3305851F6 (PANCNOT01)	552	1152
				3305851F6 (PANCNOT01)	39	491
				3305851F6 (PANCNOT01)	892	1508
				3305851F6 (PANCNOT01)	1	278
				3305851F6 (PANCNOT01)	792	1218
				3305851F6 (PANCNOT01)	1730	2250

WO 01/83524

PCT/US01/13862

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
67				2921502HL (SINRN004)	1457	1740
68	3074572CBL	1890	1-24, 1637-1890	4356445HL (OVARN003)	1553	1740
				4356332HL (S2RN004)	1345	1435
				4356332HL (S2RN004)	1345	1890
				13776386 (WEP1PLB02)	221	725
				386A01507HL	769	1343
				3173158HL (UTRSTW04)	1	273
69	1437695CBL	2893	845-1749, 1-43, 2872-2893	386A01593HL	286	843
				2530813HL (COLATW015)	2042	2283
				4919338HL (TBSRN0011)	1334	1615
				3205068HL (PENSNO03)	1523	1888
				3877384P6 (HEARN006)	305	799
				1493116R6 (PROSN001)	2324	2884
				4840534HL (OSTRN001)	903	1171
				3550395HL (BROND001)	1	244
				5453094HL (COLAD001)	192	757
				1400075FL (BRATW08)	1045	1606
				91377484	2337	2893
				1637895TL (PARCNO08)	2234	2893
				796774FL (PROSN005)	2107	2893
				138477HL (RODNO01)	1394	2893
				91377484	2337	2893
				3526811P6 (PARCNO06)	273	2976
				3526811P6 (PARCNO048)	273	2976
				732459HL (MYCNO001)	523	835
70	14E0656CBL	885	1-47	933801HL (SERNNO01)	805	875
				97691691 (LUNGAS001)	280	867
				484306HL (OSTRN001)	23	314
				91617769	1	414

WO 01/83524

PCT/US01/13862

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyta Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
71	121130CB1	1269	1-42	1457566F6 (COLMET02) 915914E1 (PRSTNCT04) 373083H1 (SMCCN003) 5217827H1 (BRSTNCT33) 576327H1 (BRAANCT01) 916281G 322481G 83048722 (ADENCT01)	530 552 974 258 14 254 254 432	1055 929 1469 521 254 432
72	1257715CB1	1066	835-1066, 1-222	1480736H1 (STOMCT01) 1301049F5 (BLACT005) 1731204F5 (BRSTNCT03) 1908142H6 (CONNTNCT01) 2257143H6 (OVARCT01) 934691H1 (PROSCT07) 1701714F5 (BLACT005) 1922010F5 (BLACT005) 2384972H1 (BLACT005) 1395146H1 (THYRCT02) 2060393H6 (OVARCT03)	1 320 17 1 100 1181 32 1156 1 807	221 825 669 614 639 1420 983 1420 255 1420
73	1342023CB1	639	1-98	507270H1 (BRSTNCT01) 1558939F6 (STMMCT04) 449732E1 (GLMMCT02) 188271H1 (OCHCT01) 485443E1 (HMT32701) 2257393H1 (LANSCT03) 574338H1 (MUSCT01) 004338F6 (MUSCT01)	1 448 780 829 56 1 693	265 745 973 971 564 223 891
74	194704CB1	1430	822-852, 1051-1420	5039960H1 (COLMET01) 856043H6 (BRATCT01)	990 704	1615 952
75	607270CB1	1457	1-103, 660-892		1265	1638
76	758546CB1	1184	1-53			
77	866043CB1	1638	1-561			

WO 01/83524

PCT/US01/13862

Table 4 (cont.)

Polynucleotide Seq ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
77				2295207RS (BRSTN025)	171	638
				3396765HL (UPRSN016)	1	226
				5421555HL (PROST07)	468	715
78	927065CBL	701		004389RS (HMLN010)	838	1355
				80407420L	254	701
				80407420L	1	414
79	938071CBL	1829	1070-829, 513-561, 828-910, 689-766	1545403	1	614
				25708326S (HPO4701)	852	1407
				2344188RS (HESN003)	168	744
				2061033RS (OYASN03)	107	143
				2965597RS (ADREN03)	296	882
				2377642HL (ADREN07)	1210	1425
				9114197S	1335	1829
				3148824HL (ADREN04)	1	274
80	3295884CBL	2541	1-242, 2345-2541	SCGA02054V1	287	1013
				SCGA02437L	184	222
				SCGA02437L (HNF2A0701)	1055	1668
				SCGA02437L	1231	1778
				SCGA10160V1	1588	2118
				4526287RS (LYNHW01)	1	532
				3813241HL (TNSN003)	2255	2541
				SCGA03805V1	591	1162
81	4545237CBL	1347	525-567	15615127S (SEJN0704)	219	830
				16395237S (UPRSN0706)	1226	1647
				16494457S (PROST09)	874	1378
				2013721X28C1	560	1154
				(CESTN0703)		
82	4942964CBL	735	704-735	5832142HL (LUTRN003)	559	832
				557202HL (NSL32701)	1	252
				49429647S (BSAT2N05)	1	735

WO 01/83524

PCT/US01/13862

Table 4 (cont.)

Polynucleotide Seq ID NO:	Inocyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
83	5702144CB1	2614	1-93, 1719-1961	1481581F6 (CORENOT02)	724	1378
				2287840X12F1	464	906
				3861300J1 (BRSTNOT01)		
				3122285F6 (BRSTNOT04)	2107	2616
				2450148G6 (ENDONOT02)	1865	2326
84	5862945CB1	736	1-34	2287840X14F1	370	861
				1329248F1 (PANCNOT07)	1539	2094
				968129X11F1 (BRSTNOT05)	1093	1612
				1635830F6 (UTRNOT06)	1	454
				280637F1 (LIVRNOT02)	106	736
85	6319547CB1	1046	1-33	2244058F6 (PANCNOT02)	1	510
				2383076F6 (LSCJNOT01)	574	1046
				3314256H1 (RUSCNOT07)	509	809
				5571304F6 (TETPANCOT08)	142	782
				591553H1 (RBAVNOT02)	1	241
86	000124CB1	2366	1-32, 2085-2266	000124F6 (UP57NOT01)	1742	2366
				729548F6 (PROSTNOT03)	1103	1600
				3513558F6 (LDNCRNOT03)	1196	1782
				902190A6 (BRSTNOT03)	629	1123
				3013556F6 (BRSTNOT01)	482	1008
87	1659474CB1	1041	927-954, 677-709, 813-857	1865034F6 (BRSTNOT01)	615	1041
				3861300J1 (PROSNOT04)	352	533
				290713214F1 (ENDONOT04)	1	533
				3139431F6 (SMCCNOT02)	726	1230
				1970333F6 (UCHCL5NOT01)	529	1043
88	2267892CB1	2722	2347-2722		888	1350

WO 01/83524

PCT/US01/13862

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
88				3815067HL (YONMNOT03)	1574	1844
				082426RL (RDVZESU01)	1828	2448
				282290FL (POVESIEU01)	2007	2710
				241508RL (BRALNNO01)	1	546
				256073FL (HMTB24U01)	2581	2722
89	2670307CB1	1287	1-382	3225855HL (KIDNNOT09)	1219	1479
				1420655FL (KIDNNOT09)	1276	1828
				009443FL (HMC1NOT01)	315	1287
				1295855FL (PGANNCT03)	565	1287
				1304855FL (PLACNCT02)	722	1287
90	4524210CB1	2426	509-1515	1351855FL (LARICTU02)	1	485
				SC1A00216VL	430	1034
				SC1A00181VL	1231	1780
				2256793X39932 (OVARCTU01)	1645	2225
				SC1A03644VL	1161	1705
91	5584860CB1	2362	1-49 1066-2362	2256793X318D4 (OVARCTU01)	1	517
				SC1A0706VL (OVARCTU03)	1999	2326
				2058779HL (OVARCTU03)	565	1337
				1955765RL (BUSMCTU04)	1954	2362
				1850300UG (LARGCTU03)	1838	2340
92	5807892CB1	731	1-63	842678RL (PROSHU03)	712	1279
				71024273VL (SG0000314)	1	614
				21034418G (BRATCTU03)	1293	1767
				6AEB02217FL	1191	1752
				377638R6 (NEUTPCTU01)	541	1453
				3745265F6 (PHTACTU03)	1	495
				958750RL (KIDNNOT05)	416	731

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
93	321004C81	2088	2051-2088, 1-190, 648-1486, 1811-1843	70822015V1	537	1143
				7705613N1 (VIRETUB01)	1	541
				70821405V1	1195	1807
				70818955V1	511	1069
				6016191H1 (HN220NN03)	1227	1739
94	4942454C81	660	1-23	70821245V1 (830000394)	1653	2088
				4942454G5 (BSA1FEN03)	3	660
				4942454F6 (BSA1FEN03)	1	569

WO 01/83524

PCT/US01/13862

Table 5

Polynucleotide SEQ. ID NO.	Insys Project ID	Representative Library
48	1622129CB1	SNF2AOT01
49	1820078CB1	LUNGNOT20
50	1527017CB1	LUNGNOT14
51	1647264CB1	LUNGNOT27
52	1721989CB1	ISLNOT01
53	1710581CB1	LUNGNOT03
54	1740714CB1	GLYNNOT02
56	1856105CB1	BLA7NOT04
57	1521715CB1	BRSTNOT01
58	2098225CB1	ENANOT01
59	2416915CB1	BLA9NOT05
60	2427844CB1	BRSTNOT02
61	2558381CB1	LIVNONOT03
62	2367052CB1	BRSTNOT02
63	2266045CB1	BRSTNOT02
64	2666004CB1	LUNGNOT03
65	3178424CB1	KUNNOT03
66	2855596CB1	BLA7NOT04
67	2901076CB1	END3NOT02
68	3074572CB1	SNRFBT03
69	1437855CB1	PANCNOT08
70	1454656CB1	ADRENOT01
71	121130CB1	SPINFBT01
72	1257715CB1	MENINVT03
73	1342022CB1	COLNFET02
74	194704CB1	OVARNOT03
76	758545CB1	COLJNOT11
77	855042CB1	BRATNOT03
78	927055CB1	LUNGNOT03
79	938071CB1	EPHNONOT05
80	3292584CB1	THCNOT03

WO 01/83524

PCT/US01/13862

Table 5 (cont.)

Polymer ID NO.	Incule Project ID	Representative Library
81	4545237CB1	SPAN004
82	4542944CB1	SPAN005
83	5702744CB1	SPAN007
84	5662945CB1	SPAN001
85	6319547CB1	SPAN016
86	000124CB1	XIND005
87	1659474CB1	SPAN001
88	2267892CB1	UCCL5T01
89	2670307CB1	SPST001
90	4524210CB1	OVAR001
91	5584660CB1	SPAN005
92	5807892CB1	XIND001
93	3210044CB1	PTUM001
94	4342454CB1	BRATP003

WO 01/83524

PCT/US01/13862

Table 6

Library	Vector	Library Description
ADREN003	PSD08T1	Library was constructed using RNA isolated from the adrenal tissue of a 17-year-old Caucasian male, who died from cerebral anoxia.
BLADU004	P1N1Y	Library was constructed using RNA isolated from bladder tumor tissue removed from a 60-year-old Caucasian male during a radical cystectomy, prostatectomy, and vesectomy. Pathology indicated grade 3 transitional cell carcinoma in the left bladder wall. Carcinoma in-situ was identified in the dome and trigone. Patient history included tobacco use. Family history included Type I diabetes, malignant neoplasm of the stomach, atherosclerotic coronary artery disease, and acute myocardial infarction.
BLADU005	P1N1Y	Library was constructed using RNA isolated from bladder tumor tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy, and urinary diversion. Pathology indicated grade 3 transitional cell carcinoma on the anterior wall of the bladder. Patient history included lung neoplasm and tobacco abuse in remission. Family history included malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease, and lung cancer.
BRAIF003	P1N1Y	This normalized fetal brain tissue library was constructed from 3.26 million independent clones from a fetal brain library. Starting RNA was made from brain tissue removed from a Caucasian male fetus, who was stillborn with a hypoplastic left heart at 23 weeks' gestation. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9239 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRAIF005	P1N1Y	This normalized fetal brain tissue library was constructed from 3.26 million independent clones from a fetal brain library. Starting RNA was made from brain tissue removed from a Caucasian male fetus, who was stillborn with a hypoplastic left heart at 23 weeks' gestation. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9239 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRAIN014	P1N1Y	Library was constructed using RNA isolated from brain tissue removed from the left frontal lobe of a 40-year-old Caucasian female during excision of a cerebral meningial lesion. Pathology for the associated tumor tissue indicated grade 4 gemistocytic astrocytoma.

Table 6 (cont.)

Library	Vector	Library Description
BRAT0702	PSORT11	Library was constructed using RNA isolated from brain tumor tissue removed from the frontal lobe of a 58-year-old Caucasian male during excision of a cerebral meningioma lesion. Pathology indicated a grade 2 metastatic meningioma. Patient history included a grade 2 renal cell carcinoma, insulinoma, and chronic airway obstruction. Family history included a malignant neoplasm of the kidney.
BRAT0703	PSORT11	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 17-year-old Caucasian female during excision of a cerebral meningioma lesion. Pathology indicated a grade 4 fibrillary giant and small-cell astrocytoma. Family history included benign hypertension and cerebrovascular disease.
BRAT0702	PSORT11	Library was constructed using RNA isolated from diseased breast tissue removed from a 55-year-old Caucasian female during a unilateral axillary extended simple mastectomy. Pathology indicated invasive grade 4 mammary adenocarcinoma of mixed lobular and ductal type, extensively involving the chest wall and associated with ipsilateral axillary lymph node metastasis. Family history included a malignant neoplasm of the breast and a low and five high axillary lymph nodes were positive for tumor. Proliferative fibrocystic changes were characterized by apocrine metaplasia, sclerosing adenosis, cyst formation, and ductal hyperplasia without atypia. Patient history included atrial tachycardia, blood in the stool, and a benign breast neoplasm. Family history included benign hypertension, atherosclerotic coronary artery disease, cerebrovascular disease, and depressive disorder.
COLMF002	P1N1	Library was constructed using RNA isolated from the colon tissue of a Caucasian female fetus, who died at 30 weeks gestation.
COLMF011	PSORT11	Library was constructed using RNA isolated from the left hemocolicocolic tissue removed from a Caucasian male.
ENDAN0701	PB1/BSRCP10	Library was constructed using RNA isolated from aortic endothelial cell tissue from an excised heart removed from a male during a heart transplant.

WO 01/83524

PCT/US01/13862

Table 6 (cont.)

Library	Vector	Library Description
ENDCR002	pINCY	This normalized coronary artery endothelial cell tissue library was constructed from 444,000 independent clones from an endothelial tissue library. Starting RNA was made from coronary artery endothelial cell tissue removed from a 3-year-old Caucasian male. This library was normalized in two rounds using conditions adapted from Soares et al., <i>PNAS</i> (1994) 91:9228-9232 and Amado et al., <i>Genome Research</i> 1996; 6:74-80, using a significantly longer (48 hours/round) reannealing hybridization protocol.
EPIT00005	pINCY	This normalized prostate epithelial cell tissue library was constructed from 2.36 million independent clones from a prostate epithelial cell tissue library. Starting RNA was made from untreated prostatic epithelial cell tissue removed from a 17-year-old Hispanic male. The library was normalized in two rounds using conditions adapted from Soares et al., <i>PNAS</i> (1994) 91:9228 and Ronaldo et al., <i>Genome Research</i> (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
HT2RAT01	PRUESCRIPT	Library was constructed at Stratagene (SRR37231), using RNA isolated from the HT2 cell line derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor. Cells were treated with retinoic acid for 24 hours.
ISLET00001	pINCY	Library was constructed using RNA isolated from a pooled collection of pancreatic islet cells.
KERANT001	PRUESCRIPT	Library was constructed using RNA isolated from neonatal keratinocytes obtained from the leg skin of a spontaneously aborted black male.
KIDN00005	PSP0001	Library was constructed using RNA isolated from the kidney tissue of a 2-day-old Hispanic female, who died from cerebral anoxia. Family history included congenital heart disease.
KIDN00009	pINCY	Library was constructed using RNA isolated from the kidney tissue of a Caucasian male fetus, who died at 23 weeks' gestation.
KIDN00001	PSP0001	Library was constructed using RNA isolated from the kidney tumor tissue removed from an 8-year-old female child who died from a renal tumor. Pathology indicated Wilms' tumor (nephroblastoma) which involved 90 percent of the renal parenchyma. Prior to surgery, the patient was receiving heparin anticoagulant therapy.

Table 6 (cont.)

Library	Vector	Library Description
LIVR00008	pINCY	This normalized library was constructed from 5.7 million independent clones from a pooled liver tumor library. Sixteen RNUs were identified from pooled liver tumor tissue from 4-year-old Caucasian females who died of hepatocellular carcinoma. All 16 female fetuses who died after 16-weeks gestation from amniocentesis. Serologies were positive for cytomegalovirus in the 4-year-old. Patient history included asthma in the 4-year-old. Family history included taking daily prenatal vitamins and nitral valve prolapse in the mother of the fetus. The library was normalized in 2 rounds using conditions adapted from Soares et al., <i>NMRS</i> (1994) 91:9228 and Bernaldo et al., <i>Genome Research</i> 6 (1996) 791, except that a significantly longer (48 hours/round) denaturing hybridization was used.
LUNG00003	pINCY	Library was constructed using RNA isolated from lymph node tissue obtained from a 67-year-old Caucasian male during a segmental lung resection. Pathology for the lymph node tissue indicated metastatic bronchogenic carcinoma. Tumor pathology for the associated tumor tissue indicated invasive grade 3-4 squamous cell carcinomas. Patient history included hemangioma. Family history included atherosclerotic coronary artery disease, benign hypertension, congestive heart failure, atherosclerotic coronary artery disease.
LUNG00004	pINCY	Library was constructed using RNA isolated from lung tissue removed from the left lower lobe of a 47-year-old Caucasian male during a segmental lung resection. Pathology for the associated tumor tissue indicated a grade 4 adenocarcinoma, and the parenchyma showed calcified granuloma. Patient history included benign hypertension and chronic obstructive pulmonary disease. Family history included type 2 diabetes mellitus, myocardial infarction.
LUNG00020	pINCY	Library was constructed using RNA isolated from right upper lobe lung tissue removed from a 61-year-old Caucasian male. Pathology indicated bronchial emphysema with blebs in the right anterior upper lobe and apex, as well as emphysema in the right posterior upper lobe. Patient history included angina pectoris, and gastric ulcer. Family history included a subdural hemorrhage, cancer of an unidentified site, atherosclerotic coronary artery disease, and pneumonia.
LUNG00027	pINCY	Library was constructed using RNA isolated from lung tissue removed from a 17-year-old Hispanic female.
LUNG00003	PSF0001	Library was constructed using RNA isolated from lung tumor tissue removed from the left lower lobe of a 65-year-old Caucasian male during segmental lung resection. Pathology indicated a metastatic grade 3 invasive squamous cell carcinoma. Patient history included myocardial infarction, prostatectomy, hypoparathyroidism, malignant skin melanoma, and tobacco use.

WO 01/83524

PCT/US01/13862

Table 6 (cont.)

Library	Vector	Library Description
LUNGSTUCS	pLNCY	Library was constructed using RNA isolated from lung tumor tissue removed from a 69-year-old Caucasian male during segmental lung resection. Pathology indicated invasive grade 3 squamous cell carcinoma and a metastatic tumor. Patient history included type II diabetes, thyroid disorder, depressive disorder, hyperlipidemia, esophageal ulcer, and tobacco use.
NEURITUC03	pLNCY	Library was constructed using RNA isolated from brain meningioma tissue removed from a 35-year-old Caucasian female during craniotomy and excision of the meningioma. Pathology indicated meningioma, grade 1. Patient history included hypertension, hyperlipidemia, brain. Patient history included hypothyroidism. Family history included myocardial infarction and breast cancer.
OVAR30703	PSPORT1	Library was constructed using RNA isolated from ovarian tissue removed from a 43-year-old Caucasian female during removal of the fallopian tubes and ovaries. Pathology for the associated tumor tissue indicated grade 2 mucinous cystadenocarcinoma. Patient history included mitral valve disorder, pneumonia, and viral hepatitis. Family history included atherosclerotic coronary artery disease, pancreatic cancer, stress reaction, cerebrovascular disease, breast cancer, and uterine cancer.
OVARCTU01	PSPORT1	Library was constructed using RNA isolated from ovarian tumor tissue removed from a 43-year-old Caucasian female during removal of the fallopian tubes and ovaries. Pathology indicated grade 2 mucinous cystadenocarcinoma involving the entire left ovary. Patient history included mitral valve disorder, pneumonia, and viral hepatitis. Family history included atherosclerotic coronary artery disease, pancreatic cancer, stress reaction, cerebrovascular disease, breast cancer, and uterine cancer.
PANCR0707	pLNCY	Library was constructed using RNA isolated from the pancreatic tissue of a Caucasian male fetus, who died at 23 weeks' gestation.
PANCR0708	pLNCY	Library was constructed using RNA isolated from pancreatic tissue removed from a 65-year-old Caucasian female during radical subtotal pancreatectomy. Pathology for the associated tumor tissue indicated grade 2 adenocarcinoma. Patient history included type II diabetes, osteoarthritis, cardiovascular disease, benign neoplasm in the large bowel, and a cataract. Previous surgeries included a total splenectomy, cholecystectomy, and abdominal hysterectomy. Family history included cardiovascular disease, type II diabetes, and stomach cancer.
PITUIT0101	PBUESCRIPT	Library was constructed using RNA obtained from Clontech (Clon 5594-2, lot 35278). The RNA was isolated from the pituitary glands removed from a pool of 18 male and female Caucasian donors, 15 to 70 years old, who died from trauma.

WO 01/83524

PCT/US01/13862

Table 6 (cont.)

Library	Vector	Library Description
PROSR0716	PINCY	Library was constructed using RNA isolated from diseased prostate tissue removed from a 68-year-old Caucasian male during a radical prostatectomy. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 3-4). The patient presented with elevated prostate specific antigen (PSA). During this hospitalization, the patient was diagnosed with myelodysplastic syndrome. History included osteoarthritis and hypertension. The patient died of a myocardial infarction. Library was constructed using RNA isolated from small intestine tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation.
SELANB0701	PBUESCRIPT	Library was constructed at Stratagene, using RNA isolated from a pool of fetal spleen tissue. Following vector packaging, 2 million primary clones were then amplified to stabilize the library for long-term storage. Amplification may significantly skew sequence abundances.
SELAN0704	PINCY	Library was constructed using RNA isolated from the spleen tissue of a 2-year-old Hispanic male, who died from cerebral anoxia. Past medical history and serologies were negative.
HEPIND0703	PINCY	Library was constructed using 1 microgram of polyA RNA isolated from untreated PH2-1 cells. PH2-1 (ATCC HTB 221) is a human promonocyte line derived from the peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (ref: Int. J. Cancer (1983) 26:171).
TLVANG0702	PBUESCRIPT	Library was constructed using RNA isolated from non-adherent peripheral blood mononuclear cells. The blood was obtained from unrelated male and female donors. Cells from each donor were purified on Ficoll Hypaque, then harvested by centrifugation, lysed in a buffer containing guanidyl, and spun through CsCl to obtain RNA for library construction.
UCMCL5701	PBUESCRIPT	Library was constructed using RNA isolated from mononuclear cells obtained from the umbilical cord blood of 12 individuals. The cells were cultured for 12 days with IL-3 and IL-6. The library was constructed using RNA isolated from the uterine tissue of a 45-year-old Caucasian female during a total abdominal hysterectomy and total colectomy. Pathology for the associated tumor tissue indicated multiple leiomyomas of the myometrium and a grade 2 colonic adenocarcinoma of the cecum. Patient history included multiple sclerosis and mitral valve disorder. Family history included type 1 diabetes, cerebrovascular disease, atherosclerotic coronary artery disease, malignant skin neoplasm, hypertension, and malignant neoplasm of the colon.
UTEREC0705	PINCY	

WO 01/83524

PCT/US01/13862

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Perceel 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, blastn, and blastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability values= 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, fastx, fastn, blastx, and search.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 185:63-68; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta B value=1.0E-6 Assembled ESTs: fasta identity= 95% or greater and Match length=200 bases or greater; fasta B value=1.0E-8 or less Full Length sequences: fasta score=100 or greater
BLIMPS	A BLASTs IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DDB/PRODOM, and PIRAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1990) Methods Enzymol. 206:88-105; and Alwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1993) Our World View in a Numbell, Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-5 or less Signal peptide hits: Score= 0 or greater

WO 01/83524

PCT/US01/13862

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Grisham, M. et al. (1988) CABIOS 4:61-66; Grisham, M. et al. (1989) Methods Enzymol. 183:146-159; Baird, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores (OCG-specified "HIGT" 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.	
Phrap	A Phred Revised Assembly Program including SWAT and CrossMatch programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Scores 120 or greater; Match length= 56 or greater
Coned	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielsen, H. et al. (1997) Protein Engineering 10:1-6; Clavette, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMEMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sankhamner, 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.	
Moifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Baird, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 01/83524

PCT/US01/13862

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence selected from the group consisting of
 - 5 SEQ ID NO:1-47,
 - b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-47,
 - c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, and
 - 10 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-47.
- 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 selected from the group consisting of SEQ ID
- 20 NO:48-94.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
- 25 7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method for producing a polypeptide of claim 1, the method comprising:
 - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said
 - 30 cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
 - b) recovering the polypeptide so expressed.
- 35 10. An isolated antibody which specifically binds to a polypeptide of claim 1.

WO 01/83524

PCT/US01/13862

11. An isolated polynucleotide selected from the group consisting of:
- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:48-94,
 - b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:48-94,
 - c) a polynucleotide complementary to a polynucleotide of a),
 - d) a polynucleotide complementary to a polynucleotide of b), and
 - e) an RNA equivalent of a)-d).
12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.
13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
 - b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.
15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- 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.
16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.
17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-47.

WO 01/83524

PCT/US01/13862

18. A method for treating a disease or condition associated with decreased expression of functional RMEP, comprising administering to a patient in need of such treatment the composition of claim 16.

5 19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

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

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

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

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

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

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

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

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

- 30 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.

35 26. A method of screening for a compound that modulates the activity of the polypeptide of

WO 01/83524

PCT/US01/13862

claim 1, said method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- 5 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.

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

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- 15 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.

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

- a) treating a biological sample containing nucleic acids with the test compound;
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim
- 25 11 or fragment thereof;
- c) quantifying the amount of hybridization complex; and
- d) comparing the amount of hybridization complex in the treated biological sample with the 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
- 30 compound.

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

- a) combining the biological sample with an antibody of claim 10, under conditions suitable

WO 01/83524

PCT/US01/13862

for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and

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

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

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

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

32. A method of diagnosing a condition or disease associated with the expression of RMEP
15 in a subject, comprising administering to said subject an effective amount of the composition of claim 31.

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

20 34. A method of diagnosing a condition or disease associated with the expression of RMEP in a subject, comprising administering to said subject an effective amount of the composition of claim 33.

35. A method of preparing a polyclonal antibody with the specificity of the antibody of claim
25 10 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibodies from said animal; and

30 c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47.

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

WO 01/83524

PCT/US01/13862

37. A composition comprising the antibody of claim 36 and a suitable carrier.
38. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising:
- 5 a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47, 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-
 - 10 producing hybridoma cells;
 - d) culturing the hybridoma cells; and
 - e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47.
- 15 39. A monoclonal antibody produced by a method of claim 38.
40. A composition comprising the antibody of claim 39 and a suitable carrier.
41. The antibody of claim 10, wherein the antibody is produced by screening a Fab
- 20 expression library.
42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant immunoglobulin library.
- 25 43. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47 in a sample, comprising the steps of:
- a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
 - b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide
 - 30 having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47 in the sample.
44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47 from a sample, the method comprising:
- a) incubating the antibody of claim 10 with a sample under conditions to allow specific

WO 01/83524

PCT/US01/13862

binding of the antibody and the polypeptide; and

b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-47.

- 5 45. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.
46. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.
47. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.
- 10 48. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.
49. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.
- 15 50. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.
51. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.
52. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.
- 20 53. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.
54. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.
- 25 55. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.
56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
- 30 58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

WO 01/83524

PCT/US01/13862

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.
61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.
- 5 62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.
63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.
64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.
- 10 65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.
66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.
67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.
- 15 68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.
69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.
- 20 70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.
71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.
- 25 72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.
73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.
74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.
- 30 75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.
76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.

WO 01/83524

PCT/US01/13862

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:33.
78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:34.
- 5 79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:35.
80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:36.
81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:37.
- 10 82. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:38.
83. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:39.
84. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:40.
- 15 85. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:41.
86. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:42.
- 20 87. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:43.
88. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:44.
- 25 89. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:45.
90. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:46.
91. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:47.
- 30 92. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:48.
93. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:49.

WO 01/83524

PCT/US01/13862

94. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:50.
95. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:51.
- 5 96. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:52.
97. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:53.
98. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:54.
- 10 99. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:55.
100. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:56.
- 15 101. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:57.
102. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
20 NO:58.
103. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:59.
- 25 104. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:60.
105. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:61.
- 30 106. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:62.
107. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID

WO 01/83524

PCT/US01/13862

NO:63.

108. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:64.

5

109. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:65.

110. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
10 NO:66.

111. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:67.

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

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

20

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

115. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
25 NO:71.

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

117. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
30 NO:73.

118. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:74.

WO 01/83524

PCT/US01/13862

119. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:75.

120. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
5 NO:76.

121. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:77.

10 122. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:78.

123. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:79.

15 124. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:80.

125. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
20 NO:81.

126. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:82.

25 127. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:83.

128. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:84.

30 129. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:85.

130. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID

WO 01/83524

PCT/US01/13862

NO:86.

131. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:87.

5

132. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:88.

133. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
10 NO:89.

134. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:90.

15 135. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:91.

136. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:92.

20

137. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:93.

138. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
25 NO:94.

WO 01/83524

PCT/US01/13862

<110> INCYTE GENOMICS, INC.

LAL, Preeti

YUB, Henry

TANG, Y. Tom

LU, Dying Anna M.

AZIMZAI, Yalda

AU-YOUNG, Janice

HILLMAN, Jennifer L.

BAUGHN, Mariah R.

YAO, Monique G.

BIRFORD, Neil

BATHA, Sajeev

POLJCKY, Jennifer J.

<120> RNA METABOLISM PROTEINS

<130> PF-0771 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/200,184; 60/201,875; 60/202,090; 60/210,232; 60/220,553

<151> 2000-04-28; 2000-05-04; 2000-05-04; 2000-06-06; 2000-07-28

<160> 94

<170> ESRL Program

<210> 1

<211> 245

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1622129CD1

<400> 1

```

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

```

WO 01/83524

PCT/US01/13862

Thr Leu Leu Asp Arg Val Val Arg Gln His Val Arg Glu Ala Ser
 230 240
 Ile Leu Leu Gly Asp
 245

<210> 2
 <211> 118
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1820078CD1

<400> 2
 Met Thr Asp Thr Ala Glu Ala Val Pro Lys Phe Glu Glu Met Phe
 1 5 10 15
 Ala Ser Arg Phe Thr Glu Asn Asp Lys Glu Tyr Gln Glu Tyr Leu
 20 25 30
 Lys Arg Pro Pro Glu Ser Pro Pro Ile Val Glu Glu Trp Asn Ser
 35 40 45
 Arg Ala Gly Gly Asn Gln Arg Asn Arg Gly Asn Arg Leu Gln Asp
 50 55 60
 Asn Arg Gln Phe Arg Gly Arg Asp Asn Arg Trp Gly Trp Pro Ser
 65 70 75
 Asp Asn Arg Ser Asn Gln Trp His Gly Arg Ser Trp Gly Asn Asn
 80 85 90
 Tyr Pro Gln His Arg Gln Glu Pro Tyr Tyr Pro Gln Gln Tyr Gly
 95 100 105
 His Tyr Gly Tyr Asn Gln Arg Pro Pro Tyr Gly Tyr Tyr
 110 115

<210> 3
 <211> 179
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1527017CD1

<400> 3
 Met Phe Gly Ser Ser Arg Arg Leu Ser Ser Ser Lys Leu Leu Gln
 1 5 10 15
 Gln Gly Lys Thr Ser Ser Val Phe Glu Asp Pro Val Ile Ser Lys
 20 25 30
 Phe Thr Asn Met Met Met Ile Gly Gly Asn Lys Val Leu Ala Arg
 35 40 45
 Ser Leu Met Ile Gln Thr Leu Glu Ala Val Lys Arg Lys Gln Phe
 50 55 60
 Glu Lys Tyr His Ala Ser Ala Glu Glu Gln Ala Thr Ile Glu
 65 70 75
 Arg Asn Pro Tyr Thr Ile Phe His Gln Ala Leu Lys Asn Cys Glu
 80 85 90
 Pro Met Ile Gly Leu Val Pro Ile Leu Lys Gly Gly Arg Phe Tyr
 95 100 105
 Gln Val Pro Val Pro Leu Pro Asp Arg Arg Arg Arg Phe Leu Ala
 110 115 120
 Met Lys Trp Met Ile Thr Glu Cys Arg Asp Lys Lys His Gln Arg
 125 130 135
 Thr Leu Met Pro Glu Lys Leu Ser His Lys Leu Leu Glu Ala Phe
 140 145 150
 His Asn Gln Gly Pro Val Ile Lys Arg Lys His Asp Leu His Lys
 155 160 165
 Met Ala Glu Ala Asn Arg Ala Leu Ala His Tyr Arg Trp Trp
 170 175

WO 01/83524

PCT/US01/13862

<210> 4
 <211> 101
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1647264CD1

<400> 4
 Met Glu Arg Pro Asp Lys Ala Ala Leu Asn Ala Leu Gln Pro Pro
 1 5 10 15
 Glu Phe Arg Asn Glu Ser Ser Leu Ala Ser Thr Leu Lys Thr Leu
 20 25 30
 Leu Phe Phe Thr Ala Leu Met Ile Thr Val Pro Ile Gly Leu Tyr
 35 40 45
 Phe Thr Thr Lys Ser Tyr Ile Phe Glu Gly Ala Leu Gly Met Ser
 50 55 60
 Asn Arg Asp Ser Tyr Phe Tyr Ala Ala Ile Val Ala Val Val Ala
 65 70 75
 Val His Val Val Leu Ala Leu Phe Val Tyr Val Ala Trp Asn Glu
 80 85 90
 Gly Ser Arg Gln Trp Arg Glu Gly Lys Gln Asp
 95 100

<210> 5
 <211> 145
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1721989CD1

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

<210> 6
 <211> 249
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1730581CD1

<400> 6
 Met Ala Ala Gln Ser Ala Pro Lys Val Val Leu Lys Ser Thr Thr

WO 01/83524

PCT/US01/13862

```

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

```

<210> 7

<211> 265

<212> PRT

<213> Homo sapiens

<220>

<221> misc feature

<223> Incyte ID No: 1740714CD1

<400> 7

```

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

```

WO 01/83524

PCT/US01/13862

```

170      175      180
Glu Tyr Leu Ser Asn Trp Phe Gly Glu Arg Asp Phe Thr Pro Glu
185      190      195
Leu Gly Arg Trp Leu Tyr Ala Leu Leu Ala Cys Leu Glu Lys Pro
200      205      210
Leu Leu Pro Glu Ala His Ser Leu Ile Arg Gln Leu Ala Arg Arg
215      220      225
Cys Ser Glu Val Arg Leu Leu Val Asp Ser Lys Asp Asp Glu Arg
230      235      240
Val Pro Ala Leu Asn Leu Leu Ile Cys Leu Val Ser Arg Tyr Phe
245      250      255
Asp Gln Arg Asp Leu Ala Asp Glu Pro Ser
260      265

```

```

<210> 8
<211> 306
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1850596CD1

```

```

<400> 8
Met Ser Leu Lys Leu Gln Ala Ser Asn Val Thr Asn Lys Asn Asp
1      5      10      15
Pro Lys Ser Ile Asn Ser Arg Val Phe Ile Gly Asn Leu Asn Thr
20      25      30
Ala Leu Val Lys Lys Ser Asp Val Glu Thr Ile Phe Ser Lys Tyr
35      40      45
Gly Arg Val Ala Gly Cys Ser Val His Lys Gly Tyr Ala Phe Val
50      55      60
Gln Tyr Ser Asn Glu Arg His Ala Arg Ala Ala Val Leu Gly Glu
65      70      75
Asn Gly Arg Val Leu Ala Gly Gln Thr Leu Asp Ile Asn Met Ala
80      85      90
Gly Glu Pro Lys Pro Asp Arg Pro Lys Gly Leu Lys Arg Ala Ala
95      100      105
Ser Ala Ile Tyr Ser Gly Tyr Ile Phe Asp Tyr Asp Tyr Tyr Arg
110      115      120
Asp Asp Phe Tyr Asp Arg Leu Phe Asp Tyr Arg Gly Arg Leu Ser
125      130      135
Pro Val Pro Val Pro Arg Ala Val Pro Val Lys Arg Pro Arg Val
140      145      150
Thr Val Pro Leu Val Arg Arg Val Lys Thr Asn Val Pro Val Lys
155      160      165
Leu Phe Ala Arg Ser Thr Ala Val Thr Thr Ser Ser Ala Lys Ile
170      175      180
Lys Leu Lys Ser Ser Glu Leu Gln Ala Ile Lys Thr Glu Leu Thr
185      190      195
Gln Ile Lys Ser Asn Ile Asp Ala Leu Leu Ser Arg Leu Glu Gln
200      205      210
Ile Ala Ala Glu Gln Lys Ala Asn Pro Asp Gly Lys Lys Lys Gly
215      220      225
Asp Gly Gly Gly Ala Gly Gly Gly Gly Gly Gly Gly Ser Gly
230      235      240
Gly Gly Gly Ser Gly Gly Gly Gly Gly Gly Gly Ser Ser Arg Pro
245      250      255
Pro Ala Pro Gln Glu Asn Thr Thr Ser Glu Ala Gly Leu Pro Gln
260      265      270
Gly Glu Ala Arg Thr Arg Asp Asp Gly Asp Glu Glu Gly Leu Leu
275      280      285
Thr His Ser Glu Glu Glu Leu Glu His Ser Gln Asp Thr Asp Ala
290      295      300
Asp Asp Gly Ala Leu Gln
305

```

WO 01/83524

PCT/US01/13862

<210> 9
 <211> 332
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1856109CD1

<400> 9
 Met Ala Ser Gly Leu Val Arg Leu Leu Gln Gln Gly His Arg Cys
 1 5 10 15
 Leu Leu Ala Pro Val Ala Pro Lys Leu Val Pro Pro Val Arg Gly
 20 25 30
 Val Lys Lys Gly Phe Arg Ala Ala Phe Arg Phe Gln Lys Glu Leu
 35 40 45
 Glu Arg Gln Arg Leu Leu Arg Cys Pro Pro Pro Pro Val Arg Arg
 50 55 60
 Ser Glu Lys Pro Asn Trp Asp Tyr His Ala Glu Ile Gln Ala Phe
 65 70 75
 Gly His Arg Leu Gln Glu Asn Phe Ser Leu Asp Leu Leu Lys Thr
 80 85 90
 Ala Phe Val Asn Ser Cys Tyr Ile Lys Ser Glu Glu Ala Lys Arg
 95 100 105
 Gln Gln Leu Gly Ile Glu Lys Glu Ala Val Leu Leu Asn Leu Lys
 110 115 120
 Ser Asn Gln Glu Leu Ser Glu Gln Gly Thr Ser Phe Ser Gln Thr
 125 130 135
 Cys Leu Thr Gln Phe Leu Glu Asp Glu Tyr Pro Asp Met Pro Thr
 140 145 150
 Glu Gly Ile Lys Asn Leu Val Asp Phe Leu Thr Gly Glu Glu Val
 155 160 165
 Val Cys His Val Ala Arg Asn Leu Ala Val Glu Gln Leu Thr Leu
 170 175 180
 Ser Glu Glu Phe Pro Val Pro Pro Ala Val Leu Gln Gln Thr Phe
 185 190 195
 Phe Ala Val Ile Gly Ala Leu Leu Gln Ser Ser Gly Pro Glu Arg
 200 205 210
 Thr Ala Leu Phe Ile Arg Asp Phe Leu Ile Thr Gln Met Thr Gly
 215 220 225
 Lys Glu Leu Phe Glu Met Trp Lys Ile Ile Asn Pro Met Gly Leu
 230 235 240
 Leu Val Glu Glu Leu Lys Lys Arg Asn Val Ser Ala Pro Glu Ser
 245 250 255
 Arg Leu Thr Arg Gln Ser Gly Gly Thr Thr Ala Leu Pro Leu Tyr
 260 265 270
 Phe Val Gly Leu Tyr Cys Asp Lys Lys Leu Ile Ala Glu Gly Pro
 275 280 285
 Gly Glu Thr Val Leu Val Ala Glu Glu Glu Ala Ala Arg Val Ala
 290 295 300
 Leu Arg Lys Leu Tyr Gly Phe Thr Glu Asn Arg Arg Pro Trp Asn
 305 310 315
 Tyr Ser Lys Pro Lys Glu Thr Leu Arg Ala Glu Lys Ser Ile Thr
 320 325 330
 Ala Ser

<210> 10
 <211> 279
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1921719CD1

<400> 10

WO 01/83524

PCT/US01/13862

```

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

```

<210> 11
 <211> 239
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2099829CD1

```

<400> 11
Met Pro Lys Ser Lys Arg Asp Lys Lys Val Ser Leu Thr Lys Thr
1      5      10      15
Ala Lys Lys Gly Leu Glu Leu Lys Gln Asn Leu Ile Glu Glu Leu
20     25     30
Arg Lys Cys Val Asp Thr Tyr Lys Tyr Leu Phe Ile Phe Ser Val
35     40     45
Ala Asn Met Arg Asn Ser Lys Leu Lys Asp Ile Arg Asn Ala Trp
50     55     60
Lys His Ser Arg Met Phe Phe Gly Lys Asn Lys Val Met Met Val
65     70     75
Ala Leu Gly Arg Ser Pro Ser Asp Glu Tyr Lys Asp Asn Leu His
80     85     90
Gln Val Ser Lys Arg Leu Arg Gly Glu Val Gly Leu Leu Phe Thr
95     100    105
Asn Arg Thr Lys Glu Glu Val Asn Glu Trp Phe Thr Lys Tyr Thr
110    115    120
Glu Met Asp Tyr Ala Arg Ala Gly Asn Lys Ala Ala Phe Thr Val
125    130    135

```

WO 01/83524

PCT/US01/13862

```

Ser Leu Asp Pro Gly Pro Leu Glu Gln Phe Pro His Ser Met Glu
140 145 150
Pro Gln Leu Arg Gln Leu Gly Leu Pro Thr Ala Leu Lys Arg Gly
155 160 165
Val Val Thr Leu Leu Ser Asp Tyr Glu Val Cys Lys Glu Gly Asp
170 175 180
Val Leu Thr Pro Glu Gln Ala Arg Val Leu Lys Leu Phe Gly Tyr
185 190 195
Glu Met Ala Glu Phe Lys Val Thr Ile Lys Tyr Met Trp Asp Ser
200 205 210
Gln Ser Gly Arg Phe Gln Gln Met Gly Asp Asp Leu Pro Glu Ser
215 220 225
Ala Ser Glu Ser Thr Glu Glu Ser Asp Ser Glu Asp Asp Asp
230 235

```

<210> 12

<211> 291

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2416915CD1

<400> 12

```

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

```

<210> 13

WO 01/83524

PCT/US01/13862

<211> 451
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2472784CD1

<400> 13
 Met Ala Gly Ala Gly Pro Ala Pro Gly Leu Pro Gly Ala Gly Gly
 1 5 10 15
 Pro Val Val Pro Gly Pro Gly Ala Gly Ile Pro Gly Lys Ser Gly
 20 25 30
 Glu Glu Arg Leu Lys Glu Met Glu Ala Glu Met Ala Leu Phe Glu
 35 40 45
 Gln Glu Val Leu Gly Ala Pro Val Pro Gly Ile Pro Thr Ala Val
 50 55 60
 Pro Ala Val Pro Thr Val Pro Thr Val Pro Thr Val Glu Ala Met
 65 70 75
 Gln Val Pro Ala Ala Pro Val Ile Arg Pro Ile Ile Ala Thr Asn
 80 85 90
 Thr Tyr Gln Gln Val Gln Gln Thr Leu Glu Ala Arg Ala Ala Ala
 95 100 105
 Ala Ala Thr Val Val Pro Pro Met Val Gly Gly Pro Pro Phe Val
 110 115 120
 Gly Pro Val Gly Phe Gly Pro Gly Asp Arg Ser His Leu Asp Ser
 125 130 135
 Pro Glu Ala Arg Glu Ala Met Phe Leu Arg Arg Ala Ala Ala Val
 140 145 150
 Pro Arg Pro Met Ala Leu Pro Pro Pro His Gln Ala Leu Val Gly
 155 160 165
 Pro Pro Leu Pro Gly Pro Pro Gly Pro Pro Met Met Leu Pro Pro
 170 175 180
 Met Ala Arg Ala Pro Gly Pro Pro Leu Gly Ser Met Ala Ala Leu
 185 190 195
 Arg Pro Pro Leu Glu Glu Pro Ala Ala Pro Arg Glu Leu Gly Leu
 200 205 210
 Gly Leu Gly Leu Gly Leu Lys Glu Lys Glu Glu Ala Val Val Ala
 215 220 225
 Ala Ala Ala Gly Leu Glu Glu Ala Ser Ala Ala Val Ala Val Gly
 230 235 240
 Ala Gly Gly Ala Pro Ala Gly Pro Ala Val Ile Gly Pro Ser Leu
 245 250 255
 Pro Leu Ala Leu Ala Met Pro Leu Pro Glu Pro Glu Pro Leu Pro
 260 265 270
 Leu Pro Leu Glu Val Val Arg Gly Leu Leu Pro Pro Leu Arg Ile
 275 280 285
 Pro Glu Leu Leu Ser Leu Arg Pro Arg Pro Arg Pro Pro Arg Pro
 290 295 300
 Glu Pro Pro Pro Gly Leu Met Ala Leu Glu Val Pro Glu Pro Leu
 305 310 315
 Gly Glu Asp Lys Lys Lys Gly Lys Pro Glu Lys Leu Lys Arg Cys
 320 325 330
 Ile Arg Thr Ala Ala Gly Ser Ser Trp Glu Asp Pro Ser Leu Leu
 335 340 345
 Glu Trp Asp Ala Asp Asp Phe Arg Ile Phe Cys Gly Asp Leu Gly
 350 355 360
 Asn Glu Val Asn Asp Asp Ile Leu Ala Arg Ala Phe Ser Arg Phe
 365 370 375
 Pro Ser Phe Leu Lys Ala Lys Val Ile Arg Asp Lys Arg Thr Gly
 380 385 390
 Lys Thr Lys Gly Tyr Gly Phe Val Ser Phe Lys Asp Pro Ser Asp
 395 400 405
 Tyr Val Arg Ala Met Arg Glu Met Asn Gly Lys Tyr Val Gly Ser
 410 415 420
 Arg Pro Ile Lys Leu Arg Lys Ser Met Trp Lys Asp Arg Asn Leu
 425 430 435

WO 01/83524

PCT/US01/13862

Asp Val Val Arg Lys Lys Gln Lys Glu Lys Lys Lys Leu Gly Leu
 440 445 450
 Arg

<210> 14
 <211> 600
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2598981CD1

<400> 14
 Met Pro Glu Ile Arg Val Thr Pro Leu Gly Ala Gly Gln Asp Val
 1 5 10 15
 Gly Arg Ser Cys Ile Leu Val Ser Ile Ala Gly Lys Asn Val Met
 20 25 30
 Leu Asp Cys Gly Met His Met Gly Phe Asn Asp Asp Arg Arg Phe
 35 40 45
 Pro Asp Phe Ser Tyr Ile Thr Gln Asn Gly Arg Leu Thr Asp Phe
 50 55 60
 Leu Asp Cys Val Ile Ile Ser His Phe His Leu Asp His Cys Gly
 65 70 75
 Ala Leu Pro Tyr Phe Ser Glu Met Val Gly Tyr Asp Gly Pro Ile
 80 85 90
 Tyr Met Thr His Pro Thr Gln Ala Ile Cys Pro Ile Leu Leu Glu
 95 100 105
 Asp Tyr Arg Lys Ile Ala Val Asp Lys Lys Gly Glu Ala Asn Phe
 110 115 120
 Phe Thr Ser Gln Met Ile Lys Asp Cys Met Lys Lys Val Val Ala
 125 130 135
 Val His Leu His Gln Thr Val Gln Val Asp Asp Glu Leu Glu Ile
 140 145 150
 Lys Ala Tyr Tyr Ala Gly His Val Leu Gly Ala Ala Met Phe Gln
 155 160 165
 Ile Lys Val Gly Ser Glu Ser Val Val Tyr Thr Gly Asp Tyr Asn
 170 175 180
 Met Thr Pro Asp Arg His Leu Gly Ala Ala Trp Ile Asp Lys Cys
 185 190 195
 Arg Pro Asn Leu Leu Ile Thr Glu Ser Thr Tyr Ala Thr Thr Ile
 200 205 210
 Arg Asp Ser Lys Arg Cys Arg Glu Arg Asp Phe Leu Lys Lys Val
 215 220 225
 His Glu Thr Val Glu Arg Gly Gly Lys Val Leu Ile Pro Val Phe
 230 235 240
 Ala Leu Gly Arg Ala Gln Glu Leu Cys Ile Leu Leu Glu Thr Phe
 245 250 255
 Trp Glu Arg Met Asn Leu Lys Val Pro Ile Tyr Phe Ser Thr Gly
 260 265 270
 Leu Thr Glu Lys Ala Asn His Tyr Tyr Lys Leu Phe Ile Pro Trp
 275 280 285
 Thr Asn Gln Lys Ile Arg Lys Thr Phe Val Gln Arg Asn Met Phe
 290 295 300
 Glu Phe Lys His Ile Lys Ala Phe Asp Arg Ala Phe Ala Asp Asn
 305 310 315
 Pro Gly Pro Met Val Val Phe Ala Thr Pro Gly Met Leu His Ala
 320 325 330
 Gly Gln Ser Leu Gln Ile Phe Arg Lys Trp Ala Gly Asn Glu Lys
 335 340 345
 Asn Met Val Ile Met Pro Gly Tyr Cys Val Gln Gly Thr Val Gly
 350 355 360
 His Lys Ile Leu Ser Gly Gln Arg Lys Leu Glu Met Glu Gly Arg
 365 370 375
 Gln Val Leu Glu Val Lys Met Gln Val Glu Tyr Met Ser Phe Ser
 380 385 390

WO 01/83524

PCT/US01/13862

Ala His Ala Asp Ala Lys Gly Ile Met Gln Leu Val Gly Gln Ala
 395 400 405
 Glu Pro Glu Ser Val Leu Leu Val His Gly Glu Ala Lys Lys Met
 410 415 420
 Glu Phe Leu Lys Gln Lys Ile Glu Gln Glu Leu Arg Val Asn Cys
 425 430 435
 Tyr Met Pro Ala Asn Gly Glu Thr Val Thr Leu Pro Thr Ser Pro
 440 445 450
 Ser Ile Pro Val Gly Ile Ser Leu Gly Leu Leu Lys Arg Glu Met
 455 460 465
 Ala Gln Gly Leu Leu Pro Glu Ala Lys Lys Pro Arg Leu Leu His
 470 475 480
 Gly Thr Leu Ile Met Lys Asp Ser Asn Phe Arg Leu Val Ser Ser
 485 490 495
 Glu Gln Ala Leu Lys Glu Leu Gly Leu Ala Glu His Gln Leu Arg
 500 505 510
 Phe Thr Cys Arg Val His Leu His Asp Thr Arg Lys Glu Gln Glu
 515 520 525
 Thr Ala Leu Arg Val Tyr Ser His Leu Lys Ser Val Leu Lys Asp
 530 535 540
 His Cys Val Gln His Leu Pro Asp Gly Ser Val Thr Val Glu Ser
 545 550 555
 Val Leu Leu Gln Ala Ala Ala Pro Ser Glu Asp Pro Gly Thr Lys
 560 565 570
 Val Leu Leu Val Ser Trp Thr Tyr Gln Asp Glu Glu Leu Gly Ser
 575 580 585
 Phe Leu Thr Ser Leu Leu Lys Lys Gly Leu Pro Gln Ala Pro Ser
 590 595 600

<210> 15

<211> 217

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2738075CD1

<400> 15

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

WO 01/83524 PCT/US01/13862

200 205 210
 Asp Glu Glu Glu Leu Ser Asp
 215

<210> 16
 <211> 319
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2279049CD1

<400> 16
 Met Lys Ile Glu Leu Ser Met Gln Pro Trp Asn Pro Gly Tyr Ser
 1 5 10 15
 Ser Glu Gly Ala Thr Ala Gln Glu Thr Tyr Thr Cys Pro Lys Met
 20 25 30
 Ile Glu Met Glu Gln Ala Glu Ala Gln Leu Ala Glu Leu Asp Leu
 35 40 45
 Leu Ala Ser Met Phe Pro Gly Glu Asn Glu Leu Ile Val Asn Asp
 50 55 60
 Gln Leu Ala Val Ala Glu Leu Lys Asp Cys Ile Glu Lys Lys Thr
 65 70 75
 Met Glu Gly Arg Ser Ser Lys Val Tyr Phe Thr Ile Asn Met Asn
 80 85 90
 Leu Asp Val Ser Asp Glu Lys Met Ala Met Phe Ser Leu Ala Cys
 95 100 105
 Ile Leu Pro Phe Lys Tyr Pro Ala Val Leu Pro Glu Ile Thr Val
 110 115 120
 Arg Ser Val Leu Leu Ser Arg Ser Gln Gln Thr Gln Leu Asn Thr
 125 130 135
 Asp Leu Thr Ala Phe Leu Gln Lys His Cys His Gly Asp Val Cys
 140 145 150
 Ile Leu Asn Ala Thr Glu Trp Val Arg Glu His Ala Ser Gly Tyr
 155 160 165
 Val Ser Arg Asp Thr Ser Ser Ser Pro Thr Thr Gly Ser Thr Val
 170 175 180
 Gln Ser Val Asp Leu Ile Phe Thr Arg Leu Trp Ile Tyr Ser His
 185 190 195
 His Ile Tyr Asn Lys Cys Lys Arg Lys Asn Ile Leu Glu Trp Ala
 200 205 210
 Lys Glu Leu Ser Leu Ser Gly Phe Ser Met Pro Gly Lys Pro Gly
 215 220 225
 Val Val Cys Val Glu Gly Pro Gln Ser Ala Cys Glu Glu Phe Trp
 230 235 240
 Ser Arg Leu Arg Lys Leu Asn Trp Lys Arg Ile Leu Ile Arg His
 245 250 255
 Arg Glu Asp Ile Pro Phe Asp Gly Thr Asn Asp Glu Thr Glu Arg
 260 265 270
 Gln Arg Lys Phe Ser Ile Phe Glu Glu Lys Val Phe Ser Val Asn
 275 280 285
 Gly Ala Arg Gly Asn His Met Asp Phe Gly Gln Leu Tyr Gln Phe
 290 295 300
 Leu Asn Thr Lys Gly Cys Gly Asp Val Phe Gln Met Phe Phe Gly
 305 310 315
 Val Glu Gly Gln

<210> 17
 <211> 108
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2660904CD1

WO 01/83524

PCT/US01/13862

```

<400> 17
Met Ser His His Ala Glu Ile Gln Arg Asp Ile Leu Glu Ser Cys
1      5      10      15
Asn His Val Arg Lys Lys Val Pro Val Thr Phe Val Gly Ala Gly
20     25     30
Gly Gln Asp Pro Glu Val Pro Glu Glu Leu Leu His Leu Leu Gln
35     40     45
Pro Gly Gln Arg Val Pro Gln Asp Val Gln His His Leu Leu Glu
50     55     60
Pro Arg Asp Arg Trp Ala His Leu Glu Val Leu Lys Lys Val Asp
65     70     75
Leu Leu Leu Gln Val Met Ala Ala Thr Gly Tyr Phe His Ala Ser
80     85     90
Leu Gln Arg Gly Glu Ile Met Arg Ser Pro Gly Pro Val Ala Arg
95     100    105
Asn Ser Pro

```

```

<210> 18
<211> 92
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3179424CD1

```

```

<400> 18
Met Ala Val Leu Ala Gly Ser Leu Leu Gly Pro Thr Ser Arg Ser
1      5      10      15
Ala Ala Leu Leu Gly Gly Arg Trp Leu Gln Pro Arg Ala Trp Leu
20     25     30
Gly Phe Pro Asp Ala Trp Gly Leu Pro Thr Pro Gln Gln Ala Arg
35     40     45
Gly Lys Ala Arg Gly Asn Glu Tyr Gln Pro Ser Asn Ile Lys Arg
50     55     60
Lys Asn Lys His Gly Trp Val Arg Arg Leu Ser Thr Pro Ala Gly
65     70     75
Val Gln Val Ile Leu Arg Arg Met Leu Lys Gly Arg Lys Ser Leu
80     85     90
Ser His

```

```

<210> 19
<211> 268
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2885096CD1

```

```

<400> 19
Met Ala Gly Gly Val Pro Gly Gln Pro Ala Gly Val Gly Leu Ala
1      5      10      15
Leu Ile Ala Thr Asp Ser Gln Glu Thr Arg Pro Gly Arg Ala Gly
20     25     30
Pro Gly Ser Gly Glu Ser Leu Ser Ala Ser His Leu Phe Ile Ser
35     40     45
Asp Phe Ala Tyr Cys Trp Glu Asn Phe Val Cys Asn Glu Gly Gln
50     55     60
Pro Phe Met Pro Trp Tyr Lys Phe Asp Asp Asn Tyr Ala Ser Leu
65     70     75
His Arg Thr Leu Lys Glu Ile Leu Arg Asn Pro Met Glu Ala Met
80     85     90
Tyr Pro His Ile Phe Tyr Phe His Phe Lys Asn Leu Leu Lys Ala
95     100    105

```

WO 01/83524

PCT/US01/13862

```

Cys Gly Arg Asn Glu Ser Trp Leu Cys Phe Thr Met Glu Val Thr
110 115 120
Lys His His Ser Ala Val Phe Arg Lys Lys Gly Val Phe Arg Asn
125 130 135
Glu Val Asp Pro Glu Thr His Cys His Ala Glu Arg Cys Phe Leu
140 145 150
Ser Trp Phe Cys Asp Asp Ile Leu Ser Pro Asn Thr Asn Tyr Glu
155 160 165
Val Thr Trp Tyr Thr Ser Trp Ser Pro Cys Pro Glu Cys Ala Gly
170 175 180
Glu Val Ala Glu Phe Leu Ala Arg His Ser Asn Val Asn Leu Thr
185 190 195
Ile Phe Thr Ala Arg Leu Cys Tyr Phe Trp Asp Thr Asp Tyr Glu
200 205 210
Glu Gly Leu Cys Ser Leu Ser Gln Glu Gly Ala Ser Val Lys Ile
215 220 225
Met Gly Tyr Lys Asp Phe Val Ser Cys Trp Lys Asn Phe Val Tyr
230 235 240
Ser Asp Asp Glu Pro Phe Lys Pro Trp Lys Gly Leu Gln Thr Asn
245 250 255
Phe Arg Leu Leu Lys Arg Arg Leu Arg Glu Ile Leu Gln
250 255

```

<210> 20
 <211> 624
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2901076CD1

```

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

```

WO 01/83524

PCT/US01/13862

```

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

```

<210> 21
 <211> 419
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> InCyte ID No: 3074572CD1

```

<400> 21
Met Ala Ala Glu Val Leu Pro Ser Ala Arg Trp Gln Tyr Cys Gly
1      5      10      15
Ala Pro Asp Gly Ser Gln Arg Ala Val Leu Val Gln Phe Ser Asn
20      25      30
Gly Lys Leu Gln Ser Pro Gly Asn Met Arg Phe Thr Leu Tyr Glu
35      40      45

```

WO 01/83524

PCT/US01/13862

```

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

```

<210> 22
<211> 743
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1437895CD1

<400> 22
Met Glu Glu Glu Gly Leu Glu Cys Pro Asn Ser Ser Ser Glu Lys
1 5 10 15
Arg Tyr Phe Pro Glu Ser Leu Asp Ser Ser Asp Gly Asp Glu Glu
20 25 30
Glu Val Leu Ala Cys Glu Asp Leu Glu Leu Asn Pro Phe Asp Gly
35 40 45

WO 01/83524

PCT/US01/13862

Leu Pro Tyr Ser Ser Arg Tyr Tyr Lys Leu Leu Lys Glu Arg Glu
 50 55 60
 Asp Leu Pro Ile Trp Lys Glu Lys Tyr Ser Phe Met Glu Asn Leu
 65 70 75
 Leu Gln Asn Gln His Val Ile Val Ser Gly Asp Ala Lys Cys Gly
 80 85 90
 Lys Ser Ala Gln Val Pro Gln Trp Cys Ala Glu Tyr Cys Leu Ser
 95 100 105
 Ile His Tyr Gln His Gly Gly Val Ile Cys Thr Gln Val His Lys
 110 115 120
 Gln Thr Val Val Gln Leu Ala Leu Arg Val Ala Asp Glu Met Asp
 125 130 135
 Val Asn Ile Gly His Glu Val Gly Tyr Val Ile Pro Phe Glu Asn
 140 145 150
 Cys Cys Thr Asn Glu Thr Ile Leu Arg Tyr Cys Thr Asp Asp Met
 155 160 165
 Leu Gln Arg Glu Met Met Ser Asn Pro Phe Leu Gly Ser Tyr Gly
 170 175 180
 Val Ile Ile Leu Asp Asp Ile His Glu Arg Ser Ile Ala Thr Asp
 185 190 195
 Val Leu Leu Gly Leu Leu Lys Asp Val Leu Ala Arg Pro Glu
 200 205 210
 Leu Lys Leu Ile Ile Asn Ser Ser Pro His Leu Ile Ser Lys Leu
 215 220 225
 Asn Ser Tyr Tyr Gly Asn Val Pro Val Ile Glu Val Lys Asn Lys
 230 235 240
 His Pro Val Glu Val Val Tyr Leu Ser Glu Ala Gln Lys Asp Ser
 245 250 255
 Phe Glu Ser Ile Leu Arg Leu Ile Phe Ile Ile His His Ser Gly
 260 265 270
 Glu Lys Gly Asp Ile Val Val Phe Leu Ala Cys Glu Gln Asp Ile
 275 280 285
 Glu Lys Val Cys Glu Thr Val Tyr Gln Gly Ser Asn Leu Asn Pro
 290 295 300
 Asp Leu Gly Glu Leu Val Val Val Pro Leu Tyr Pro Lys Glu Lys
 305 310 315
 Cys Ser Leu Phe Lys Pro Leu Asp Glu Thr Glu Lys Arg Cys Gln
 320 325 330
 Val Tyr Gln Arg Arg Val Val Leu Thr Thr Ser Ser Gly Glu Phe
 335 340 345
 Leu Ile Trp Ser Asn Ser Val Arg Phe Val Ile Asp Val Gly Val
 350 355 360
 Glu Arg Arg Lys Val Tyr Asn Pro Arg Ile Arg Ala Asn Ser Leu
 365 370 375
 Val Met Gln Pro Ile Ser Gln Ser Gln Ala Glu Ile Arg Lys Gln
 380 385 390
 Ile Leu Gly Ser Ser Ser Ser Gly Lys Phe Phe Cys Leu Tyr Thr
 395 400 405
 Glu Glu Phe Ala Ser Lys Asp Met Thr Pro Leu Lys Pro Ala Glu
 410 415 420
 Met Gln Glu Ala Asn Leu Thr Ser Met Val Leu Phe Met Lys Arg
 425 430 435
 Ile Asp Ile Ala Gly Leu Gly His Cys Asp Phe Met Asn Arg Pro
 440 445 450
 Ala Pro Glu Ser Leu Met Gln Ala Leu Glu Asp Leu Asp Tyr Leu
 455 460 465
 Ala Ala Leu Asp Asn Asp Gly Asn Leu Ser Glu Phe Gly Ile Ile
 470 475 480
 Met Ser Glu Phe Pro Leu Asp Pro Gln Leu Ser Lys Ser Ile Leu
 485 490 495
 Ala Ser Cys Glu Phe Asp Cys Val Asp Glu Val Leu Thr Ile Ala
 500 505 510
 Ala Met Val Thr Ala Pro Asn Cys Phe Ser His Val Pro His Gly
 515 520 525
 Ala Glu Glu Ala Ala Leu Thr Cys Trp Lys Thr Phe Leu His Pro
 530 535 540
 Glu Gly Asp His Phe Thr Leu Ile Ser Ile Tyr Lys Ala Tyr Gln

WO 01/83524

PCT/US01/13862

```

545      550      555
Asp Thr Thr Leu Asn Ser Ser Ser Glu Tyr Cys Val Glu Lys Trp
560      565      570
Cys Arg Asp Tyr Phe Leu Asn Cys Ser Ala Leu Arg Met Ala Asp
575      580      585
Val Ile Arg Ala Glu Leu Leu Glu Ile Ile Lys Arg Ile Glu Leu
590      595      600
Pro Tyr Ala Glu Pro Ala Phe Gly Ser Lys Glu Asn Thr Leu Asn
605      610      615
Ile Lys Lys Ala Leu Leu Ser Gly Tyr Phe Met Gln Ile Ala Arg
620      625      630
Asp Val Asp Gly Ser Gly Asn Tyr Leu Met Leu Thr His Lys Gln
635      640      645
Val Ala Gln Leu His Pro Leu Ser Gly Tyr Ser Ile Thr Lys Lys
650      655      660
Met Pro Glu Trp Val Leu Phe His Lys Phe Ser Ile Ser Glu Asn
665      670      675
Asn Tyr Ile Arg Ile Thr Ser Glu Ile Ser Pro Glu Leu Phe Met
680      685      690
Gln Leu Val Pro Gln Tyr Tyr Phe Ser Asn Leu Pro Pro Ser Glu
695      700      705
Ser Lys Asp Ile Leu Gln Gln Val Val Asp His Leu Ser Pro Val
710      715      720
Ser Thr Met Asn Lys Glu Gln Gln Met Cys Glu Thr Cys Pro Glu
725      730      735
Thr Glu Gln Arg Cys Thr Leu Gln
740

```

<210> 23
 <211> 284
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc.feature
 <223> Incyte ID No: 1454656CD1

```

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

```

WO 01/83524

PCT/US01/13862

215	220	225
Ala His Arg Arg Phe Val Thr Lys Lys	Ala Leu Cys Ile Arg Val	
230	235	240
Phe Gln Glu Thr Gln Lys Leu Lys Lys	Arg Arg Arg Ala Leu Lys	
245	250	255
Ala Ala Ala Ala Gln Lys Gln Ala	Lys Arg Arg Asn Pro Asp	
260	265	270
Ser Pro Ala Lys Ala Ile Pro Lys Thr	Leu Lys Asp Ser Gln	
275	280	

<210> 24
 <211> 248
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 121130CD1

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

<210> 25
 <211> 214
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1257715CD1

<400> 25
 Met Arg Pro Gly Gly Phe Leu Gly Ala Gly Gln Arg Leu Ser Arg
 1 5 10 15

WO 01/83524

PCT/US01/13862

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

<210> 36

<211> 184

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1342022CD1

<400> 36

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

<210> 27

WO 01/83524

PCT/US01/13862

<211> 371
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 194704CD1

<400> 27
 Met Ser Ala Gln Ala Gln Met Arg Ala Leu Leu Asp Gln Leu Met
 1 5 10 15
 Gly Thr Ala Arg Asp Gly Asp Glu Thr Arg Gln Arg Val Lys Phe
 20 25 30
 Thr Asp Asp Arg Val Cys Lys Ser His Leu Leu Asp Cys Cys Pro
 35 40 45
 His Asp Ile Leu Ala Gly Thr Arg Met Asp Leu Gly Glu Cys Thr
 50 55 60
 Lys Ile His Asp Leu Ala Leu Arg Ala Asp Tyr Glu Ile Ala Ser
 65 70 75
 Lys Glu Arg Asp Leu Phe Phe Glu Leu Asp Ala Met Asp His Leu
 80 85 90
 Glu Ser Phe Ile Ala Glu Cys Asp Arg Arg Thr Glu Leu Ala Lys
 95 100 105
 Lys Arg Leu Ala Glu Thr Gln Glu Glu Ile Ser Ala Glu Val Ser
 110 115 120
 Ala Lys Ala Glu Lys Val His Glu Leu Asn Glu Glu Ile Gly Lys
 125 130 135
 Leu Leu Ala Lys Ala Glu Gln Leu Gly Ala Glu Gly Asn Val Asp
 140 145 150
 Glu Ser Gln Lys Ile Leu Met Glu Val Glu Lys Val Arg Ala Lys
 155 160 165
 Lys Lys Glu Ala Glu Glu Glu Tyr Arg Asn Ser Met Pro Ala Ser
 170 175 180
 Ser Phe Gln Gln Gln Lys Leu Arg Val Cys Glu Val Cys Ser Ala
 185 190 195
 Tyr Leu Gly Leu His Asp Asn Asp Arg Arg Leu Ala Asp His Phe
 200 205 210
 Gly Gly Lys Leu His Leu Gly Phe Ile Gln Ile Arg Glu Lys Leu
 215 220 225
 Asp Gln Leu Arg Lys Thr Val Ala Glu Lys Gln Glu Lys Arg Asn
 230 235 240
 Gln Asp Arg Leu Arg Arg Arg Glu Glu Arg Glu Arg Glu Glu Arg
 245 250 255
 Leu Ser Arg Arg Ser Gly Ser Arg Thr Arg Asp Arg Arg Arg Ser
 260 265 270
 Arg Ser Arg Asp Arg Arg Arg Arg Ser Arg Ser Thr Ser Arg
 275 280 285
 Glu Arg Arg Lys Leu Ser Arg Ser Arg Ser Arg Asp Arg His Arg
 290 295 300
 Arg His Arg Ser Arg Ser Arg Ser His Ser Arg Gly His Arg Arg
 305 310 315
 Ala Ser Arg Asp Arg Ser Ala Lys Tyr Lys Phe Ser Arg Glu Arg
 320 325 330
 Ala Ser Arg Glu Glu Ser Trp Glu Ser Gly Arg Ser Glu Arg Gly
 335 340 345
 Pro Pro Asp Trp Arg Leu Glu Ser Ser Asn Gly Lys Met Ala Ser
 350 355 360
 Arg Arg Ser Glu Glu Lys Glu Ala Gly Glu Ile
 365 370

<210> 28
 <211> 396
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> misc_feature

WO 01/83524

PCT/US01/13862

<223> Incyte ID No: 607270CD1

<400> 28

```

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

```

<210> 29

<211> 184

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 758546CD1

WO 01/83524

PCT/US01/13862

```

<400> 29
Met Val Arg Lys Leu Lys Phe His Glu Gln Lys Leu Leu Lys Gln
1      5      10      15
Val Asp Phe Leu Asn Trp Glu Val Thr Asp His Asn Leu His Glu
20     25     30
Leu Arg Val Leu Arg Arg Tyr Arg Leu Gln Arg Arg Glu Asp Tyr
35     40     45
Thr Arg Tyr Asn Gln Leu Ser Arg Ala Val Arg Glu Leu Ala Arg
50     55     60
Arg Leu Arg Asp Leu Pro Glu Arg Asp Gln Phe Arg Val Arg Ala
65     70     75
Ser Ala Ala Leu Leu Asp Lys Leu Tyr Ala Leu Gly Leu Val Pro
80     85     90
Thr Arg Gly Ser Leu Glu Leu Cys Asp Phe Val Thr Ala Ser Ser
95     100    105
Phe Cys Arg Arg Arg Leu Pro Thr Val Leu Leu Lys Leu Arg Met
110    115    120
Ala Gln His Leu Gln Ala Ala Val Ala Phe Val Glu Gln Gly His
125    130    135
Val Arg Val Gly Pro Asp Val Val Thr Asp Pro Ala Phe Leu Val
140    145    150
Thr Arg Ser Met Glu Asp Phe Val Thr Trp Val Asp Ser Ser Lys
155    160    165
Ile Lys Arg His Val Leu Glu Tyr Asn Glu Glu Arg Asp Asp Phe
170    175    180
Asp Leu Glu Ala

```

```

<210> 30
<211> 282
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 866043CD1

```

```

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

```

WO 01/83524

PCT/US01/13862

```

215          220          225
Leu Val Thr Gly Ser Ala Asp Gly Arg Val Leu Met Tyr Ser Phe
230          235          240
Arg Thr Ala Ser Arg Ala Cys Thr Leu Gln Gly His Thr Gln Ala
245          250          255
Cys Val Gly Thr Thr Tyr His Pro Val Leu Pro Ser Val Leu Ala
260          265          270
Thr Cys Ser Trp Gly Gly Asp Met Lys Ile Trp His
275          280

```

<210> 31
 <211> 125
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 827065CD1

```

<400> 31
Met Pro Ala Pro Ala Ala Thr Tyr Glu Arg Val Val Tyr Lys Asn
1      5      10      15
Pro Ser Glu Tyr His Tyr Met Lys Val Cys Leu Glu Phe Gln Asp
20     25     30
Cys Gly Val Gly Leu Asn Ala Ala Gln Phe Lys Gln Leu Leu Ile
35     40     45
Ser Ala Val Lys Asp Leu Phe Gly Glu Val Asp Ala Ala Leu Pro
50     55     60
Leu Asp Ile Leu Thr Tyr Glu Glu Lys Thr Leu Ser Ala Ile Leu
65     70     75
Arg Ile Cys Ser Ser Gly Leu Val Lys Leu Trp Ser Ser Leu Thr
80     85     90
Leu Leu Arg Ile Pro Ile Lys Gly Lys Lys Cys Ala Phe Arg Val
95     100    105
Ile Gln Val Ser Pro Phe Leu Leu Ala Leu Ser Gly Asn Ser Arg
110    115    120
Glu Leu Val Leu Asp
125

```

<210> 32
 <211> 365
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 938071CD1

```

<400> 32
Met Ala Pro Val Ser Gly Ser Arg Ser Pro Asp Arg Glu Ala Ser
1      5      10      15
Gly Ser Gly Gly Arg Arg Ser Ser Ser Lys Ser Pro Lys Pro
20     25     30
Ser Lys Ser Ala Arg Ser Pro Arg Gly Arg Arg Ser Arg Ser His
35     40     45
Ser Cys Ser Arg Ser Gly Asp Arg Asn Gly Leu Thr His Gln Leu
50     55     60
Gly Gly Leu Ser Gln Gly Ser Arg Asn Gln Ser Tyr Arg Ser Arg
65     70     75
Ser Arg Ser Arg Ser Arg Glu Arg Pro Ser Ala Pro Arg Gly Ile
80     85     90
Pro Phe Ala Ser Ala Ser Ser Ser Val Tyr Tyr Gly Ser Tyr Ser
95     100    105
Arg Pro Tyr Gly Ser Asp Lys Pro Trp Pro Ser Leu Leu Asp Lys
110    115    120
Glu Arg Glu Glu Ser Leu Arg Gln Lys Arg Leu Ser Glu Arg Glu
125    130    135

```


WO 01/83524

PCT/US01/13862

```

Arg Ile Gly Glu Leu Gly Ala Pro Glu Val Trp Gly Leu Ser Pro
140 145 150
Lys Asn Pro Glu Pro Asp Ser Asp Glu His Thr Pro Val Glu Asp
155 160 165
Glu Glu Pro Lys Lys Ser Thr Thr Ser Ala Ser Thr Ser Glu Glu
170 175 180
Glu Lys Lys Lys Lys Ser Ser Arg Ser Lys Glu Arg Ser Lys Lys
185 190 195
Arg Arg Lys Lys Lys Ser Ser Lys Arg Lys His Lys Lys Tyr Ser
200 205 210
Glu Asp Ser Asp Ser Asp Ser Asp Ser Glu Thr Asp Ser Ser Asp
215 220 225
Glu Asp Asn Lys Arg Arg Ala Lys Lys Ala Lys Lys Lys Glu Lys
230 235 240
Lys Lys Lys His Arg Ser Lys Lys Tyr Lys Lys Lys Arg Ser Lys
245 250 255
Lys Ser Arg Lys Glu Ser Ser Asp Ser Ser Ser Lys Glu Ser Gln
260 265 270
Glu Glu Phe Leu Glu Asn Pro Trp Lys Asp Arg Thr Lys Ala Glu
275 280 285
Glu Pro Ser Asp Leu Ile Gly Pro Glu Ala Pro Lys Thr Leu Thr
290 295 300
Ser Gln Asp Asp Lys Pro Leu Lys His Arg Arg Met Glu Ala Val
305 310 315
Arg Leu Arg Lys Glu Asn Gln Ile Tyr Ser Ala Asp Glu Lys Arg
320 325 330
Ala Leu Ala Ser Phe Asn Gln Glu Glu Arg Arg Lys Arg Glu Asn
335 340 345
Lys Ile Leu Ala Ser Phe Arg Glu Met Val Tyr Arg Lys Thr Lys
350 355 360
Gly Lys Asp Asp Lys
365

```

<210> 33

<211> 672

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3295984CD1

<400> 33

```

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

```

WO 01/83524

PCT/US01/13862

```

Val Ser Met Lys Asp Pro Pro Asp Leu Leu Asp Arg Gln Lys Cys
185 190 195
Leu Asn Ala Leu Ala Ser Leu Arg His Ala Lys Trp Phe Gln Ala
200 205 210
Arg Ala Asn Gly Leu Lys Ser Cys Val Ile Val Leu Arg Ile Leu
215 220 225
Arg Asp Leu Cys Asn Arg Val Pro Thr Trp Ala Pro Leu Lys Gly
230 235 240
Trp Pro Leu Glu Leu Ile Cys Glu Lys Ser Ile Gly Thr Cys Asn
245 250 255
Arg Pro Leu Gly Ala Gly Glu Ala Leu Arg Arg Val Met Glu Cys
260 265 270
Leu Ala Ser Gly Ile Leu Leu Pro Gly Gly Pro Gly Leu His Asp
275 280 285
Pro Cys Glu Arg Asp Pro Thr Asp Ala Leu Ser Tyr Met Thr Ile
290 295 300
Gln Gln Lys Glu Asp Ile Thr His Ser Ala Gln His Ala Leu Arg
305 310 315
Leu Ser Ala Phe Gly Gln Ile Tyr Lys Val Leu Glu Met Asp Pro
320 325 330
Leu Pro Ser Ser Lys Pro Phe Gln Lys Tyr Ser Trp Ser Val Thr
335 340 345
Asp Lys Glu Gly Ala Gly Ser Ser Ala Leu Lys Arg Pro Phe Glu
350 355 360
Asp Gly Leu Gly Asp Asp Lys Asp Pro Asn Lys Lys Met Lys Arg
365 370 375
Asn Leu Arg Lys Ile Leu Asp Ser Lys Ala Ile Asp Leu Met Asn
380 385 390
Ala Leu Met Arg Leu Asn Gln Ile Arg Gly Leu Gln Tyr Lys
395 400 405
Leu Leu Ser Gln Ser Gly Pro Val His Ala Pro Val Phe Thr Met
410 415 420
Ser Val Asp Val Asp Gly Thr Thr Tyr Glu Ala Ser Gly Pro Ser
425 430 435
Lys Lys Thr Ala Lys Leu His Val Ala Val Lys Val Leu Gln Ala
440 445 450
Met Gly Tyr Pro Thr Gly Phe Asp Ala Asp Ile Glu Cys Met Ser
455 460 465
Ser Asp Glu Lys Ser Asp Asn Glu Ser Lys Asn Glu Thr Val Ser
470 475 480
Ser Asn Ser Ser Asn Asn Thr Gly Asn Ser Thr Thr Glu Thr Ser
485 490 495
Ser Thr Leu Glu Val Arg Thr Gln Gly Pro Ile Leu Thr Ala Ser
500 505 510
Gly Lys Asn Pro Val Met Glu Leu Asn Glu Lys Arg Arg Gly Leu
515 520 525
Lys Tyr Glu Leu Ile Ser Glu Thr Gly Gly Ser His Asp Lys Arg
530 535 540
Phe Val Met Glu Val Glu Val Asp Gly Gln Lys Phe Arg Gly Ala
545 550 555
Gly Pro Asn Lys Lys Val Ala Lys Ala Ser Ala Ala Leu Ala Ala
560 565 570
Leu Glu Lys Leu Phe Ser Gly Pro Asn Ala Ala Asn Asn Lys Lys
575 580 585
Lys Lys Ile Ile Pro Gln Ala Lys Gly Val Val Asn Thr Ala Val
590 595 600
Ser Ala Ala Val Gln Ala Val Arg Gly Arg Gly Arg Gly Thr Leu
605 610 615
Thr Arg Gly Ala Phe Val Gly Ala Thr Ala Ala Pro Gly Tyr Ile
620 625 630
Ala Pro Gly Tyr Gly Thr Pro Tyr Gly Tyr Ser Thr Ala Ala Pro
635 640 645
Ala Tyr Gly Leu Pro Lys Arg Met Val Leu Leu Pro Val Met Lys
650 655 660
Phe Pro Thr Tyr Pro Val Pro His Tyr Ser Phe Phe
665 670

```

WO 01/83524

PCT/US01/13862

<210> 34
 <211> 430
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4545237CD1

<400> 34
 Met Ala Thr Ala Val Arg Ala Val Gly Cys Leu Pro Val Leu Cys
 1 10 15
 Ser Gly Thr Ala Gly His Leu Leu Gly Arg Gln Cys Ser Leu Asn
 20 25 30
 Thr Leu Pro Ala Ala Ser Ile Leu Ala Trp Lys Ser Val Leu Gly
 35 40 45
 Asn Gly His Leu Ser Ser Leu Gly Thr Arg Asp Thr His Pro Tyr
 50 55 60
 Ala Ser Leu Ser Arg Ala Leu Gln Thr Gln Cys Cys Ile Ser Ser
 65 70 75
 Pro Ser His Leu Met Ser Gln Gln Tyr Arg Pro Tyr Ser Phe Phe
 80 85 90
 Thr Lys Leu Thr Ala Asp Glu Leu Trp Lys Gly Ala Leu Ala Glu
 95 100 105
 Thr Gly Ala Gly Ala Lys Lys Gly Arg Gly Lys Arg Thr Lys Lys
 110 115 120
 Lys Lys Arg Lys Asp Leu Asn Arg Gly Gln Ile Ile Gly Glu Gly
 125 130 135
 Arg Tyr Gly Phe Leu Trp Pro Gly Leu Asn Val Pro Leu Met Lys
 140 145 150
 Asn Gly Ala Val Gln Thr Ile Ala Gln Arg Ser Lys Glu Glu Gln
 155 160 165
 Glu Lys Val Glu Ala Asp Met Ile Gln Gln Arg Glu Glu Trp Asp
 170 175 180
 Arg Lys Lys Lys Met Lys Val Lys Arg Glu Arg Gly Trp Ser Gly
 185 190 195
 Asn Ser Trp Gly Gly Ile Ser Leu Gly Pro Pro Asp Pro Gly Pro
 200 205 210
 Cys Gly Glu Thr Tyr Glu Asp Phe Asp Thr Arg Ile Leu Glu Val
 215 220 225
 Arg Asn Val Phe Thr Met Thr Ala Lys Glu Gly Arg Lys Lys Ser
 230 235 240
 Ile Arg Val Leu Val Ala Val Gly Asn Gly Lys Gly Ala Ala Gly
 245 250 255
 Phe Ser Ile Gly Lys Ala Thr Asp Arg Met Asp Ala Phe Arg Lys
 260 265 270
 Ala Lys Asn Arg Ala Val His His Leu His Tyr Ile Glu Arg Tyr
 275 280 285
 Glu Asp His Thr Ile Phe His Asp Ile Ser Leu Arg Phe Lys Arg
 290 295 300
 Thr His Ile Lys Met Lys Lys Gln Pro Lys Gly Tyr Gly Leu Arg
 305 310 315
 Cys His Arg Ala Ile Ile Thr Ile Cys Arg Leu Ile Gly Ile Lys
 320 325 330
 Asp Met Tyr Ala Lys Val Ser Gly Ser Ile Asn Met Leu Ser Leu
 335 340 345
 Thr Gln Gly Leu Phe Arg Gly Leu Ser Arg Gln Glu Thr His Gln
 350 355 360
 Gln Leu Ala Asp Lys Lys Gly Leu His Val Val Glu Ile Arg Glu
 365 370 375
 Glu Cys Gly Pro Leu Pro Ile Val Val Ala Ser Pro Arg Gly Pro
 380 385 390
 Leu Arg Lys Asp Pro Glu Pro Glu Asp Glu Val Pro Asp Val Lys
 395 400 405
 Leu Asp Trp Glu Asp Val Lys Thr Ala Gln Gly Met Lys Arg Ser
 410 415 420
 Val Trp Ser Asn Leu Lys Arg Ala Ala Thr

WO 01/83524

PCT/US01/13862

425

430

```

<210> 35
<211> 137
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4942964CD1

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

```

```

<210> 36
<211> 380
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5702144CD1

<400> 36
Met Arg Ser Arg Val Leu Trp Gly Ala Ala Arg Trp Leu Trp Pro
1 5 10 15
Arg Arg Ala Val Gly Pro Ala Arg Arg Pro Leu Ser Ser Gly Ser
20 25 30
Pro Pro Leu Glu Glu Leu Phe Thr Arg Gly Gly Pro Leu Arg Thr
35 40 45
Phe Leu Glu Arg Gln Ala Gly Ser Glu Ala His Leu Lys Val Arg
50 55 60
Arg Pro Glu Leu Leu Ala Val Ile Lys Leu Leu Asn Glu Lys Glu
65 70 75
Gln Glu Leu Arg Glu Thr Glu His Leu Leu His Asp Glu Asn Glu
80 85 90
Asp Leu Arg Lys Leu Ala Glu Asn Glu Ile Thr Leu Cys Gln Lys
95 100 105
Glu Ile Thr Gln Leu Lys His Gln Ile Ile Leu Leu Leu Val Pro
110 115 120
Ser Glu Glu Thr Asp Glu Asn Asp Leu Ile Leu Glu Val Thr Ala
125 130 135
Gly Val Gly Gly Gln Glu Ala Met Leu Phe Thr Ser Glu Ile Phe
140 145 150
Asp Met Tyr Gln Gln Tyr Ala Ala Phe Lys Arg Trp His Phe Glu
155 160 165
Thr Leu Glu Tyr Phe Pro Ser Glu Leu Gly Gly Leu Arg His Ala
170 175 180

```

WO 01/83524

PCT/US01/13862

```

Ser Ala Ser Ile Gly Gly Ser Glu Ala Tyr Arg His Met Lys Phe
185 190 195
Glu Gly Gly Val His Arg Val Gln Arg Val Pro Lys Thr Glu Lys
200 205 210
Gln Gly Arg Ile His Thr Ser Thr Met Thr Val Ala Ile Leu Pro
215 220 225
Gln Pro Thr Glu Ile Asn Leu Val Ile Asn Pro Lys Asp Leu Arg
230 235 240
Ile Asp Thr Lys Arg Ala Ser Gly Ala Gly Gly Gln His Val Asn
245 250 255
Thr Thr Asp Ser Ala Val Arg Ile Val His Leu Pro Thr Gly Val
260 265 270
Val Ser Glu Cys Gln Gln Glu Arg Ser Gln Leu Lys Asn Lys Asp
275 280 285
Leu Ala Met Thr Lys Leu Arg Ala Lys Leu Tyr Ser Met His Leu
290 295 300
Glu Glu Glu Ile Asn Lys Arg Gln Asn Ala Arg Lys Ile Gln Ile
305 310 315
Gly Ser Lys Gly Arg Ser Glu Lys Ile Arg Thr Tyr Asn Phe Pro
320 325 330
Gln Asn Arg Val Thr Asp His Arg Ile Asn Lys Thr Leu His Asp
335 340 345
Leu Glu Thr Phe Met Gln Gly Asp Tyr Leu Leu Asp Glu Leu Val
350 355 360
Gln Ser Leu Lys Glu Tyr Ala Asp Tyr Glu Ser Leu Val Glu Ile
365 370 375
Ile Ser Gln Lys Val
380

```

<210> 37
 <211> 206
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5862945CD1

```

<400> 37
Met Ala Ala Ala Val Leu Gly Gln Leu Gly Ala Leu Trp Ile His
1 5 10 15
Asn Leu Arg Ser Arg Gly Lys Leu Ala Leu Gly Val Leu Pro Gln
20 25 30
Ser Tyr Ile His Thr Ser Ala Ser Leu Asp Ile Ser Arg Lys Trp
35 40 45
Glu Lys Lys Asn Lys Ile Val Tyr Pro Pro Gln Leu Pro Gly Glu
50 55 60
Pro Arg Arg Pro Ala Glu Ile Tyr His Cys Arg Arg Gln Ile Lys
65 70 75
Tyr Ser Lys Asp Lys Met Trp Tyr Leu Ala Lys Leu Ile Arg Gly
80 85 90
Met Ser Ile Asp Gln Ala Leu Ala Gln Leu Glu Phe Asn Asp Lys
95 100 105
Lys Gly Ala Lys Ile Ile Lys Glu Val Leu Leu Glu Ala Gln Asp
110 115 120
Met Ala Val Arg Asp His Asn Val Glu Phe Arg Ser Asn Leu Tyr
125 130 135
Ile Ala Glu Ser Thr Ser Gly Arg Gly Gln Cys Leu Lys Arg Ile
140 145 150
Arg Tyr His Gly Arg Gly Arg Phe Gly Ile Met Glu Lys Val Tyr
155 160 165
Cys His Tyr Phe Val Lys Leu Val Glu Gly Pro Pro Pro Pro Pro
170 175 180
Glu Pro Pro Lys Thr Ala Val Ala His Ala Lys Glu Tyr Ile Gln
185 190 195
Gln Leu Arg Ser Arg Thr Ile Val His Thr Leu
200 205

```

WO 01/83524

PCT/US01/13862

```

<210> 38
<211> 190
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6319547CD1

<400> 38
Met Glu Ala Glu Thr Lys Thr Leu Pro Leu Glu Asn Ala Ser Ile
1      5      10      15
Leu Ser Glu Gly Ser Leu Gln Glu Gly His Arg Leu Trp Ile Gly
20     25     30
Asn Leu Asp Pro Lys Ile Thr Glu Tyr His Leu Leu Lys Leu Leu
35     40     45
Gln Lys Phe Gly Lys Val Lys Gln Phe Asp Phe Leu Phe His Lys
50     55     60
Ser Gly Ala Leu Glu Gly Gln Pro Arg Gly Tyr Cys Phe Val Asn
65     70     75
Phe Glu Thr Lys Gln Glu Ala Glu Gln Ala Ile Gln Cys Leu Asn
80     85     90
Gly Lys Leu Ala Leu Ser Lys Lys Leu Val Val Arg Trp Ala His
95     100    105
Ala Gln Val Lys Arg Tyr Asp His Asn Lys Asn Asp Lys Ile Leu
110    115    120
Pro Ile Ser Leu Glu Pro Ser Ser Ser Thr Glu Pro Thr Gln Ser
125    130    135
Asn Leu Ser Val Thr Ala Lys Ile Lys Ala Ile Glu Ala Lys Leu
140    145    150
Lys Met Met Ala Glu Asn Pro Asp Ala Glu Tyr Pro Ala Ala Pro
155    160    165
Val Tyr Ser Tyr Phe Lys Pro Pro Asp Lys Lys Arg Thr Thr Pro
170    175    180
Tyr Ser Arg Thr Ala Trp Lys Ser Arg Arg
185    190

```

```

<210> 39
<211> 434
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 000124CD1

<400> 39
Met Leu Arg Cys Leu Tyr His Trp His Arg Pro Val Leu Asn Arg
1      5      10      15
Arg Trp Ser Arg Leu Cys Leu Leu Lys Gln Tyr Leu Phe Thr Met
20     25     30
Lys Leu Gln Ser Pro Glu Phe Gln Ser Leu Phe Thr Glu Gly Leu
35     40     45
Lys Ser Leu Thr Glu Leu Phe Val Lys Glu Asn His Glu Leu Arg
50     55     60
Ile Ala Gly Gly Ala Val Arg Asp Leu Leu Asn Gly Val Lys Pro
65     70     75
Gln Asp Ile Asp Phe Ala Thr Thr Ala Thr Pro Thr Gln Met Lys
80     85     90
Glu Met Phe Gln Ser Ala Gly Ile Arg Met Ile Asn Asn Arg Gly
95     100    105
Glu Lys His Gly Thr Ile Thr Ala Arg Leu His Glu Glu Asn Phe
110    115    120
Glu Ile Thr Thr Leu Arg Ile Asp Val Thr Thr Asp Gly Arg His
125    130    135
Ala Glu Val Glu Phe Thr Thr Asp Trp Gln Lys Asp Ala Glu Arg
140    145    150

```

WO 01/83524

PCT/US01/13862

```

Arg Asp Leu Thr Ile Asn Ser Met Phe Leu Gly Phe Asp Gly Thr
155 160 165
Leu Phe Asp Tyr Phe Asn Gly Tyr Glu Asp Leu Lys Asn Lys Lys
170 175 180
Val Arg Phe Val Gly His Ala Lys Gln Arg Ile Gln Glu Asp Tyr
185 190 195
Leu Arg Ile Leu Arg Tyr Phe Arg Phe Tyr Gly Arg Ile Val Asp
200 205 210
Lys Pro Gly Asp His Asp Pro Glu Thr Leu Glu Ala Ile Ala Glu
215 220 225
Asn Ala Lys Gly Leu Ala Gly Ile Ser Gly Glu Arg Ile Trp Val
230 235 240
Glu Leu Lys Lys Ile Leu Val Gly Asn His Val Asn His Leu Ile
245 250 255
His Leu Ile Tyr Asp Leu Asp Val Ala Pro Tyr Ile Gly Leu Pro
260 265 270
Ala Asn Ala Ser Leu Glu Glu Phe Asp Lys Val Ser Lys Asn Val
275 280 285
Asp Gly Phe Ser Pro Lys Pro Val Thr Leu Leu Ala Ser Leu Phe
290 295 300
Lys Val Gln Asp Asp Val Thr Lys Leu Asp Leu Arg Leu Lys Ile
305 310 315
Ala Lys Glu Glu Lys Asn Leu Gly Leu Phe Ile Val Lys Asn Arg
320 325 330
Lys Asp Leu Ile Lys Ala Thr Asp Ser Ser Asp Pro Leu Lys Pro
335 340 345
Tyr Gln Asp Phe Ile Ile Asp Ser Arg Glu Pro Asp Ala Thr Thr
350 355 360
Arg Val Cys Glu Leu Leu Lys Tyr Gln Gly Glu His Cys Leu Leu
365 370 375
Glu Met Gln Gln Trp Ser Ile Pro Pro Phe Pro Val Ser Gly
380 385 390
His Asp Ile Arg Lys Val Gly Ile Ser Ser Gly Lys Glu Ile Gly
395 400 405
Ala Leu Leu Gln Gln Leu Arg Glu Gln Trp Lys Lys Ser Gly Tyr
410 415 420
Gln Met Glu Lys Asp Glu Leu Leu Ser Tyr Ile Lys Lys Thr
425 430 435

```

<210> 40

<211> 339

<212> FRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1659474CD1

<400> 40

```

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

```

WO 01/83524

PCT/US01/13862

```

Asn Lys Val Ser Ser His Ile Gly Cys Leu Val His Gly Cys
140 145 150
Phe Asn Ala Ser Ile Pro Lys Pro Glu Gln Leu Ser Ala Glu Gln
155 160 165
Trp Gln Thr Met Glu Ile Asn Met Gly Asp Glu Leu Glu Phe Glu
170 175 180
Val Phe Arg Leu Asp Ser Asp Ala Ala Gly Val Phe Cys Ile Arg
185 190 195
Gly Lys Leu Asn Ile Thr Ser Leu Gln Phe Lys Arg Ser Glu Val
200 205 210
Ser Glu Glu Val Thr Glu Asn Gly Thr Glu Glu Ala Ala Lys Lys
215 220 225
Pro Lys Lys Lys Lys Lys Lys Lys Asp Pro Glu Thr Tyr Glu Val
230 235 240
Asp Ser Gly Thr Thr Lys Leu Ala Asp Asp Ala Asp Asp Thr Pro
245 250 255
Met Glu Glu Ser Ala Leu Gln Asn Thr Asn Asn Ala Asn Gly Ile
260 265 270
Trp Glu Glu Glu Pro Lys Lys Lys Lys Lys Lys Lys Lys His Gln
275 280 285
Glu Val Gln Asp Gln Asp Pro Val Phe Gln Gly Ser Asp Ser Ser
290 295 300
Gly Tyr Gln Ser Asp His Lys Lys Lys Lys Lys Glu Lys Lys Thr
305 310 315
Asn Ser Glu Glu Ala Glu Phe Thr Pro Pro Leu Lys Cys Ser Pro
320 325 330
Lys Arg Lys Gly Lys Ser Asn Phe Leu
335

```

<210> 41

<211> 599

<212> ERT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2267892CD1

<400> 41

```

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

```


WO 01/83524

PCT/US01/13862

```

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

```

<210> 42

<211> 334

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2670307CD1

<400> 42

```

Met Ala Ala Ser Gly Ser Gly Met Ala Gln Lys Thr Trp Glu Leu
1 5 10 15
Ala Asn Asn Met Gln Glu Ala Gln Ser Ile Asp Glu Ile Tyr Lys
20 25 30

```

WO 01/83524

PCT/US01/13862

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

<210> 43

<211> 448

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4524210CD1

<400> 43

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

WO 01/83524

PCT/US01/13862

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

<210> 44

<211> 420

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5584350CD1

<400> 44

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

WO 01/83524

PCT/US01/13862

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

<210> 45
 <211> 137
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5807892CD1

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

WO 01/83524

PCT/US01/13862

Trp	Ile	Gly	Cys	Gly	Ala	His	Leu	Ser	Lys	Pro	Met	Glu	Lys	Leu
80														
Leu	Gly	Leu	Ala	Gly	Phe	Phe	Pro	Leu	His	Pro	Met	Met	Ile	Thr
95														
Asn	Ala	Glu	Arg	Leu	Arg	Arg	Lys	Arg	Ala	Arg	Glu	Val	Leu	Leu
110														
Ala	Ser	Gln	Lys	Thr	Asp	Ala	Glu	Ala	Thr	Asp	Thr	Glu	Ala	Thr
125														
Glu	Thr													

<210> 46

<211> 556

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3210044CD1

<400> 46

Met	Met	Asn	Leu	Pro	Phe	Asn	Arg	Asp	Ala	Val	Phe	Tyr	His	Glu
1														
Asp	Glu	Thr	Asn	Cys	Leu	Leu	Leu	Ile	Met	Ala	Pro	Ser	Phe	Thr
20														
Ala	Arg	Ile	Gln	Leu	Phe	Leu	Leu	Arg	Ala	Leu	Gly	Phe	Leu	Ile
35														
Gly	Leu	Val	Gly	Arg	Ala	Ala	Leu	Val	Leu	Gly	Gly	Pro	Lys	Phe
50														
Ala	Ser	Lys	Thr	Pro	Arg	Pro	Val	Thr	Glu	Pro	Leu	Leu	Leu	Leu
65														
Ser	Gly	Met	Gln	Leu	Ala	Lys	Leu	Ile	Arg	Gln	Arg	Lys	Val	Lys
80														
Cys	Ile	Asp	Val	Val	Gln	Ala	Tyr	Ile	Asn	Arg	Ile	Lys	Asp	Val
95														
Asn	Pro	Met	Ile	Asn	Gly	Ile	Val	Lys	Tyr	Arg	Phe	Glu	Glu	Ala
110														
Met	Lys	Glu	Ala	His	Ala	Val	Asp	Gln	Lys	Leu	Ala	Glu	Lys	Gln
125														
Glu	Asp	Glu	Ala	Thr	Leu	Glu	Asn	Lys	Trp	Pro	Phe	Leu	Gly	Val
140														
Pro	Leu	Thr	Val	Lys	Glu	Ala	Phe	Gln	Leu	Gln	Gly	Met	Pro	Asn
155														
Ser	Ser	Gly	Leu	Met	Asn	Arg	Arg	Asp	Ala	Ile	Ala	Lys	Thr	Asp
170														
Ala	Thr	Val	Val	Ala	Leu	Leu	Lys	Gly	Ala	Gly	Ala	Ile	Pro	Leu
185														
Gly	Ile	Thr	Asn	Cys	Ser	Glu	Leu	Cys	Met	Trp	Tyr	Glu	Ser	Ser
200														
Asn	Lys	Ile	Tyr	Gly	Arg	Ser	Asn	Asn	Pro	Tyr	Asp	Leu	Gln	His
215														
Ile	Val	Gly	Gly	Ser	Ser	Gly	Gly	Glu	Gly	Cys	Thr	Leu	Ala	Ala
230														
Ala	Cys	Ser	Val	Ile	Gly	Val	Gly	Ser	Asp	Ile	Gly	Gly	Ser	Ile
245														
Arg	Met	Pro	Ala	Phe	Phe	Asn	Gly	Ile	Phe	Gly	His	Lys	Pro	Ser
260														
Pro	Gly	Val	Val	Pro	Asn	Lys	Gly	Gln	Phe	Pro	Leu	Ala	Val	Gly
275														
Ala	Gln	Glu	Leu	Phe	Leu	Cys	Thr	Gly	Pro	Met	Cys	Arg	Tyr	Ala
290														
Glu	Asp	Leu	Ala	Pro	Met	Leu	Lys	Val	Met	Ala	Gly	Pro	Gly	Ile
305														
Lys	Arg	Leu	Lys	Leu	Asp	Thr	Lys	Val	His	Leu	Lys	Asp	Leu	Lys
320														
Phe	Tyr	Trp	Met	Glu	His	Asp	Gly	Gly	Ser	Phe	Leu	Met	Ser	Lys

WO 01/83524

PCT/US01/13862

```

335          340          345
Val Asp Gln Asp Leu Ile Met Thr Gln Lys Lys Val Val Val His
350          355          360
Leu Glu Thr Ile Leu Gly Ala Ser Val Gln His Val Lys Leu Lys
365          370          375
Lys Met Lys Tyr Ser Phe Gln Leu Trp Ile Ala Met Met Ser Ala
380          385          390
Lys Gly His Asp Gly Lys Glu Pro Val Lys Phe Val Asp Leu Leu
395          400          405
Gly Asp His Gly Lys His Val Ser Pro Leu Trp Glu Leu Ile Lys
410          415          420
Trp Cys Leu Gly Leu Ser Val Tyr Thr Ile Pro Ser Ile Gly Leu
425          430          435
Ala Leu Leu Glu Gln Lys Leu Arg Tyr Ser Asn Glu Lys Tyr Gln
440          445          450
Lys Phe Lys Ala Val Glu Glu Ser Leu Arg Lys Glu Leu Val Asp
455          460          465
Met Leu Gly Asp Asp Gly Val Phe Leu Tyr Pro Ser His Pro Thr
470          475          480
Val Ala Pro Lys His His Val Pro Leu Thr Arg Pro Phe Asn Phe
485          490          495
Ala Tyr Thr Gly Val Phe Ser Ala Leu Gly Leu Pro Val Thr Gln
500          505          510
Cys Pro Leu Gly Leu Asn Ala Lys Gly Leu Pro Leu Gly Ile Gln
515          520          525
Val Val Ala Gly Pro Phe Asn Asp His Leu Thr Leu Ala Val Ala
530          535          540
Gln Tyr Leu Glu Lys Thr Phe Gly Gly Trp Val Cys Pro Gly Lys
545          550          555
Phe

```

```

<210> 47
<211> 111
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4942454CD1

```

```

<400> 47
Met Lys Phe Val Ala Ala Tyr Leu Leu Ala Val Leu Ala Gly Asn
1      5      10      15
Ser Ser Pro Ser Ala Glu Asp Leu Thr Ala Ile Leu Glu Ser Val
20     25     30
Gly Cys Glu Val Asp Asn Glu Lys Met Glu Leu Leu Leu Ser Gln
35     40     45
Leu Ser Gly Lys Asp Ile Thr Glu Leu Ile Ala Ala Gly Arg Glu
50     55     60
Lys Phe Ala Ser Val Pro Cys Gly Gly Gly Gly Val Ala Val Ala
65     70     75
Ala Ala Ala Pro Ala Ala Gly Gly Ala Pro Ala Ala Glu Ala Lys
80     85     90
Lys Glu Glu Lys Val Glu Glu Lys Glu Glu Ser Asp Asp Asp Met
95     100    105
Gly Phe Ser Leu Phe Asp
110

```

```

<210> 48
<211> 882
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1622129CB1

```

WO 01/83524

PCT/US01/13862

```

<400> 48
ccacacgggtc cggggagcgg cggggagctg tagttctccc ggggtcaga gaagtaggca 60
gagagcgggac ctggcggcgg ggcagcatgg cggggctgga gctctgtcg gaccagggtc 120
accgggtgga cgggaggcgg gccggggagc tggcgaagat ccaggcggcg atggscatgt 180
tcgcgcaggc tgaacggctcg gctacatctg agcaggggaa caccaggcca ctggctatgg 240
tcacaggccc gacagagatc cggggctccc gggctcgagg cctgcaggac agggcctag 300
tgaactgtca alatagttca ggcaccttca gcaacagtgga gggcaagcga cggccacatg 360
gggacggtaa gtccgtgtgag atgggcctgc agctccgcca gaatttcgaa gacgcaccc 420
tcacacagct gacccacagg tccagatctg atatatatgt gcaggtgcta caggcagatg 480
gtgggacctt tgcagcttgt gtgaalgcag ccacgctggc agtgcctgat gccgggatac 540
ccatgagaga ctttgttgtt gctgtctcag ctggctctgt ggaaggcaca gccctggcgg 600
acctcagcca tgtggaggaa gcaagctggg gcccccagct ggcctggccc ctgctgccag 660
ctccaggaca gatctgcctg ctgagatagg atggccggtt gcaacaggac cacttgagag 720
gggtgttga ggcgtgtgac caggctgccc gagatgtgca caccctctta gatcgatgg 780
tcggcgagca tgtgcgtgag gctctcatct tgcctggggg ctgacacccc agccacccat 840
gtccagcata aaacccctct ctgccacaaa aaaaaaaaaa aa 882

```

```

<210> 48
<211> 1220
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1820078CB1

```

```

<400> 49
ctcaacttta gccgcggga agcggaagtc aggtggttgt cggatttttag aggaaggcgc 60
tgggttaaat tggggaactg ggttggtctg gatttccacg gtgtagtga ccaggaggca 120
cctctcctgg gcttctcagt gtctgcggcg cggggttcga cctgagctcg attgacatag 180
cccttgaggg atttaaacaa cctaacattt agcagctaca gctgctctaa acctttggga 240
tttccagaaat gaactgacct gccgaacgtg ttccnaagtt tgaagagatg ttgtctagta 300
gattccacaga naatgacaag gagtatcagg natarctgna acgcccctct gagtctcttc 360
caatttgtga ggaatggaat agcagagctg gtgggaacca aagaaacaga ggcacatcgt 420
tgcagaacaa cagacagttc aagaggcagg acaacagatg ggggttgcca agtgacatc 480
gatccaaatc agggcaatga cgaacctggg glaacaacta ccgcacacac agcaaacaa 540
cttactatcc ccagcaatat ggaacttatg gttacaacca cgggctctct taccgttact 600
actgatagaa atgttggcag ctttttagtaa aagcatttac tctgttacea tgaganaagt 660
ttgggtgtct tctgttggtc atagttttac atctgatttt acagatgga ttattgattt 720
tttggagatt gaggacttta aaaaaataga tcttacttgc gaactgcgat ggtgtctggg 780
aataactgaa acgtgtgatt atattgcttg acttctacot cagaactctt ttgtttcat 840
gaacttaatag tgccttaagt ttggtatatt alitgaactc taggaattct ttgttttaca 900
cugaaataaa aattttaaaa tagaanaatgc ttttacttly tauggtlaaga gattatccat 960
atgcttagat gtgctgtgtt ctanaattct agnngttgat ataacagct catgattga 1020
cagctatgct ttttgtgata gattgtacat aacatcagca gttgaagggt aaacacattg 1080
cttttttttt tttttgcatt tgttaagtga ctatggtact ttgtgattcc ttaattctata 1140
gatgagtcag ctccacacatt gagtctcttt ttagagggaa atcagtaata aagctgtanaa 1200
atagggaagg aaaaaaaaaa 1220

```

```

<210> 50
<211> 2020
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1527017CB1
<230>
<231> unsure
<232> 1905, 1909, 1936, 1942, 1959, 1967, 1972, 1979, 1989, 1999, 2002
<233> a, t, c, g, or other

```

```

<400> 50
gagaggagtg agtgcctga ccgaggcccg ccgcagactg cgaaggatac agggagggca 60
aggttttctt tttggcctt ccttttggac ccggagtgga aaaactctaa agtccagatc 120
agtggagaga aacgcagatt taggacctg aggagtcttt ttcacccgtt tccgtcaact 180

```

WO 01/83524

PCT/US01/13862

```

cgctcaggcg cgcgagggcg agtccttctg gggctectcg ggcagagcaa gatggttgac 240
cccgagtgga aggttgcccg aggatggctg ggcctggcgt tggcgctgcg ggggctgtgc 300
ttgcagcttc caggggctaa ctcaggtgag atggagccgc talagtctcg aattcaagga 360
tcctctgatt gacaaggaat attatcgcaa gccagtggag gagctaacgt agggagagaa 420
atactgttcg gagctcaaga agactcagct catcaaaagt gctccagcag gggaaaaaaa 480
gttctgtgtt tgaagaccca gtcatcagta aattcaacaa catgatgatg ataggagaa 540
acaaagtaact ggcagagatcc ctcatgattc agactctgga agctgtgaaa aggaagcagt 600
ttgagaagta ccabgcgcct tctgcagagg aacaggcaac catcgaaacg aaccttaca 660
ccalcctbca tcaagcactg aaaaactgtg agcctatgac tgggctggta cccatctoca 720
agggagggcg ttctaccag gtccctgtac cctaccoga cggcgctgcg cgcttcttag 780
ccatgaagtg gatgacact dagtgccggg ataaaaagca ccagcgaca ctgatccgg 840
agaaagctgc acaaaagctg ctggaggttt tccataaaca gggcccgctg atcaagagaa 900
agcatgactt gcacaagatg gcagagggca accgtgcccc gggccacatc cgctgtgtgt 960
agagctctca ggaagagccc agggccctct gcgcaagaa aaggtgtgag ctactgcaac 1020
gtcgaaactt acctgtgggt taaggatgta gttcctttgt aagggtgggc aggcctcgta 1080
agaaagatgt agcagcatat tcaatcccg ttaactcttc ttcttttag gctggaactt 1140
gctctctctg cccctatttc cttgtaagga gggagcacat tgaactggga atttctcca 1200
ggaaactcag ggctgttttc tcttccctta ggttggggcg gacctttgga catataaagg 1260
aagcagtttt agtatcagaa aagatttatt agaaaattct cagcgtgaac tgggttagca 1320
tgtgtgtcag cattcagtg aactggctgg aggaatagg ctgttttoca ggtttgtctt 1380
tatataaaat gtatataaag cagtttcttg tgtgacttgt gctctgtctc cacccttga 1440
catcccaaaa tatcccaaaa gtggtatggt ttacccattt taacagtgag taactgtagg 1500
caccaagata gtagttgcac taatggttac acagtgcagt ggcctttggg agttgcccct 1560
ctctgctctg ccgtgtgtgg ttgtgtgtgg gaagggggct cagggcagga ccacggcata 1620
agtggggaac atctccacg gagatgggaa agtctagaag ggaagcacat caagctctgg 1680
aagggaaaag tctttgggtg aggcagagac tccactgcra gctttagagg tgggttagag 1740
aaagccaggt gctgtgtgag aaacctgtat ctggagagcta gtccgagact tgcgtgtagt 1800
atactgttgg cactggcgtt gcttccagcc gttggcgtgt gttctttccc aagcccgggc 1860
cgcgcccgag gaacttcca atgaaatttt tccaaagaa agcnaagcna accttggggg 1920
cccaagggaa agctnaaag gnoccaattt ttggggaanc cagggnaaaa gnoccatttc 1980
ccccggggtt ttaaggccnc cnaaaaagaa ggttccctt 2020

```

<210> 51
 <211> 637
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1647264CB1

```

<400> 51
cggccagctg aagctaaaat taacctcac taagggaat aagcttgagg cgcgcagacc 60
cagctcagcc cgcgagagcc ttgtccggca ccgtaacaa tggagcgccc ggataagggc 120
ggcgtgaacg cactgaagcc tctgagttc agaaatgaaa gctcatttag attctacactg 180
aagacgctcc tgtttctcac agctttaatg atcactgttc ctattgggtt atattttaca 240
actaaattct acatatttga agcgccctct gggatgtcca atagggaacg ctatttttac 300
gctgctattg ttgcagtggt cgcgcgtccat gttgtgtctg cctcttttgt gtatgtggcc 360
tggaaatgaag gctcagagca gtggcgtgaa ggcacacagg attaaagtga acatcacctt 420
tttatgcatc taatttcatt ttttaaatg ataatgtctg gagggggcca tctgatttga 480
atcaagctga agaaacatgt taagtcaggt ctttaaggagt cagcttttag tatgtaaatt 540
ttgatctctc taatgtgttg ggtttgatat tcaagttttac tgtatgaatc gattgcaatg 600
agaaattggaa aagtagtaca agaatatgta attatta 637

```

<210> 52
 <211> 717
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1721598CB1

```

<400> 52
ggctggggcct ggcgcgcagg cgctaggaag agggcgcgct gggcgagggc ggcgcttggc 60
tgggtggggcc cgcgagggga ttttccgggg cggcgagagc ggaatctatc tgggatacca 120
tggctttctt tactgggctc tggggccccc tcaactgtgt aagcagagtg ctgagccatc 180

```


WO 01/83524

PCT/US01/13862

```

acgtttttcag caccactggg agtctgagtg cgattcagaa gatgacggcg gtacgagtg 240
tgacacacag tgcctctggg aacagcccat accatcgggc tctcgcctgc atccatgtct 300
ataagaagaa tggagtgggc aaggtgggag accagatact actggccatc aagggaacaga 360
agaaaaaggg gctcattgtg gggcaactgca tgcctggccc ccgaatgacc ccagattttg 420
actccacaaa cgtgggtcctc attgaggaca acgggaaccc tggggggaca cgaattaaaga 480
caccctacc caccagcctg cgaagcggg aaggcgagta ttccagggtg ctggccattg 540
ctcagaactt tgtgtgagtt gaggccaggc ctctgtgtgc aggaactcgt atggaagcag 600
ttctgagAAC cacccttttg ctaaggagag ttgggagcca catggctcgt ccttccacac 660
tgggtaacag ttagtatacc tgtgagagaa taaatgtatt catttaaaaa aaaaaaa 717

```

```

<210> 53
<211> 2061
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> Incyte ID No: 1730581CB1

```

```

<400> 53
ggctcgcaca cacttcggca ctaggaagag caggaagagg caggaagagg cagccgggt gagcagaagg 60
atcggtccag tagacagcaa accagcaaca accgaactgag ctgcataact accgttaag 120
catctaacca accgcccatc tagctaaccc gagccctccc accgttaact caggttcagg 180
cgttcctccgg ccgcctctgc gtagccgttg tggcagcccc gggagagaca ctggcgtctg 240
tttctctoga ttctcgggat tcgaagatgg ctgcacagtc aggcgcgaaa gttgtgtctg 300
aagcaccac caagatgtct ctaaatgagc gctttactaa tatgctgaag aacaaacagc 360
cgagccaggt gaattattcg gcttcgatgc agcaacaaca gcagctagcc agtgccagaa 420
acgaagact ggtccagcag atggagaata gacctctgt coaggcagca ttaaaaotta 480
agcagaagag cttaagcag cgtctgggta agagtaacat ccaggcaggg ttaggccagc 540
cctaggggc cctggccagg gtagcaatcg gagccagag cctacccata attcagagag 600
gcttcccag aggagagact cgtgggggac gtgccaccag aacctactt agggggcggg 660
tgtcactccg aggtcaaaac ctgctccgag gtggccagag cgtagctccc cgaatggggt 720
taagaagagg tgggtgttga ggtcgtggag gtctcgggag agggggccta gggcgtggag 780
ctatgggtcg tggcgggaac ggtggttagg gtccgggtat gataggtcgg ggaagagggg 840
gctttggagg ccgagggcga ggcctgtgac gaggagagag tgcccttgtc cgccctgtat 900
tgaccaagga gcagctggac aaccaattgg atgcataat gtcaaaaaa aaaggaaccc 960
tgagatctga gttggtgac taactggcgc agccagatcc cgaaccaat gattgaagcc 1020
tgcccatcct ccatgagag actctgttta ctcaacacat ctgtaataa ccttgagata 1080
acagatgaga agnaactctga ttgatgtcgt atggacctat cacaatagcg tgtggacata 1140
cttgccacca gcttgtgcat ttagtgtgtt ccttttactt ttgtatactg tgttgtatga 1200
aacctttttt tcttttgatt tgggtttttt tttttgtttt ttttaggggg aggggggggt 1260
tcctctctct tgcaccagct tctctttgaa cacaatgca tttagcctgt ggttagaaca 1320
cctctctctc acctctgtct cccctcactt gtcatactgt ctgacatgct aacatttctt 1380
ttgttatccc ctgttgccc cagagaaaca accggtcagt gttcttctct 1440
ccctgatcct taggtttctg aaatagggt ctgtttactc ctcttgata gctgtgttaa 1500
aatgtttaga aggtctggag ctcaaaaatg cgtttctcca cattgataat ttagtanaat 1560
gagaacattg acatcactac agggcagcag aagaggttgc ttacatgtg tagcagctct 1620
ggtttgatcc aagttgctac catgtacatt gacagcacat ataccataac cagcgtgtgt 1680
ggttgaattg cactttctac ctttgtatga gatttacaga ctttctctct ggttttgtat 1740
catgaccaga ggggtactat aggggtgtgt tatactgcaa tatagaggat cagaagccat 1800
ttgatttggt aggtgtgtca gaaggagaa tcatgtcaga cgaactgtct gaagaggtca 1860
gaagabagc algttaaaat gcaattatct cctctatgtt atcccaact aactctggac 1920
tttccactc attagctttt ttttgcctc gtctcccttg aaggttttaag ttcaacata 1980
ttctgtcaac tgttcagttt cagtggaaac ttgtatttct ggttcattat acaaaactgt 2040
tcgtctaaaa aaaaaaaaaa a 2061

```

```

<210> 54
<211> 1307
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> Incyte ID No: 1740714CB1

```

```

<220>
<221> unsure

```

WO 01/83524

PCT/US01/13862

<222> 1301

<223> a, t, c, g, or other

<400> 54

```

ggcgtgtgac ctagaatggg cgcattgcgc gaggcgaact ggcgtgttg aaaaaccatgg 60
cgtgggtacc agcggagtcgc gcaatggaaag agttgatgco tgggtattg ccgttaagc 120
cttgogactt gaaggaggtt ttogatccct cgttaccccc gaggacgctt caggaaatacc 180
tgagtcgggt ccagatcgaa gcagctcaat gtccagatgt tgtggtagct caaatggacc 240
caaaagagtt gaaaaggaag caaagtgtga atatttctct tttaggatgc caaccgccgc 300
ctgaaggtta ttcccacaac cttcaatggc aacagcaaca agtggccagc ttttcaactg 360
ttgcacagaa tgtgaacaaa catagaagtc actggaaatc acaacagttg gatagtaatg 420
tcgcaatggc aaaaatcgaa gatgaagagc gctggaaaga atttgcctg ggtgaaagtc 480
tatgtgctga cggggtgttt ggaacagcca caaatgaagc tccgtgaaat gattatgtac 540
aagcaacagt aactagtgtc ttggaatc tcagtaattg gtttggagaa agagacttta 600
ctccagcaatt gggaagatgg ctttatgctt tattggcttg tctgaaaag ccttggttac 660
ctgaggtcca ttcaatgatt cggcagcttg caagaagttg cctgaaagt aggcctttag 720
tggtatagcaa agatgatgag agggttcctg ctttgaattt attaatctgc ttggttagca 780
ggtattttga ccaacgtgat tttagctgat agccatcttg atgtagctga tctctcagg 840
atagaagata tttctcatga aggcagccta actctgagga aaacaatgcc aattcaagta 900
cagatttcaa cacatcttca acaatctatg aagggttcac atctaacctt gtcacattca 960
gatgtgaact cagaatatgg gttgatttga atatctgaaa tatcaattga aattccact 1020
cagtttttga tgaaacagttt gaacagtttt ctgtaactaa gcagcttga tagaaattgt 1080
atgtgaaat ttacatagg ttcttggctg tgttttgtt ttttttgtt ttttgttgt 1140
ttgttattta cttatataca tataaaattt tattgaaat atgttttgtt tactaaaaat 1200
ttgtttgact cctaacaanaa gacaatggat ggctttagca aataactcgg 1260
atcaatggc aatgtgttca tagtcagcaa taaaattaaa natttta 1307

```

<210> 55

<211> 1357

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1850596CB1

<400> 55

```

ggggcgcgcg acggcgcacg ctgggggcag cgggaaccag agaagctgaa ggggcggtag 60
ggggcgcgac ggcgaogacg acgaactccg cggctgtgco cagcctctt ccgcgcagc 120
cgcccttttc ctccctccct taagtccccc agtgccgcag taccgcctcc ttcccagcc 180
cgccgcttcc tccagacctc tgggcgcggg tgagccctat tcccagagcg aggtgtgtgt 240
gacctgttaa ccaaaaggag gaacacgctg gctaaagctca tcattgttac tgggtggcac 300
aatgtccttg aagctccagg caagcaatgt aaccaacagc aatgcaccca agtccatcaa 360
ctctcgagtc ttcatgggaa acctcaaac acgtctgttg aagaactcag atgtgagac 420
catctctctt aagtatggcc gtgtggccgg ctgtttctgt cacaaggcct atgcctttgt 480
tcagtactcc aatgagcgcc atgcccgggc agctgtgtgt gtagagaaat ggccgggtgt 540
ggccgggcag accctggaca tcaaatggc tggagagcct aagcctgaca gacccaaggg 600
gctaagagaga gcaatcatct ccatatacag tggctacatc ttgtactatg attactacc 660
ggacgacttc taagcaagggc tcttctacta ccggggccgl ctgtcgcgcg tgcagtgc 720
caggcgggtc cctgtgaagc gaccccggtt cacagtccct ttgttcgggc gtgtcaaac 780
taagctactt gtcagctctt ttgcccgttc caagctgttc accccagctt cagccaagat 840
caagttaaag agcagtgagc tgcaggccat caagcggag ctgacacaga tcaagtcaa 900
tatcgatgcc ctgctgagcc gcttgagaca gatcgtctgg gagcaaaagg ccaatccaga 960
tggcaagaag aaggtgtgat gagggtggcg cggcgccggc ggcygtgtg gtggcagcg 1020
tggcgttggc agtgtgtgtg ggggtggcgg tggcagcagc cggccacagc cccccaaga 1080
gaacacaact tctgaggcag gctgcaccca gggggaagca cggaccagag acgacggcga 1140
tgaggaaaggg ctcttgacac acagcgagga agagctggaa cacagccagg acacagcgc 1200
ggaatgatgg gcectbgcagc aagcagcctg acaggagcaa tggccacag caggtgaagg 1260
gcatacgtgc cccagggcto aagccgggca ccaaacctg galgcacccc ccagcgggt 1320
accagaggaag agctggcagc aggcgcctcc tccccca 1357

```

<210> 56

<211> 1749

<212> DNA

<213> Homo sapiens

<220>

WO 01/83524

PCT/US01/13862

<221> misc_feature

<223> Incyte ID No: 1856109CB1

<400> 56

```

tcgttttctc  tcgttcaatg  gcgtccgggc  60
tggtaagatt  gctgcagag  ggcacatgct  120
tccttcaggt  tcggggaggt  taagaaggat  180
tagagcggrc  ggcctctctg  cgggtgccgc  240
actgggatta  coakgcagaa  atacaagctt  300
tagatcttct  caaaactgca  ttgttaata  360
gcacaacact  tgggatalag  aaagaagctg  420
tatecgaaac  agggacatct  ttttcacaga  480
acccacacat  gccacatgaa  ggcataaaaa  540
tcgtgtgtca  cgtgtgtgga  aaattgtgtg  600
cagtgcctcc  agctgtgtta  cagcagactt  660
gcagtggacc  tgacaggact  gcaactttca  720
gaaaagagct  ctttgagatg  tggaaagtaa  780
tgaagaaaag  gaattgttca  gctcctgaat  840
cagctttgct  tttctatttt  gttggcttat  900
ctggggaaac  agtatgtgtt  gcagaagaag  960
atggattcac  agaaataga  cggcgtggga  1020
cagaaaaaga  catcactgca  agctagcgc  1080
gaagtgaga  taaatgcoaa  aggtgtttca  1140
agattgtttt  tttctgtttt  ttaactgtgt  1200
cagtgttttt  ggtttgtttt  ttctgaatc  1260
ttgagatgga  gtcttactct  ttcgcccagg  1320
ttcaacctcc  acctcacagg  ttcaagcgat  1380
atcacaggca  cacaccacca  tacctggcta  1440
caccaagttg  gctagaccag  tcctgaactc  1500
tcccagaatg  ctggatttac  aggtgtgagc  1560
tgacattttt  gcaaaaaaat  tttcctaagt  1620
attagtgata  tttattttgc  ctctagcttt  1680
attttttgct  gtaataccaa  taaagttttt  1740
aaaaaaaa

```

<210> 57

<211> 991

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1921719CB1

<400> 57

```

cgaagatgg  cgcgcgccgt  aaggcggacg  ctgttagggg  tgscgggggg  ttggcggcgg  60
ttegagaggc  tetgggcggg  cagctctaagc  totgcagcc  tgctctttgc  agccgcaccc  120
tcagacacgc  gatcccatg  ggcctgtttg  ggcgcgttgt  gctgcagcg  gccacttgta  180
gtctccagc  cgttgacccc  attgcaggaa  gagatggcgt  ctctactgca  gcagattgag  240
atagagagaa  goctgtattc  agaaccagag  ctctgtgtct  tggatgaaa  ccagcgactg  300
gcnaagaaga  aagctgacct  tcatgatgaa  gaagatgaac  aggatataat  gctgggcgaa  360
gatttgaaa  atatgtggga  gcagaatct  ctacagttca  aacttgagc  tcgcataaca  420
gaagctgatg  aazagaatga  cgaacatcc  ctgaacagga  agctagacag  gaacctgttc  480
ctgtatgtca  gagagaagtt  tggagaccag  gatgtttgga  tactgccca  gccagagttg  540
cagctgtggg  agacccctcg  aggaacagct  gaacgaaccc  tggccacact  ctacgaacac  600
aacatggaa  ccaagtctct  aggaatgca  cctgtgggc  actacacatt  caagtccccc  660
cagggcaatg  ggcacagag  taaactcgga  gccaaaggtg  tctcttccaa  agcactgcta  720
ttaactggag  acctttccca  ggctgggaat  aagggccatc  atgtgtgggt  cactaaggat  780
gagctgggtg  actatttgaa  accaaaaaac  ctggcccaag  ttgagaggtt  tgtttcagac  840
ctctgatgg  ccgagctgca  tgttgacggt  gctcagacaa  gctctggggt  agagctcaaa  900
ggacattgtg  tgattgctc  aaatttgag  gtaatatcaa  gacgaacact  aaattctgag  960
aaataaacga  gtctattact  gaaaaaaaaa  a

```

<210> 58

<211> 1188

<212> DNA

<213> Homo sapiens

WO 01/83524

PCT/US01/13862

<220>

<221> misc_feature

<223> Incyte ID No: 2099829CB1

<400> 58

```

cggctctccg cgcacgtggt attcagcgcg atgcccacat ccaagcggga caagaaagtc 60
tcctttaacc aaactgcccac gaagggcttg gaattgaac aaactctgat agagagcttt 120
cggaatgttg tggacacctt caagtacctt ttcatctct ctgtggccaa catgaggaac 180
agcaagctga aggcacatcc gaacgcctgg aagcacagcc ggaatgtttt tggcaaaaac 240
aagytgatga tggtagcctt gggtaggagc ccattctgat aatacaaga caacctgcac 300
caggtcagca aaagggttag ggttagggtg ggtctctgt tcaaccaacg cacaagagag 360
gaggtgaatg agtgggttac gaaatacaca gaaatggact acgcccggag tggtaacaaa 420
cgagctttca ctgtagagct ggaatccagg cccctggagg agttcccca ctccatggag 480
ccacagctca ggcagctggg cctggccacc ggcctcaaga gaagtggtgt gactctgctg 540
ctgactacg aggtgtgcac ggaaggcgat gtgtgaccc cagagcaggc tgcgtctctg 600
aagctttttg ggtatgagat ggttgaatbc aaggtgacca tcaatacat gtgggattca 660
cagtcgggaa ggttccagca gatgggagac gacttgccag agagcgcatc tgagtccaca 720
gagagctcag actcagaaga tgatgactga aagggactcg ggaactgagg tctcttgaaa 780
gcttctgggt ctacttggac catcaggact gctgcgccc ctctggagag agcagctttt 840
tattttctgt tagacaggga acatgatggg caotgacctc ctgtaaagaa taaactgttg 900
ggccgggggc ggtgggtcac gcttggaaac ccagcacttt gggaaagcga gttgggcaga 960
tcaaaaggtc aggaatttaa gaactctct gtaaacaggy tctctacta 1020
aaatagaaaa aaaaactagt tgggcatagt ggcattgccc tctagtcoca gctaactagg 1080
aggttgaggg aggagaatca cttgaacccg ggaagtgagg gttgcctga gttgagattg 1140
gaccactgct ctccagccty ggcaacagag taaaactctg tcccaaaa 1188

```

<210> 58

<211> 1454

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2416915CB1

<400> 58

```

gttgtacatc tctcgggttg ttactctgta gcttcocgga tggcgaagg gaggaacctg 60
cttggtcatg gatttttgaga atcttttctc aaaaaccccc aacccggccc tggcaaaaac 120
ggccacggac tctgacgnaa gaatcgatga tgaatatgat acagaagttg aagaacacac 180
agaaagaaaa attaaactgg agtgccgaca aattcccaaa aaatttagac actctgcaat 240
atocccaaaa agttcgtctg atagaaaaac aagaagtaag gactatgatg tatatagtga 300
taatgatatc tgcagtcagg aatcagaaga taattttgcc aaagagcttc aacagtacat 360
acaaagcaga gaaatggcaa atgctgtctc acctgaagaa tctacaaga aagaaggagt 420
aaagagctcc ccacaggtct ctacacaaa aaataaaaat cttaagcty gtcacagaa 480
tggcaaacag aagaataaga agcgaataa gctcggcctt ggaacaaaag gatcaatgc 540
tttgctgagg aacagcggct ccaaggaaga ggaatgttaa cctaaagaga agcagcagca 600
tttgagtcag gcaatctatc accaacatc agtggaacgc aagggaaaac aaatttgtaa 660
atattttctt gaaaggaaat gtattaaagg agaccagtyt aaatttgatc atgatgcaga 720
gatagaaaaa aaaaaggaaa tgtgttaagt ttatgtacaa ggaatttgta ccagaggtga 780
aaactgtcty tatttgcata atgaatatcc ttgtaagttt taacatacag gaacaaaatg 840
ttatcagggg gaatactgca agtttttctc tgcctccacty actcctgaaa cacaagaatt 900
gttggttaaa gttttggata ctgaagaaga gtcattgtaaa taatatagac ataaaaaggt 960
agcaatgtac agataaagag tactttaacg ccactgctgt ttcaagctgt ttcaagctgt 1020
gtgatttgga gtatgttaca agattcctca ttcagagtyc cctcttgyt gactgggggt 1080
atggtcagct tccataatgg atgggacaga gagctgggat ctaatgtaca agtgaagggc 1140
ttggtcttcc ctgagacatt ccagccattg gaataggaga ggaatcata tggcagaggt 1200
gatggctggt ggttaaatgt gatagtaaat tgtagaacc tctctgtatt gattggattt 1260
cottaataaa atcggaaagca aggttagggt gagttaggtt gagtaagag gtagagaggy 1320
tttgaggaga gagaactgct cgggaagcat tgytagatgg accataaaaa cagagtaggt 1380
tctcttctat acattaaata gtttlaaac atatttttaa tgyttacaaa tttaatttta 1440
ggtcttaaaa taaa 1454

```

<210> 60

<211> 1588

<212> DNA

<213> Homo sapiens

WO 01/83524

PCT/US01/13862

<220>

<221> misc_feature

<223> Incyte ID No: 2472784CB1

<400> 60

```

ccccggacgaa gggggagagat agacagcaga accaggggcy ggggctaagc agagactgta 60
gtacggcgga cagcgagcag ggcagcagtg gctgggggcy ggcagcccc ggcactcccg 120
ggtcagagag gaccgctgtt cccgggtcct ggcgttgca tccggggcaa aagcggcgag 180
gaacgcttga aggaatgtga ggcggagatg gccctgttt agcagggaat tctgggggct 240
ccagtcctg gaatcccaac tgcgtgtcct ggcgtgcca ctgtccccc ggtcccccaca 300
gtagaagcga tgcagytccc agcgggtcct gtgatccgcc caattatcgc gaacaaacaa 360
taccagcagg tccagcagac tctggagggc cagcagcgtg ctgcagccac agtagttcct 420
ccccgtgtgg gtggccctcc ttttctaggc cctgttggct ttggccctgg tgatcggagt 480
ccccgggaca gcccgagggc togaagaagc atgttccctg ggcgggcagc agcgttccc 540
cgccctatgg ccttaccgac cctcacag ggcctcgtgg gccccctct gcttggggcc 600
cctggaccac ccatgatgct gccaccaatg gctcgggctc caggggcccc gctgggctcc 660
atggctgacac tgaagccccc tctggaagag ccagcagcac ccgagagct gggcctaggc 720
ctggggttgg gcttgaaga gaaggaagag gcagtggtgg cggcgggcgg tgggtggag 780
gaggttagcg cggctgttgg cgtgggggca ggaagtgcgc cagctggccc tgaagtcat 840
gggcccagcc tgcgcttggc cctggccatg ccattgccc agcctggacc cctgcccctc 900
ccgttggagg tcttccggcg cctcctgccc ccgttgcca ttcctgaact cctgtccctg 960
cgtctcgggc ccggcccc cctggcagag ccaccccag gctcctaggc tcttgaagtc 1020
ccagggcccc tgggtgaaga caagaagaag ggaagccag agaaattgaa acggtgcatt 1080
cgccacggcg caggggagcag ctgggaggac ccagcctgc tggagtggga tgcagtgac 1140
ttccggttct tcttggggga tctgggaagt gagggtgaag atgacattt ggcacggccc 1200
ttccgctcgt tcccatcctt ccttaaggcc aagtggtatc gtgacaagcg cacaggcaag 1260
accagggctt acggtctcgt cagcttcaag gcccccagcg actactgctg cgcctatgct 1320
gagatgaatg gaaagtatgt gggctcggcg cccatcaagc ttcgcaagag catgtggaag 1380
gacgggaatc tggagctggt ccgcaagagc cagaagagaa aagagaagct gggcctgaga 1440
taggtctggt ggcaggggaa ccgttccca ctggccgggc gctgggtcct ccttagtctc 1500
tctttggaaa acccccagct gtccacccat cccctgcccc aaaaacagtt caataaatt 1560
tcgttccatt tccaaaaaaa aaaaaaaa 1588

```

<210> 61

<211> 2111

<212> DKA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2598981CB1

<400> 61

```

cgccaggcggg cctcgcgggt ccggggagcgc gggggagagc atgcttgaga tcagagtacc 60
gccccttgggg gcccggcagg acgttgggcg aagctgcata ctgttcca tgcgggcaa 120
gaatgtcatg ctggactgtg gaatgcaat gggcttcaat gacagccagc gcttccctga 180
cttctctac atcaccocaga ccggccgccc aacagacttc ctggactgtg tgatcattag 240
ccacttcac ctggaccact ggggggcaact cccctacttc agcagatgtg tgggtatcga 300
cgggcccatc tacatgactc accccaccca ggcctatctg cccatcttgc tggaggacta 360
ccgcaagatc ggcgtagaca agaagggcga ggcacacttc ttcactccc agatgatcaa 420
agactyctatg aagaaggttg tggctgtcca cctccacag acggtccagg tagatgatga 480
gcttggagatc aaggtctact atgcaggcca cytcttgggg gcagccatgt tccagatcaa 540
agtggtgtca gagtctgtgg tctcacggg tgattataa atgacccag acgcaact 600
agagatctcc tggattgaca agtgcgcgcc caactgtctc atcacagagt ccagctacgc 660
cacgacatc cgtgactcca agcgtgcgag ggcagagagc ttcctgaaga aagtcacaga 720
gacgttgag cgtgttggga aggtgctgat acctgtgttc ggcctgggac ggcgccagga 780
gctctgcatc ctcttgga ccttctggga ggcctgaac ctgaaggtgc ccatctactt 840
ctccacgggg ctgacccaga aggcacacca ctactacaag ctgttcatcc cctggaccaa 900
ccgaagalc cgaagactt tctgtcagag gaacatgttt gagtccaagc acatcaggc 960
cttgcacgg gcttctgtg acacccaggg ccgatgttt gttgttgcga ggcgggact 1020
gctgcaagct ggcagctccc tgcagctctt ccggaatgg gccgggaagc aagagacat 1080
ggtcatcag cccgctact gcttgcaggg ccactcagcc cacagatc tccaggggca 1140
gggttaagctc gatatggagg ggcggcaggt cctggaggtc aagatgcagg tggagtacat 1200
gtcatcagc gacacggcg acgccaaggg catcaltgag ctggtggggc aggcagagcc 1260
ggagagcgtg ctgctgttgc atggcagggc caagaagatg gagtctctga agcagaagat 1320
cgagcaggag ctcgggtcca actgtacat gccggccant ggcagagcgg tgacgtgccc 1380
cacaaagccc agcatcccg taggcactct gctggggctg ctgaagcggg agatggcgca 1440

```

WO 01/83524

PCT/US01/13862

```

ggggctgttc cctgaggcca agaagcctcg gctcctgcac ggcacctga tcatgaagga 1500
cagcaacttc cggctgtgtt cctcagagca agccctcaaa gagctgggtc tggctyagca 1560
ccagctgggc ttcaactgcc gctgtcacct gcatgacaca cgcaaggagc aggaagagyc 1620
attgctgggc tacagccacc tcaagagcgt cctgaaggac cactgtgtgc agcaocctccc 1680
ggacggctct gtgactgttg agtccgtcct cctccaggcc gcgcgcctct ctgagagccc 1740
aggacccaag gtgctgtgtg tctcctggac ctaccaggac gagagctgg gtagcttct 1800
cacatctctg ctgagagagg gctctccca ggcctccagc tgaaggcggc aactaccca 1860
gcgcgcacct ctgcctcttc ccagctggac agacctggg cctgcacttc aggaactgtg 1920
gtgcctggg tgaacagacc ctgcaggctc catcctggg gacagaggcc ttgtgtcacc 1980
tgcttccca ggcagctgtt tgcagctgaa gaaacaaact ggtctccagg ctgtcttggc 2040
tttattctcg gttaggycag gtggtcctag acagcagttt ccagtaaaag ctgaacaaaa 2100
gaaaaaaaaa a 2111

```

```

<210> 62
<211> 1155
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> incyte ID No: 2738075CB1

```

```

<220>
<221> unsure
<222> 1150-1155
<223> a, t, c, g, or other

```

```

<400> 62
ccccccggc ccccccagg tccggccacg cgtccggccg ggaagaagca cagtggctgc 60
tattatctgc tctccgggcc tgacccctcc caggactcgt gatgcgaag cctctggag 120
oggctacgaa gagtgggggt tgagcccccag ctgagccgag gctcgcact cttctggctc 180
ccccggccca acccaccctga agaaaatgagt ggtggattgg ctccagtaa gagcacagt 240
tatgtatcca acttgccctt ttccctgaca aacatgact tgtaccgat attttccag 300
tatggcaag ttgtaaggt taccatcatg aaagataaag ataccaggaa gactaaagg 360
gttgcattta ttttattttt ggataaagac tctgcacaaa actgtaccag ggcataaac 420
aacaacagct tatttggtag agtgataaaa gcaagcattg ctattgacaa tggagagca 480
gctgagttca tccgaggcgg aaactacttt gataactta agtgttatg atgtgggaa 540
agtggacact taagttagc ctgtccgaaa aatatgctcg gagaacgtga gctccaaag 600
aagaagaaa aaaaagaaa aaagaangct cctgacccag aagaagaact tgaagagta 660
gaagaagtg aagatgaagg ggaagatctc gctcttgaca gctcagta ggcctagta 720
ttccagcaag ccaaaattga agaagaacaa aaaaatgga aacccagttc agggatccc 780
tcaacatcag atgattcaag acgcccgaag ataaagaaa gcacatatt cagtgtatg 840
gaagaactta gtgattaaaa tcttgcocca gcacagtaat aaaaatcag atttggtag 900
aacaaacttg aagagtaaat tttaaaaaa ataaagaaa ttaatactat catgttaata 960
ctatttatgt catcccaaga aaaaagatat tttaaaatt tattgnaaa gttcattata 1020
agggctttat tcatgctgaa tttgtttaca tgaaggactt tgaattaat ctttaaaaa 1080
aacttctga agaccgaaaa gttgaatgat ttattgttac ttatattaat aaactttta 1140
agagaaaaaa nnnnn 1155

```

```

<210> 63
<211> 1673
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> incyte ID No: 2279049CB1

```

```

<400> 63
gttttgggtc gcaatgatgt agaattttga ggtcccttc tgatgaatat tgagctgtcc 60
atgcagcact ggaacccggg ttacagcagt gagggggcca cggclcaaga aacttacaa 120
tgtccaaaaa tgatttgagt ggaagcggcg gaggccacgc ttgtctgatt agacctgta 180
gccagtatgt tccctgggtg gaatgagctc atagtgaaag accagctggc tgaagcga 240
ctgaagatgt gttatgaaaa gaagacaaat gaggggcgat ctccaanagt ctacttact 300
atcaatatga acctggatgt atctgacgaa aaaatggcga tgttttctct ggctgtatt 360
cttcccttta aatacccggc agttctgctt gaatttaact tcaagatcagt attattgagt 420
agatccacgc agactcagct gaacacagat ctgactgcat tcttgcacaa acattgtcat 480

```

WO 01/83524

PCT/US01/13862

```

ggagatgttt gtatactgaa tgcacacag tgggttagag aacacgctc tggctatgtc 540
agcagagata ctcatctctc accacacaca ggaagcacag tccagttagt tgcctctc 600
tccagagac tctgatatc cagcatcat atclataaca aatgcaaaag aaagaatatt 660
ctagagtggtg caaaggagct tccctgtct ggttttagca tgcctggaaa acctgtgttt 720
gtttgtgttg aaggcccaaca aagtgcctgt gaagaattct ggtcaagact cagaaaatta 780
aactgggaaga gaattttaat tgcctatcga gaagactttc cttttagatg tacaaaatgat 840
gaacgggaaa gacaaggaa attttcaat ttggaagaa aagtgttcag tgttaaggaa 900
gcccagggaa accacatgga ctttgggtcag ctctatcagt tcttaaacac caaaggatgt 960
ggggatgttt tccagatgtt ctttgggtga gaaggacaat gacatcaaga gtagttgaaa 1020
gtatcttgcc actgtttggt tttttagttt tttttccaac ttttttttga aagatttaag 1080
aattttattt tagttccatt ctagaattgt ggggagtggt gacacagaaa aatagtata 1140
gctgaaatgc atctgtttaa aatgtcatga ttgaaagcag aactgagttt caaattacaa 1200
ccttaaaatt gttgttagat atttcttcc atactagctg cccattttga aaagaatttt 1260
atccataaag gtaattgttg tgcctcaatt tcccaacccc ctctccctt 1320
acctgcttc actaaggaa ccagaaaagc taattgtctc cctttcagcc tctgttgcaa 1380
ctaaccaalc tcaatggcct caggacacag ctttggcctt ggggaattcty ggaatacttt 1440
taactctga taaagatac atatgcagct aggcacatc ctccccctt taatgcata 1500
aacaccaaag tgaatgactg agctggagga gttatttgaa ccacagagaa gggccaaag 1560
aacacagaaag atgcagtg tccaatgttt gagctgclga cccacaccca gccattgctt 1620
gtctctaaac atcttatgaa ataaaaaccg tttgttttaa aaaaaaaaaa aaa 1673

```

<210> 64
 <211> 584
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2650904CB1

```

<400> 64
aaaataagtc atttgggatg ccgatgttaa aaggagaagg taatcagaga agaaaaagg 60
aagtcgggtg agagtctctg gaagattttg aagatttcta gagaagcagg ggttcttggg 120
gaagggaata gacttggcca gataccagg aagacttctc caacagatgt cccatcatgc 180
tgagattcaa cgcgaatttt tagagtcatg caacatgty agaaaaaag tccagtaac 240
ctttgttggg gctggagggc aggatccoga ygtcccgag gactgtctcc acctcttcca 300
gcaaggagag cgtgtgcttc aggaacttca gcaacactt ctggagccc gagaagggtg 360
ggctcaactg gaggtgctga agaaggtcga cctctcttct caggtcatgg ctgcaacagg 420
atattttcat gcaagcctgc aaagagtgga gatcatgagg agccagggcc ctgtggccag 480
aaatagcccc tgactgaacc ttctgaggaa agcgtttctg actctgccac gaagcaggca 540
ggcaccgata ttggcgcctc tgggaggact cagttttttt cttt 584

```

<210> 65
 <211> 978
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3179424CB1

```

<400> 65
ccggctaacct gttggtgtgt atgcagagca tccctgtgct ccgggatab agactgggag 60
gcctctgttg cgcaggcgca gaactacaac ttcaagggtt tcccaacgg cctctttttt 120
gcacgttagg agaaactaca ttcccatata tctctttgtc cagggtctga ggcgtctgtg 180
gctccggaat cgcgccgagc cggtaactgg ggacccaact cggatatgga tgccttggct 240
ggatccctgt tgggccccac gagttagtgc gcagcgttgc tgggtggcag gttggtccag 300
ccccggcctc ggtgggggtt cccagacgcc tggggccttc ccccccgca gcaggccgg 360
ggcaaggctc gggggaatga gtabcagccg agcaacatca aacgcaagaa caagcaaggc 420
tgggtccggc gcttgagcac gccggccggc gtgcagttca tctctggccg aatgctcaag 480
ggcgcgaagt cgtcagaca ttgaggttcg gtagcagatc ggcgggaccc taatggagag 540
atcgcctcgc cctcgacact tgcctggcgc tatttttga gggagctggg gaggcggaac 600
gcctccgacc tgaagtctct ccatattgtg ggtttgaagt ctggatggga gcttgcgaag 660
tcccttttta ggtcttttaa tttagaagca ttctgaacct cgcgaacaga ccaagaacaa 720
gtacnaagaa catcgtgta ccaagtaacc tgaactacca ctacctnaa cccgtccctg 780
ccccatctg agttcttttg aagctgatct caggcatcgg attatttctt ctgtaaatat 840
ttcagaatgt atctctccaa gatgagagct cattaaaaga caattacaaa gcttatcaca 900

```

WO 01/83524

PCT/US01/13862

tccaaagaa ttatcaatca ttttgaata ttattaaag tgtaataaat gttcaaaagt 960
ccacttgca aaaaaaaa 978

<210> 66
<211> 1055
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2885096CB1

<400> 66
cagctccaga gatggatgga tctgcaatgc catggctggg ggtgttccag ggcagccg 60
aggggtgggg ctggcactga ttgcaactga cagccaggag accagcccg ggaggccagg 120
cccagggtca ggggagagcc tgagtgtctc ccacctcttc atctcagact ttgcatactg 180
ctgggaaaac ttgtgtgca atgaagggtca gccattcatg ccttgytaca aattcgatga 240
caattatgca tccctgcacc gcaagctaaa ggagattctc agaaacccga tggaggcaat 300
gtaccacacc atattctact tccactttaa aaacctactg aaagccctgt gtcggaaaga 360
aagctggctg tgcctcccca tgggaagttac aaagaccac tccagctgtt tccggaaaga 420
ggcgctcttc cgaacccagg tggatctga gacctatg calgagaaaa ggtgcttct 480
ctcttggttc tgtgagaca tactgtctcc taacacaaac tacyaggtca cctgttaaac 540
atcttgagac ccttgcccag agtgtgcaag ggagggtgac gaggttctgt ccaggccag 600
caacgtgaat ctacacatct tcccgccag cctctgctac ttctgggata cagatttacc 660
ggagggtctc tgcagcctga gtcagggaag ggctctcgtg aagatcatgg gctacaaaga 720
tttctatctt ttttggaaaa actttgtgta cagtgtgat gagccatca agccttggaa 780
gggctacaaa accaaacttc gacttctgaa aagaaggcta cggggagatc tccagtggg 840
gtctctcttg ggcctcatgg tctgtctctt ctgacctct gctcaalgct ccagggtctc 900
ccctccatcc tgcacacagct gtgcttttgc ctgctctac tggagccctc ctggcctcag 960
ggcatttcaa laggccccc ctgctctacc acctactctc cgtctctcca ggttcttctc 1020
gcaggagcct ctctctgctt ccattggctat cctac 1055

<210> 67
<211> 2220
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2901076CB1

<400> 67
cggctccgtc gctgacgcgt cgtgacgctt ggggagcggg aaggcaacgg cagcgggato 60
gggatgaaca gggcgggcgg ctctgctttg gctcttagct tgggctccac ccccgctgag 120
gtgattcagg tgacgaatct gctctcggcg gtgacccagg agcagatggc gaagcttttt 180
tctctctcag gagaalloga gtagctggcg ctctaccctc cggacaacgc acctctgtgt 240
tttctctcaa aagtatgtta tgttaagtgt cgtgacccat caagtgttgg cgtggccag 300
caatcaacta acacgggtttt lattgaacga gctctgatag ttgttctctg tgcagaaggt 360
aaaaaccag aggaatccaa agccctctct tlatgtgctc ctgctccaac catgacaagt 420
ctgagtccctg gtgcaggatt gcttccaaia ccgaccacca atcctttgac tactcttgg 480
gttccactta gcaatttggg agctalacaa gcagcagcac tagacccaaa cattgcaaca 540
cttggagaga taccacagcc accactatg ggaacgltgg atcctccaaa aatgagatga 600
alttgaggaa cgttttatgt tggaaatctg aaltccaga caacgcagc tgaatcaact 660
cttgaalttt ttaaaccaagt tggagaagtg aagtttgtc ggaatgcagg tgaatgaact 720
cagccaactc ggtttgtgtt tgtggaattt gcagaccaaa attctgtacc aagggccttt 780
gcttttaatg gagttatgtt tggagacagg ccactgaaaa taaatcactc caacaatgca 840
atagttaaac cccctgagat gacacctcag gctgcagcta aggagttaga agaagtaagt 900
aagcaggtac gagaagctca gtcatttate tccagcagcta ttgaaccaga gctctgaaag 960
agcaatgaaa gaaagggcgg tgcattctgt tccatactac gctcaaaatc caggtctagc 1020
tcaaatccc attctagaag gaaagatca caatcaaac accaggatag atccataat 1080
agatcacgtt caagacagaa agacacagct agatctcaga gccacataa aanaagctct 1140
aaatcaaggg agacacagaa gtcaaggagt cgttccgcat caccggacaa gagaanaagac 1200
actcgagaaa agatcaaggaa aaaggaagaa gtgaagagaa aagacaggga aaaggaagaa 1260
gagcgggaaa aggaacgtga aaagaaaaag gacgggggta aaacacaaag ccgggacaaag 1320
gaacgggaaa aggaacggga aaagacaaag gaaagagaca gaggagagaa accgggaaaaa 1380
gagcatgaga aggatcgaga caaagagaag gaaaggaac aggcacaaag aaaggaacga 1440
gaaaagaca gatccaaaga gatagatgaa aaagagaaga aggataaaaa atccagaaac 1500

WO 01/83524

PCT/US01/13862

```

ccaccacagg gttacaatgc atcgcgagga tctcgtagtt ccagcaggga aaggcgtagg 1560
agggaggagca gtagttcttc cagatcgcca agaaccatca aaaccatcaa aaggaaatct 1620
tctagatctc cgtccccag gagcagaaat aagaaggata aaaggagaga aaaggaaagg 1680
gaccacatca gtgaagaag agagagagaa cgttcaactt ctatgagaaa gagttctaat 1740
gataagagatg ggaaggagaa gttaggagaa aacagtaact cacttaaaaga gaaagagcac 1800
aataaagac cagattcaag tgtgagcaaa gaagtagatg acaggatgc accaggact 1860
gaggaaacaa aatatcacga caatgggaat tgtcagctga atgaagaaaa cctctctacc 1920
aaaaacagag cagtatagga ccgacaagtg taactctgca ctcaatgctg gaatacaatc 1980
caaaagctttt aattctctca acaagatgta aacaggaaag aaatctagtt gagcaatgaag 2040
ataggatctca acagcttttc cagttgttag atgaacttgt ggcactcttg ttatttgatg 2100
agaaaataaaa gcatggacat catgaaaata acagatgtta cccaaactca tctttotaaa 2160
tctgtgcatt tcaatgggtg ctgacacact tgcatalgtg tctgttagtg ttggcaaga 2220

```

```

<210> 68
<211> 1890
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3074572CB1

```

```

<400> 68
ggcgggtgcc ggcgggggcc agcgcctttt cggccgcag cggcgcttg gctcccgct 60
gtttaaaagt gcgctgttgg ctgctgctgt cttaactct gtgcttggcg gacagacagg 120
cyagatggcg gcggaggtgt tgcgagtgcc gaggtagcag taattgtggg cggccgacgg 180
gagccagaga gctgtactgg tccagttctc caacgggaag ctacagagtc caggcaacat 240
gcgcctttacc ttgtatgaga acaagatgt caccacccc aggaagagga altcaacggat 300
cctggcgact gaacagata ggtctctcta tgtgggaac aaatttggga ctggagccct 360
caaatgcacac actttgtgca ggcactttgt gggaaatttg aacaagacct ctggccaat 420
ggaagtatat gatcgtgaat tgttcaatat gcagcacta ttttcagatg tatcagttga 480
gagtgacatg gcgctagaga gtacagacca aacttacaga gaaagatgg attcttgtat 540
tgaagccttt ggtacacaca aacagagcg agctctgacc accaggagaa tgaacagagt 600
tggcaatgaa tctttgaatc gtgcagtggc taagctgca gagactatca ttgatagcaa 660
gggtgtgact gctctgttca gcatgctat ccacactgac ttgcuagntg actcncctca 720
ccttctcccc ttgtatgat atgcagcaca gccctgaagc gtgtataaat ttgaagntct 780
tccttccccc gcggagtatg aagctcttca gagccatct gaagcttcca ggaacgtcac 840
gttcgaagaa atactgaaga tgalbaggaa gaacagccat tgcaccttg tcatagaagc 900
gttcgaagtct ttgcataag atgtggagag ccagagccgc caggcccgat gcataatggt 960
tctggatacc ctcatcaaat ttctagatca taggttagtt aaggcgaaaa gtgctctggg 1020
acttggagtt cccacacatca tcaaacacca actgctgaag cactttactt gcttgacta 1080
caacaatggc agattacyga acttaatttc ggtttctatg aaggcgaaaga ttaactgcata 1140
tgtatcatata ctgtgcttgc acatacatga ctctcaaatg gactgacag tgttacagag 1200
ggacttgag ctgactgaga aaagattgat gtagatagcc aaagcctgga ggtgaagat 1260
ctccaaaaga aggtgtgtgt tggcgcgcgg cagtgaagaa gatcaaaaac tgggcacct 1320
gtccctcccg ctgcctccag ccagacacct agccgccttg gcaagcaga ggaagattac 1380
ctagacgcat gctttccaga caggcgcttt tggctgcata acagccactg gctggctcta 1440
tcaatttcca tttttatgta tgttttgaaa agaaaaggtc cggggatggt ggtctacacc 1500
tgaatccca gcaatttggg aggcgcgggc aggaagatca ttgagctcag gagtttgaaa 1560
ccagtctgga caacataggg agaccccatc tctaccggag gaaaaaaaa agagtccggc 1620
ctgggtgtgt gcgcgtgtaa tccacgtcac tgggaaggt gaggcaggac gattactaga 1680
gcttgggaaa tcaaggttgc agttagctat gatgtgtg coaacctca tcttgggta 1740
cagagttaga ccttgtctca aaaaagtaa cataaggaaa aaagaagct tgcctttagc 1800
caggtatgaa gccagagcc agcatctcaa ctgtgcttgt ctatgcaga aatataaagc 1860
gatggccagg ttggaattca aaaaaaaaaa 1890

```

```

<210> 69
<211> 2893
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1437895CB1

```

```

<400> 69
aattgggctc accaggtatg tccaggataa tcttccaate tcaagtgtgg tttattgaca 60

```

WO 01/83524

PCT/US01/13862

```

atcatttaca atgcggaaga gtgctgtagt gagccagcac agtgggtaac acagcaacgg 120
agaaacagatg caggttttagg gaatttaact tgcataaaccc ttgaactgaa gtcttagaga 180
ttggaaacata cgggttttcta taaataggct ttttaagccct gtttgcattg ggtaactgat 240
aggagaacact tgccttggga atgtcagctg cgtgagctca ctgtcagaca agatgggaaga 300
agaaaggctg gagggtccaa actcttctc tgaaaaaacg tatitctctg aatccctgga 360
ttccagcgat ggggagaggg aagaggttct ggctgtgag gatttggaa ttaacccct 420
tgatggattg ccatattcat cactgtatta taaacttctg aaagaagag aagatcttcc 480
tataatggaa gaaaaatact cctttatgga gaacctgtct caaaatcaaa tctgtattgt 540
ttcaggagat gctaaatgtg gtaagagcgc tcaagttcct cagtggtgtg ctgaatattg 600
tctttccatc cactaccagc acgggggcgt gatatgcaca caggtccaca agcagactgt 660
ggtcacagct cccctggcgg tggcggatga aatggatgtt aacattggtc atgaggttgg 720
ctacgtgac cctttcgaga actgctgtac caacgaaaca atcctgaggt attgtactga 780
tgatctgtg caaagagaaa tgaatgcaca tcccttttgg ggtagctatg gggtaact 840
cttagatgat attaatgaaa gaagcattgc aactgatgtg ttaactggac ttcttaaga 900
tgttttacta gcaagccag aactgaagct cataattaac tctcactctc aactgatcag 960
caaaactcaat tcttattatg gaaactgtgc tgcataagaa gtgaaaaata aacacctgt 1020
ggaggttctg taccttagtg aggtcaaaaa ggtattcttt gactctattt taagccttat 1080
ctttgaaatt caccactcgg gtgagaaagg tgacattgta gtctttctgg cctgtgaaa 1140
agatatttag aaagtctgtg aaactgtcta agaggaatgt taattgttca agcaacttga 1200
agaaactgtg gtgttctctt tgaatccaaa agagaatgtt tcatgttcca agcaacttga 1260
tgaaacagaa aaaaagtgtc aggtttatca aagagaggtg gtgttaacta ctagctctgg 1320
agagtttttg atctggagca actcagtcag atttgttata atgtgggtg tggaaagaga 1380
aaaggtgtac aaccogagaa taagagcaaa ctctctctg atgcagccca tcaagccag 1440
ccaggtcagag alacgcaagc agatctcttg ctoacttctc loaggaaat ttttctgct 1500
gtacactgaa gaatttgctt caaagacat gacgcaactg aggcagcag aatgcagga 1560
agccaacta caaagcattg gtctttttat gaagaggata gaacttctgg gcttaggca 1620
ctgtgacttc atgaacagac cagcaacaga aagtttgatg caggaacttg aagacttga 1680
ttactgtgca gcaatggata atgaggaata ttttctgaa ttgggaatca tcatgtcaga 1740
gttctctctt gatcccaac tctcgaagtc taacttagct tctgtgaa ttgactgtgt 1800
agatgaagtg caaacactcg cagcactggt nacagctcca aatltgttct cactgtgct 1860
aacgtgagct gaagaggtg ccttgacttg ttggaagaaa tttttacatc ccgaagagga 1920
tcaacttacc ctacatagca ttlaaagtc ttaccaagac acaactctga attctagcag 1980
tgatgtactg gtggaaggt gggtctgtga ttacttctc aactgttcag cactcagaat 2040
ggcagatgtt attcagctg aactcttata aattatcaag cgaatcgagc ttccctatgc 2100
agaaactgtc ttgtgtctca aggaacacac cotaacata aagaaagctc ttctgtcgg 2160
ttactttatg cagatgtctc ggaatgttga tggatcaggt aactacttga tctgacaca 2220
taagcaggtt gctcagctgc actccctgtc tggttactca atcccaaga agatgcaga 2280
gtggtgtctc ttccataaat tcaagatttc tgagaacaaa tacaatcaga ttactcaga 2340
aatctctctt gaactattta tgaagctgtt accacaatac tatttcagta alctgtctcc 2400
tagtgaaagt aaggaacttc tacagcaagt agtggatcac ctatccctg tgcacaacat 2460
gaataaggaa cagcaaatgt gtgagcgtg ccttgaacat gaacagagat gcactctcca 2520
gtgactctcc agcaaacaca aggtgcagca ggggtccaaa ggtagctgga tgggtgaact 2580
gctgtgatg ggaatcacat gacgcaaga cggatttccac atccacagga cgggtctgaa 2640
gaataacaca ctgtgtatct ttttttaaaa taataaagag aagtttttat tgggtcttct 2700
aaattactac tccatgtctt ttttctctct ggaaggttct taaatcaac cactcataak 2760
ttgcacaaaa ttttaaaaaa ctggtatttt gtaaatgtgt cagagacana tgggacagaa 2820
cctacttttt tgttagaggaa cttaattotga ataaagtctg agtttttcag taaaaaaaaa 2880
aaaaaaaag aag 2893

```

<210> 70
 <211> 885
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1454656CB1

```

<400> 70
ccagcagtcg ggcgccatgt aaccgggtcc gtgccgcaaa cgcgaacggc gccgcggcgc 60
ggggcccgcc ggggttagag gtaaccatgc tgagggtcgc gtaggagacg ctgagtttca 120
lccggaccgg ggcagttacc caggtctctag taccgggtgc cccgggcggt gggagcgcca 180
agtttctctt caacagtggt ggcctgcagc ctggaagtct cctcctcag gccgcgcgcg 240
gatattctgt ccggaaccca gccagctcta ggcgtgata tgaacacct ccttctacgc 300
tgctcaagaa ctacagaaat gtccctggan ttgagaaggt tgatgatgtc gtgaaagac 360
tcttctcttt ggaactggcc aacaagaagg agatgtcaaa aatcaagcaa gaacagttaa 420
tgaagaagat tgttgcaaac ccagaggaca ccagatccct ggaggctcga attattgct 480

```

WO 01/83524

PCT/US01/13862

```

tgtotgtcaa gatccgcagt tatgaagAAC acttggagaa acatcgaaag gacaaagccc 540
acaaacgcta tctgttaaty agcatiganc agaggaaaaa gatgetcaaa aacctccgta 600
acacaaacta tgaigtotftt gagaagataat gctgggggctt ggggaattgag lacacottcc 660
ccccctctyta ttaccgaaga gcccaacgcc gatctgtgac caaggaagctt ctgtgcattc 720
gggtttttca ggaagactcaa agcttgaaga agcgaagaag agcctlaaag gctgcagcag 780
cagcccaaaa acagcaaaa gggaggaacc cagacagccc lgcacaagcc ataccaaga 840
cactcaaaag cagccaataa attctgtttca atcattttaa aaaaa 885

```

```

<210> 71
<211> 1269
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 121130CB1

```

```

<400> 71
tcagacaagc actggacgtg gggggcattt tgttttggac accgagcagg agctggcggc 60
cgctgcagac gaaaggcagg aaagggcagg ccgggtgagc agacggatcg gccgactaga 120
cagccaaaca gaaacaaga actgagctcg catambaccy cttaacgcatc taacaaaccg 180
ccalcetage taacocgagc cctccacccg tcaactcagg ttccggcgggt ccccgccccc 240
ccctccggag ccgtgggtgc agcccccggg ggaagcactgg agtctgttcc ctccgatttc 300
cggtgatctga agatgggtgc acagtcaagc cgaagaattg tgcataaagc caccaccaaag 360
atgctctcaa atgagcgttt tactaatatg ctgaagaaca aacagccgac gccagtgaa 420
atlcgggctt cgatgcagca acaacagcag ctgcccagtg ccagaaacag aagactggcc 480
cagcagatgg agaataagac ctclgtccag gcagcattaa aacttaagca gagcttaaa 540
cagcgctggg plaagagtaa catccaggca cggllaggcc gacccatagg ggcctggcc 600
aggggagcaa lcgagagcag aggcctaccc ataatcaga gaggcttgc ccagagagga 660
clacgtgggg gacgtgcac cngaacccca cttagggggc ggaatgcaat ccgaggtcaa 720
aaactgtccc gaggtgagc agcctagct cccgaatgg gcttaagaa aggtgggtgt 780
caggtctgtg gaggctcgtg gaggggggc ctaggccgtg gactatngg tctgtggcga 840
atcgglggtg gaggctgggg tatgataggt cggggaagag ggggctttgg aggcgcagcc 900
cyagggcgtg gccgagggag aggtgcccct gctgcacctg taatgacaa ggagcagctg 960
gacaaactaat tggatgcata tatgtcgaaa acaaaaggac acctggatgc tgagltggat 1020
gcctcaatgg cgcagcaga tcccgaaac aatggttga gactgcacat ctcccaiga 1080
ggagactcttg ttagtcaca catctgtaaa taactttgag ataacagat agaagaaatc 1140
tgatgtatgc tggatggacc tatcaacata ggcgttggac ttacttgcac ccagctctgt 1200
caatttagtct gttcccttta ctttttgata ctgtgttcta tgaacccctt ctgtcctttg 1260
aaaaaaaaa

```

```

<210> 72
<211> 1066
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1257715CB1

```

```

<400> 72
cggctcgagg tgaatggggg cagcatgagg ccgggcggct ttttggggc cgaacagcgg 60
ctgagtagag ccatgagccg algtgttttg ggcctcgcc ccccggggaa gccgtgagtg 120
gtggctcgcc tggggaatcc cggactgccc ggcacggcac acagcgtggg catggcggtg 180
clggggcagc tggcgcggcg gctgggtgtg cgggagagtt ggaacgcgga ccggcactgt 240
gccgcagacc tgcgcctggc ccgcgtgggg gatgccaac tggctctgct ccggcagcgg 300
cggcttatga acgccaacgg gccgagcgtg gcccgggctg cggagctgct tgggtgact 360
gccaggaag tclacctggg gcaltgatag ctggacaagc cctgggggag actggctctg 420
aagclggggg gcaatgcccag gggccacaa ggaatccgtt cctgcattag ctgcctcagc 480
tccaatgcaa tgcgaagctt gccgggtggg atccggagcc cggccgaccc tgaggcgggt 540
caggcccatg tgcctgggtg ctctccctct gctgagcagg agctgtgtcc ctgtgtgtgt 600
gatcgagcca ccgacctgat ctggaccac ctccgtgagc gaagccagg gccctcactg 660
ggcgcgtgac actagtggcc atgggtgcct gcctgactgt agtgcacac aacccagcca 720
ctgcacaga gctgcacagc cagccttggt atctactttt tatcaaatc tccctagac 780
tgttccagcc tgcctgcgga ttaaagtggg ggtgactgtg actggacag tccattctgt 840
gagtaggttc tctctctgt gtccacttg ggaactaggg gaacttcagg aagactaac 900
ttttcaagcc tttttagaga accaggggca cgcattcttc ctgggttggg caatgggact 960

```

WO 01/83524

PCT/US01/13862

```

gtgactccttg gtggggacac gcagcctttt gaggtctctg ggcacacgtg gagctgagca 1020
tgaccagcag ttgtctgcagc atctccttgt gccatggctg gaaact 1066

```

```

<210> 73
<211> 639
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1342022CB1

```

```

<400> 73
ggggagacac gtgcctcttg tactatgacc actagaccag oattcatatt acacacag 60
gactctctct cgagccgctc gagccgaatt cggcacgagg gactctggag acgacgtgca 120
gaaatggcac ctggaagagg gaaggaaaag aaggaaagac aggtctacag cctcggacct 180
caggtggctg aaggagagaa tgtatttggg gtctgccata tctttgcac cttcaatgac 240
acttttctcc atgtcactga tctttctggc aaagaaccca tctgcctgtg gactgggtgg 300
atgaaggtaa aggcagaccg agatgaatcc tcaccataty ctgctatgtt ggctgcccag 360
gatgtggccc agaggtgcaa gtagctgggt atcaccgccc cacacatcaa actcggggcc 420
aagaggagaa ataggaccaa gacccctgga cctggggccc agtcggccct cagagccctt 480
gcccgctcgg gtafgaagat cggcgggatt gaggtatgca ccccatccc ctctgacagc 540
actgcagga aggggggtcg ccgtggtcgc cgtctgtgaa naagattcct caaaatattt 600
tctgttaata aatgtccttc atgtaaaaaa aaaaaaaa 639

```

```

<210> 74
<211> 1420
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 194704CB1

```

```

<400> 74
ggcgcagcgcg acacatcttt gtgcagcgcg ctgcacacgc ctgcctgaga gaagtcgtcg 60
cgccgcaccc cgtgcctccc gccgcctacc atgtccgccc aggcgcagat gcgggcccctg 120
cggacccagc tcatgggccc ggctcgggac ggagacgaaa ccagacagag ggtaagtttt 180
acagatgacc gtgtctgcaa ggtcaccctt ctggactgct gcccccata caatcctgct 240
gggacgcgca tggatttagg agaattgacc aaatccacg acttggccct ccgagcagat 300
tatgagattg caagtaaaag aagagacctg ttttttgaat tagatgcaat ggtacacttg 360
gagtccttta ttgctgaaty tgatcggaga actgagctcg ccaagaagcg gctggcagaa 420
acacagggag aaatcagtyc gaaagtctct gcaaaaggcag aaaaagtaca tgagttaaat 480
gaagaataag gaanaactct tgctaaagcc gaacagctag ggtctgaagg taatgtggat 540
gaaccccaga agattcttat ggaagtggaa aaagtctgtg cgaagaaaaa agaagctgag 600
gaagaalaca gaattcttat gctcgcctcc agttttcagc agcaaaagct gctgtctctg 660
gaagttctgt cagctacact tggctctccat gacaatgacc gtgcctgggc agaccacttc 720
ggtggcaagt tacacttggt gttcattcag atccgagaga agcttgatca gttgagggaaa 780
actgtcctgy aaaaagcagg gaagagaaat caggatcgtc tgaggaggag agaggaggag 840
gaacggggag agcgtctgag caggaggtcg ggtatcagaa ccagagatcg caggaggtca 900
cgtctccggg atcgggctcg gaggcagtca agatctact ccagagagcg acggaaattg 960
tcocggctcc ggtcccgaga tagaacctcg cgcacccgca gcctgtcccg gagcacacagc 1020
cggggacatc gtccggcttc ccgggaccca agtcgcaaat caaagtcttc cagagagcgg 1080
gcattcacag agggagctct ggagagcggg cggagcgagc gagggccccc ggaactggag 1140
cttgagagct ccaacgggaa gatggtctca cggaggtcag aaggaagga ggcgggcgag 1200
atctgaaccc gtcctccggg tgcgttaaat agttctgata acgttcacac agttctaaat 1260
tacctcttat atttgctgaa tacaaactat ctttttagt ttaaaatttc tattgttttg 1320
gagctagctg tgagttctta gaagtgtaca gactgtctcc tgtgttcccg ggtcatgttg 1380
agtuggaata aataaatctg atgtgcacaa aaaaaaaa 1420

```

```

<210> 75
<211> 1457
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature

```

WO 01/83524

PCT/US01/13862

<223> Incyte ID No: 607270CB1

```

<400> 75
gggcatttag agcttgccg gctcttaggc cccgccccct caccctccg gtcctggagc 60
tcccacagct aacatggcgg ggcctgtgtt gttctacagg gttacccggt 120
ttctctcggg ggttagggta gctctggccg gaagcaaggc ctctggaaaa ccgctggccc 180
tgagttgcga acaaatgtca gctcccagat attaaagcta agacatacty catttgaat 240
accaaagaaa aacgttccta cctcaaaacy tgaacttac acagaggatt ttattaaaaa 300
gcagattgaa gagttcaaca taggaaagag acatttagcc aacatgatgg gagaagatcc 360
agaaactttc actcaagaag atattgacag agctattgct tacottttcc caagtgtgtt 420
gtttgagaaa cgaagccaggc cagtaatgaa gcacccctgaa cagatttttc caagacaaag 480
agcaatccag tggggagaaag atggccgtcc atttccactat ctcttcctata ctggcaaaa 540
gtcaactat tcaataatgc atgatgtata tgggaatgta ctcaattag aaaaacatca 600
aagtcaactg caagcaaaaa gtctgtctcc agaaaaaact gtaaccagag acgtgattgg 660
cagcaatgag ctgattaaag aggaactaga agaaatgtta gtgaaaaaac tgtcagatct 720
agatttatag oagttcaatc ggtgtctaga aaagttaact acatgcaggt gttgtgtctg 780
tgagggaagaa ttgtgcaga ggtttcgaaag aagtgttaact ottgaatcaa aaaaacagct 840
gattgaacct gtacagtatg atgagcaagg aatggccctt agcaaaagcg aaggtaaaag 900
aaagcagatga aaagcagaag caattgttta caaacatgga agtgggaagaa taagaataaa 960
tggaaattgat taacagcttt acttcccgat- caaacaggag agagaacagc tgaagtctcc 1020
tttccacttt gttgaacggc tgggaaagca cgaagtgaat tgcacagctc cagggggggg 1080
gaggtcagcg caggctggag caalacgaat ggcaatgcaa aagaccttgt gaaactttgt 1140
caaccgagaa caggtctgagt ggaatgagaa agcttgacta cttactactg atccacgtgt 1200
aggggaaggg aagaagccag gcaagagagg gcaaccgaga aagtttaact ggaagaaagc 1260
ctaagggttt gctcccagga aaggagagga agagctatata atatgtgcgg aactgtggca 1320
gacacacagt aaataatggc tgaccagcat gagggcagta ctgtcagaaa ttctcttgag 1380
ctgtgagatg gatltatgtt taatatgtac ttgttaaggt tgacctttaa aaataaaaag 1440
gaaaaaaaag aaaaaaaa 1457

```

```

<210> 76
<211> 1184
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 758546CB1

```

```

<400> 76
caggccgtcc aggtctctggg gcgcgcgggc ggaatccgcy cggatgcag aacgcgtctc 60
cagcttcggg tccctcgggt gcggctgcgg ccatcatggt gcggaagctt aagttccacg 120
agcagaagct gctgaagcag gtggacttcc tgaactggga ggtcaacgac caaacctbpc 180
acgagctgcy cgtgtctgcy cyttacccgg tcagcggcgg ggaagactac acgctgctaca 240
acacagctgag cgtgtccggt cgtgagctgg cgcggcgccg ggtcgactac ccgcaagcgg 300
accagttccg cgtgtccgct tccgcgcgcg tcttggacaa gctgtatgct ctccgcttgg 360
tgcccccagc cgttctgcty gactctctgg acttctgcaa ggcctctgct tctctgcgcc 420
gcgcctcccc caccgtgctc ctcaagctgc gcatggcgca gcaacttcag gctgccttgg 480
cctttgtgga gcaagggcac gtacgcgtgg gccctgacgt ggtttaccag ccgcgccttc 540
ttgtcagcgy cagcatggag gaacttctca cttgggtgga ctctgtccag atcaagcgcc 600
acgtgtctga gtacaatgag gacgcgcgat acttctgact ggaagccctag cggatctccc 660
actttgcatg gctgtgtttt acagatggga aaactgaagg ctgaagctgg agaatctatg 720
aggtgtctct cctcaagggt atcagaaggt cgtaggttct taagaatttg attcaatcag 780
ggcaggcact gcataagacc acgggaggtg cgtctctgtt ttccaggaaa tgttctbaga 840
acttggacta ctgattatta aktgactgtg ccttgggaaa cagtgggaag taacttgttg 900
cagcactggg gtattgttgg cttctgtgtg tggaaacttt gtaactgaaa aggaaaaaact 960
ggaaatcccc acgcoctggt tccctttatc gtcttctgtt tggactgggt caatttgttt 1020
aaactgaatt cttgtctctg gccctgggtt agctgtgtac agatgatgga gagtttgtgc 1080
taaaattttt ataaactgag cgaagactag gttaaggata tectcccttg tttaaatgtc 1140
aataaatgcc ccaactgctt tgttaagtga aaaaaaaaaa aaaa 1184

```

```

<210> 77
<211> 1638
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature

```

WO 01/83524

PCT/US01/13862

<223> Incyte ID No: 866043CB1

```
<400> 77
atcggggagc ttggcccccagc cagagggtac agtggccccc gaaggagcct caagtcaccc 60
tcccacatcaa agagccctctc ttttctctctc tctggagcag ccattgttcca gccagccaca 120
tgcccctggc agctgcccgc ttaagcaag taanaacttc caggaaacttt ccaagtcac 180
ctttccgtgc tcaaggtgagc tctgaaaccc tagtaaaat ggcagctctt ttccagagaa 240
aaaatgtgag gactgtgtgg taccctatac tcccagaaga ctanagacag gccagccatt 300
aagcacggag acagggcaagg gtaagacagt ggcagccacag gggcccccct caggggcgtgc 360
cccagccctc ctctacgtgg gccccggagc gctcagattt attcagccgt atttgaatag 420
ccattataaa gaaacccacag ttcccgggaa agtgcctttc cacttgagag gccacagggg 480
ccctgtaaac acaattcagt ggtgtccagt cctttctaac agccacatgc ttctctccac 540
ttctatggat aaactttcca aggtatggag cggcctggag tccggggcact gccctgcagac 600
ctactccttg ccaacagagg caatgcgggc cggccgttgg gctccctgtg gccggggcat 660
cctcagtggt ggccttgnct tgcgcgtgca cctaacagac ttgaaacag gaaccacagt 720
attcagtggt cgaagtgcct ttgaatcac taccttgaaa ttccatccaa aagaccacaa 780
catcttttta tctggaggct tcagctctga aatgaaagct tgggataata ggaactggcaa 840
ggtgatgaga agctacaagg cgaacctcca gccgaccttg gacatcctgt tctccgggga 900
aggctccgag ttctgagca gcacagagc ttccaccccg gactcagctg accgcacct 960
tattgcccgg gatttccgga cctctgccaa aatctccaac cagatcttcc accgagaggtt 1020
caactggccc agcctggctc tgcacccggg agagccctgt ttctgggac agaccaatgg 1080
caactacctg gccctttctc ccaactgtgt gccctacccg atgagagagc ggaggggca 1140
tgaagggaac aaggtggagg gctactcagt gggctggagc tgcctccag gccgtgactt 1200
gatgttgacg ggcagcgccg atggccgggt cctgagtac agcttccgca cagccagccg 1260
agcagtcaca ctgcaggggc acacacgggc ctgtgtggc accacctatc accccgtgt 1320
gccctccgtc ctgcacacct gctcctgggg aggggacatg aagatctggc actgagcttt 1380
ttgtcacgca accctccga tgcagctgg gctcttggac tccctcttc ctcaagggtta 1440
gatgagagga acgagcacag aggttggctg tgggtcctgt gtaccacct ctgagcccca 1500
gttctcctat gctaaagtg gggagaaag tctgttgc tccagagtg gaggactaca 1560
ctagtgaag cgcctggcgg gcagccggcg atgcccata aatgtgtgt ttgctgtttg 1620
ttaagtgaaa aaaaaaaaa
```

```
<210> 78
<211> 701
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<223> Incyte ID No: 927065CB1
```

```
<400> 78
tcacgcttcg tggggcgagg cagggagag ccaaacgtaa agacaccagg agtttctcgg 60
gcccngctgt gctctgtccc ggggagcccc aagccttggc ggtctcttc gccgaatagg 120
agctctgtca ggcctcaggc tagtccgacg aagagtgggt gtgacagca ctggaaagaa 180
tgccctgccc tgcctccaca tatgaagag tagtttcaaa aaacctctcc gagtaccact 240
acatgaaggt ctgctagaaa ttccaagatt gtggagttgg actgaatgct gcacagtcca 300
aacagctgct batttccgct gtgaaggacc tgtttgggga ggttgatgcc gcttacctt 360
tggaactcct aacctatgaa gagaagacct tgtcagccat cttgagaata tgtagcagty 420
gtcttctcaa atgtggngc tctttgacct tgttaaggat cctattataa ggcaaaaaat 480
gtgtcttccc ggtgattcag gttctccat ttctctctgc attatctgt aatagttagg 540
aacctagatt ggaatgaag aatagctctc ctttttggaa acgttcacac actctaatat 600
ttattttttg gtgcctgca tgbttgaaga ctgaaagcag gctaaaagct cttgatgaaa 660
tttgagggtg ctgaaagatg ttcccactaa ttccagcca t 701
```

```
<210> 79
<211> 1829
<212> DNA
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<223> Incyte ID No: 938071CB1
```

```
<400> 79
gggtttttga gaagtaccca gaaatgtgtc caaggtttcc tcagatttgg gctgttccc 60
agcggcaggt cccgggaacc aaggaaacag acatcttcc aggtctgaga gagcgcccc 120
```

WO 01/83524

PCT/US01/13862

```

ttgtccacag gctgctgggg ccccccagta gccatggctc cgggtgctgg ctcacgcagc 180
ccggataggg aggcctcggg ctcgggggga agacgtcgca gttcgtcgaa gagtcggaa 240
cccagcaaat ctgcccgctc cccgcggggc cgcgctcttc gctcgcactc ttgctctcgg 300
tcgggggacc ggaatggact caccatcag ctgggtggcc tcagccaaag ctcgccaaac 360
cagtcctacc gctcagctc ggggtcggt tctagagagc ggccctctgc gccccggggc 420
atccctctgc ctctgctc ctgctcagtc tcttagagca gctactcggy cccctaggg 480
agcgaacagc cttggcttag ctctctcgac aagggaaggg agggagggct ggggcaagaa 540
agattaaagt acagagagagc aattggagaa ttgggagctc ctgaagtatg gggactttct 600
ccaaagaaac ctgaaccaga ttctgatgaa ctacacagc tggaggatga agagccaaag 660
aaaagcaata ctacagcttc tacttcagaa gaagaaaaaa agaagaagtc tagcggttca 720
aaagaaaggt ccaagaaaag gagaagaaa aaatcatoga aaagaaaaca taagaaagat 780
tcgaaagata ggcacagtga ctctgattct gaacacagct ccagtgtatg agataaocaa 840
agggagagaa agaaagccaa gaaaagggaa aagaagagaa aacacagctc gaagaaatct 900
aagaaaagaa gctctagaa gagcagaaaa gagtcagtg attcagctc taagaaatcc 960
caagaaaggt ttctggaaaa tccctggaaag gatcgaacaa aggtcgaaag accatcagat 1020
ttaattggcc cagaggtctc aaaaacactt acctctcaag atgataaac ttgaaagcat 1080
cgccgaatgg aggcctgtgc actcgaaaaa gagaaccaga tctacagtc tgatgagaag 1140
agugcccttg catcctttaa ccaagaagag agacgaagaa gagagaacaa gattctggcc 1200
agttttcagc aaatgggtta cagaagagcc aaagggaaag atgacaaata aagattttct 1260
gattgtccag agacatttt taacaacaaa aaagaaagtc tgggttccac acatacatag 1320
aaaaagttta ttatgttctg agaaagcttt acagtgctac tggccctct atttaattct 1380
ttcagctctt caataaaaa ctgcttatct atataactt agcaagttct ttggttatct 1440
ttggttatct catagtaact ttctggttta aaaaocaaa ttatgggtcg ggcagcgttg 1500
ctcagccctg tagtctcagc ctctgaaat gctgggattg cagggtgtgag caacgtgtcc 1560
tggcgttttt tcttaaggtt atttgatctg cattattatt acatgctcat gatnaatttt 1620
taattacccc tgggtataaa agggccttcc gattatccta ttgggaaaaa gcccgcttgc 1680
cttatatttt taagtgttg ttcttcaaaa gtgtttaaat aaggggcgcc atatttcaaa 1740
gtattggaca aaagttttt ttaattataa tttttggaga cgggggtctc ctctgttacc 1800
caggtctagag ttcaagttag cgagatctt

```

<210> 80

<211> 2541

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> Incyte ID No: 3295984CB1

<400> 80

```

caagaagagc ggganaggat cggaaaaaga agctaaaaa ctatagaaaa ccatgagatc 60
taltoalat ttigetaaag atgategcca lgttalgggt aaacattcaa caatctatcc 120
alotcggagc gaacttgaaag ctgttcagaa tatggatctc actgttgaat gtgctcttaa 180
aaatgtctca gattgtgtgg atgaaacaaa taaggagacc aaacagagg gtgagacaga 240
agtgaagaaa gatgagcccg gagaabaata ttccaaagat caagtggttc gacatttgtg 300
tgggttaalg aggtttggcc tggttgcaaa aggtctgtct attaaagatg atatggactt 360
ggagctgggt tttaattgca aagacaaccc cccagagacc ctgttaataa cagtcacaga 420
taattctctt attcagatc agaaactcac agaaagagaa tatcaagtug aacaatgtgt 480
aaatgaggca totattataa ttccgaatcc aaagagagcc acgtcaactt tgaaggtgat 540
acttaactca cctclaatca gggacgaatt gggagaagag gatggagaaa atgtttctag 600
gaagatctct ccggaactat tggacagcca gaattgcctg aacgccttgg cgtctctctg 660
caatgcacaa tggtttcagg caaggccaaa tggattaaaa tcatgtgtaa ttgtctctcg 720
cattctcgct gatttgtgca acagagctcc cacatggcca ccattgaaag gatggccact 780
agaaactata tgtgaaaagt ctataggtac ttgtaataga ctttggggcg ctggggagggc 840
cttgagacga gtaattgggt gtttggcctc tggaaacta cttctctggg gtccctggctc 900
tcattgctct tgtgagcgag acccaacaga tgcctctgag tatatgacca tccagcaaaa 960
agaagatatt acccagagtg cagacgatgc actcagacta tcagcctttg gtcagattta 1020
caagtgctg gagatggacc cccctccatc tagtaagcct ttccagaagt attcctggtc 1080
agttactgat aaagaaggtg ctggttcttc agctctaaag aggcacttgg aagatggatt 1140
aggggtgat aaagacccca caagagagat gaaagcaaa ttaagaaaaa ttctgttagg 1200
taagcaata gaacttatga atgcaataat gaggtcaaat cagatcagyc ctggtgttca 1260
gtatgaagctc ctatctcagc ctggcccccgt tcatcccca gtcttcacaa tgtctgtaga 1320
tgttgatggc acaacatagc aagcctcagc accatccaag aaacacgcaa aacttcagct 1380
agcgttgaa gttatgcagc caatgggata tccacagggc ttgtatgagc atattgaagt 1440
tatgagttcc gatgaaaaat cagataatga aagtaaaat gaacagtggt cttcaaatc 1500
aagcaataat actggaatt ctacaactga aactccagt accttagagg taagaaatca 1560
gggcctatc ctacacagaa gtggcaaaaa cctgttaatg gagctcaatg aaaaaagag 1620

```

WO 01/83524

PCT/US01/13862

```

aggtctcaag tatgaactca tctcagmgac tgggtggaagc catgacaagc gctttgtaat 1680
ggaggttagaa gtagatggac agaatattcag agggcgaggt ccaataaaga aagtggcaaa 1740
ggcgagtgca gctttagctg ccttggagaa actgttttct ggacccaatg cggcaataaa 1800
taagaaaaag aagattatcc ctcaggcaaa gggcggttgtg aatacagctg tgtctgcagc 1860
agtcacagct gtcgggggca gagggaagag aacttaaca agggagctt ttttggggc 1920
gaacgtgtct cctggctaca tagctccag ctatggaaa ccatatggt acagcaagc 1980
tgccccctgc tatgttttac ccaagagaat ggttctgtta cccgtatga aatttcacac 2040
atacctctgt cccactact cactctttta gcaaatgaca gaagctaat cctattgaac 2100
aacaatcacg tacacacag aatgttagag aaaaagcctt ttttctctg tttctttgaa 2160
cacatacttg atcaaaatta tttgtaagaa acatctttcc tacttttga ttttaacaaa 2220
tgcaaattta gttctctaaa acttgaaaa aaaaaagaa accggtctg tgaaaacgtt 2280
acctcaatttc tggaaaataa ottalaccag ccttctgtt ctagggaat aaaagtctag 2340
cagtcacaaag ttaagtttt aagagacgta coagattatg taanaataa ttttgaagg 2400
algtatagag tctcaaacac tgaatacaaa taactgtct tgtgttaaa caggtctg 2460
cotgttact gatgcagta ctgattotta gttgattgat atgtatttgc cccagggcac 2520
tttaatttgg gctgtagtta t 2541

```

<210> 81
 <211> 1647
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4545237CB1

```

<400> 81
gccccgggctc cgcctgcagt gtggccccgc cggacagctc ctcacccggc 60
cctcgctgctc tgcctgagct cgggcctcag gaattcctgt gcgtgccttt 120
tgacggagagc cagtgcaggt agcgcttgag caagctgccc ctggcggagc gccagagcgc 180
ggctcgagctt ccagctcggc gaacgcagtg cgcctgttgc gctgcctccc cgtgctgct 240
agcgggagcg caggtctatt attggggagg cagtgttccc taacacacct accagcagct 300
tcaattttgg cagggagagc tgttctcggc aatggccatt tgcctcactt gggacccaga 360
gacacccatc cctacggcag cttgagccgt gcactgcaga ccaaatgctg batttcttct 420
ccccctcacc tgatgagcca gcagtataga ccatatgtt tcttactaa attgactgca 480
galtgagctgt ggaagggcgc tttagcagag actgtgtgt gaggaaaaag aggaagagc 540
aaaggaacta aaaggaagaa aagaaggat ctgaacggg gtcagatcat tggtaaggg 600
cgttatgggtt ttataggccc cggactgaat gtccctctta tgaanaatag agcagtgca 660
acatttgccc aaagaagcaa ggaagagcag gagaagtggt aggcagacat gatccagcag 720
agagaagagat gggacggaaa gaagaagatg aaggttaaac gggagcgagg atggagtga 780
aacloaatggg gaggcatcag tcttgcccc cctgacccct gtccctgtgt agaaacatat 840
gaggattttg ataccaggat acttgaggtt agaaacgttt tcaatatgac tgcgaanag 900
ggaagaagaa aatcgatcct tctcttggtg gctgtggggg acggaanaag agctgcaggt 960
tttctcaltg ggaagctgac tgatcgggtg gatgtcttca ggaagcaaaa gaaacagaca 1020
gttccacatt tgaattatat agaacgatat gaagaccata caatattcca tgatattcca 1080
ttaagattta aaagagcgca tatcaagatg aagaacaac ccaaggttta cggcctccgc 1140
tgcccaaggg ccatcatcac catctgcctg ctcatgtgca tcaaaagcat gtatgccaa 1200
gtctctgggtt ccatctatat gctcagctc acccagggcc tcttctgtgt gctctccaga 1260
caggaacccc atcaacagct ggtctgatang aagggcctcc atgttgtga aatccgggag 1320
gaatgtggcc ctctgcccac tgtgtgtgct tccccccggg ggcctctgag gaaggatcca 1380
gagcaagaag atgagttgag agactcaaa ctggactggg aagatgtgaa gactgcacag 1440
ggaaagagag cctctgtgtg gtcataattt gaagagccg ccaagtaacc tctctgacct 1500
tgtgcagcca gtctctgtgc tgccctgcac ctggagagga ctacgcccc ccaagcttgg 1560
gatgttacct tgccttttgt ttgttttgag ggnagttaa tctttaaact ctttggnaat 1620
aataaattat agcttttcaa aaaaaaaa 1647

```

<210> 82
 <211> 735
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4942964CB1

<220>
 <221> unsure

WO 01/83524

PCT/US01/13862

<222> 723

<223> a, t, c, g, or other

```

<400> 82
tcgcgcagca cgaacctccac ttccacatct ccccccggct cggcgccgtc 60
agtggaacca tggcgggactc caaggccacc tggcggttca cccctccgac ccgcaagttc 120
atgaccaacc gccctccctggc ccgcaagcaa ttctgtcttg aggtgatcca ccccgccgac 180
gccaaagtct ccaaggcggga gttgaaggag aggtctgcca aggcgtacga ggtgaaggac 240
cccaacacca tctttgtctt caagttccgc accactctcg gaggaggaaa gtccactggt 300
ttcgccctca tctacgacaa cctcgaggct gccaaagaat tccagccgaa ataccgctc 360
atcaggaaat gctctgtctac taagggttag aagtcgccga agcaaatgaa ggagccgaag 420
aacagggcca agagagctcg tgggtcaag agaccaaa gctggtgacg caagagaag 480
taaacgttcg tttaaatctg tattaactgt ctgggtctcg ggtggctcag ctgcaatgic 540
ataattatgg tctgttttag tttgttcca cctctggcac tgaagtgtat ttttttgtaa 600
ttctctggca ctgaagtgaa gttttgtctg aatattgctt cgttaacata ttgcccgttc 660
cctgtctcag ttgtggcgca gtcctggttg ttttgacatt tctaactctg gttaatgtgg 720
ntggatcggt tcatg 735

```

<210> 83

<211> 2514

<212> PNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5702144CB1

```

<400> 83
gtgcgctctc acccttatct ccaaatctcg ggtgtgtctg cgggggctgc tgtgtccgga 60
actctccgtt ccgttcaggg tccgcgactc cgganbaagg atggcgtccc ggggtctctg 120
gggcgctgcc cgttggtctc ggcccgcgcy ggcgcttgcc ccagccgccc ggcccctgag 180
ctccggttag ccgcgctcgg aggagctgtt caccgggggc gggcccttgc ggaactctct 240
cgagcccgag cgggggtctg aagccatctt gaaggtcagg aggcccgagt tgcctggcgt 300
gatcaaacct ctgaacgaga aggagcagga gctgcgggag actgagcact tgcctgcaga 360
tgagaatgaa gatbtaagga aacttgcaga gantgaatc actttgtctc aaaaagaaat 420
aactcagctg aagostcaga ttacttact ttggttccc tccgaagaaa cagatgcaaa 480
tgatttgatc ctggaagtaa ctgcaggagt tggaggtcag gaggcaatgt tglttacatc 540
agagatattt gatattgata cgaactatgc tgcatttaaa agatggcalt ttgaaacatc 600
ggaatatttt ccaagtgaac taggtggcct tagacatgca tctgcagaca ttgggggttc 660
agaagcctat aggcacatga aatttgaagg aggtgttcac agaglacaaa gagtgcocaa 720
gacgaaaaag caaggccgca tccatactag caccatgaot gtagcaatat taaccacagc 780
tactgagatt aatctgggtga ttaatccgaa agatttgaga attgaacata agogagccag 840
tgagctgggg gggcagcatg taataaccac ggaactgtct gtcgggtag ttcatctctc 900
aacaggtgtt gttctcgat gtcaacaaag gagctctcag ctgaaaaata aagagctggc 960
tatgaacaaag ttaactgcaa aactgtacag catcatctca gaagaagaaa taataaagag 1020
acagaatgct agaaaaatc agattggaag taaggaaga tccagagaaa taagaacata 1080
taattttcca cagaacgggg tccagatcca cagaataaac aagacgctgc atgatcttga 1140
aactttttatg caaggagatt atctactgga tgaacttgta cagtcattga aggaatacgc 1200
cgattatgaa tctttagtag aaattatttc ccaaaaagtt caagltgalt tglttaltal 1260
agactttcgt agcttagaaa aattctacag tacatccaca tagggtgaaa gtacccctac 1320
tctcttgaaa aacgttgagt taacacaggtt ggaagtaata tgcacattct gaagctcag 1380
ataatttaca cagatctctc tcaatgcatt agcaaaaatc atcaaatata cagatggctc 1440
tcgatttaca ttgtggttaa ttcccaatga acccatcata agttaaaat gcatataacg 1500
ttagcaaacac agcagttctc taattaatga cagottgaot taacaatttt ccaattttac 1560
catgtgttga aagaggtatg attcctaagc cctaaggagc tccctcagctt gaaatggggc 1620
tgcattccca taaccccatc ataaagtcaa aaaaacctaa aacataagtt ggtgacatc 1680
tgltaactatg atgtggttgtt aaatcttggc cgtacactta caataactag acaaggaaa 1740
atcaactctt gctcgttctc gtgtaaatat ttaatgaatg atcaaaactt cagtttaaat 1800
attatgaaa acatttaaaa taagtagta gaataagac gtaactaat 1860
atccagtcag gatacagaaa ccataccatt aactgaaca gggataattt taatataaaa 1920
actgttaact gatnatgta ttaactttta agagggatga aagagagcta tcatgtctca 1980
ggactgagag taacccagga aagaataccc ttgaagggtt ctccctctcc ccatgttgaa 2040
gtcaggctca atggagagag tggctacagc ctactcagtg attgggaat tccctgtctt 2100
gccttgggcc agagctgggt taccgtggtt ggaatcaggt ctacagagca agaacctcac 2160
actccanct ggttaagccag aagcctcttg ctagggtgtg agcaaaactt ggaaggaaac 2220
tctcagtata tgtttgtgtt tgtcagagtt ctccagacaa acttctctaa aaggatggc 2280
ttgtgttgtt attattaaat ctaacaaatc caaaagctgg agtgtgagc aggaggtctg 2340

```

WO 01/83524

PCT/US01/13862

```

aaacccaggga aagctgctggg tgcgaaggctc agtccaaagg taletgttgg aggattctct 2400
tgcttcgggga agaggagcgt ctttttttct ctccagacct taccctgact ggatccagcc 2460
cactaacatc gagggaggaca gtctgcatta ctccaggttc actgattgat ttaaatgtca 2520
atctcattata aaacaccctc acggaacac ctgataaact gtttgacctt ataattggaa 2580
aatccaggca aaagttaaat ctctaaaaaa aaaa 2614

```

<210> 84

<211> 736

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5862945CB1

<400> 84

```

actcggcggc ttcggtagcg ggaggggcga agatggcggc gccagtgctg ggacagtgg 60
gtgcgttatg gatacalaac ctgaggagcc gggggagcct gccctgggtt gtttlacctc 120
aatcatatat ccacacaagt gctctctctg acattctctc aaaaaggagg aagaagaala 180
aaaltgttclt tctccacaca ctgcctggag aacctcggag accagcagaa atctaccact 240
gtcgaaagaca aatbaaaatc agcaagaca agatgtggtt ttggcaaaa ttgataagag 300
aatgtctctat tgaacaggct ttggtctcag ttggaatcaa tgaaaaaaa ggggccaana 360
taattaaaga ggttctctta gaagcaacag atatggcagt gagagacat aacgtggaa 420
tcagggtccaa ttatatata gctgaagtcn cctcaggacg aggcagtgcc ctgaaacgca 480
tcgcctacaa tggcagaggt cgtcttggga tcatggagaa ggtttattgc cattattttg 540
tgaagtctgtt ggaaggggcc ccacctccac ctgagccacc aaagacggca gttgcccalg 600
ccaaugagta tattcagcag ctccgcagcc ggcacctcgt tcaacctcta tgatgaggag 660
atctcagctc cncagtgat atattttgcc atttatlttc taaaaataaa caaaaatga 720
aggcaaaaaa aaaaaa 736

```

<210> 85

<211> 1046

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6319547CB1

<400> 85

```

ggcgtaacgc gtccagggcg gctcggcagc tggcggcalt gagggcgacg cgtctagagg 60
tcogtttgac cgcggcgctcg ggaacctggt tccggggcatg agctgagagc accacgcga 120
ggccacaggt attccataga catgtatgga agcagaaacc aaaaactctc ccttggagaa 180
tgcatacctc ctctccaggg gctctctgca ggaaggaccc cgttatgga ttggaacct 240
ggaccccaaaa attaccgaat accaactcct caagctctc cagaagtltg gcaaggtaaa 300
gcagtttgac ttctctctcc acaagtcagg tgccttggag ggcagcctc gaggctactg 360
tttctttaac ttgaaacta agcagggaag agagcaagcc atccagtgc tcaatggcaa 420
gttggccctg tccaaagacg tggctgtgcy atgggaacat gctcaaglaa agagatatga 480
tcatacaaa aatgataaga ttcttccant cagtctcgag ccactctca gcactgagcc 540
taactcagtc aacctagtg tcaactgcaa gataaaagcc attgaagcaa aactjanaat 600
gatgggggaa aatcctgagc cagagtalcc agcagcgcc gtttatct accitaaagc 660
acccgataaa aaaggagata ctccatctc tagaacagca tggaaatctc gaggatgag 720
gttgtgaatt actgtagcag caaaagcaaa ttggtctcca cactaaaat cplctgtctg 780
tgtactttgt agatgtgaat ggtactatct aacggagcac aatcacatgt tagcaatttg 840
taacataatg ttttggatg ttctattgga tgtttctcc ctaaacatg tatggaaattg 900
agcatcatcc agataaata gonyttgtat ccaaatgtg atttgaacco tgggatgtctc 960
taattggctg gttgtgttgg atttgtaact ccagaaacat tctatagtgt gccagagcaa 1020
aaggcaataa caaaaatat tatctt 1046

```

<210> 86

<211> 2266

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 000124CB1

WO 01/83524

PCT/US01/13862

```

<400> 86
cggtttacc agcccggaag tgcgctggc ggccgtggcg gctggcgcaa cagcgggcc 60
gatgtgtagt tgytgactgc ctctccagat gctgaggtgc ctgtatcatt ggcacaggcc 120
agtgtctaac cgttagtgga gtaggtctgt cctctggaag cagtatctat tcacaaatgaa 180
gttgccagttt ccggaattcc agtcaatttt ccacgaagga ctgaagagtc tgaagaaatt 240
atttgtcaaa gagaatcaag aattaaagat agcagggagg ggcgtgaggg atttattaaa 300
tggagttaag cctcaggata tagatttllgc caccactgtt acccttactc aaalgaaagg 360
gatgtttcag tccgtcggga ttccgatgat aaacaacaga ggaagaaagc accggaacaa 420
tactgccagg cttcatgaag aaaaatttga gattactaca ctacggattg atgtacccac 480
tgaatgaaga catgctgagg tagaatttgc aactgactgg cagaagagtg cggagcgcag 540
agatctcaat ataatttcta tgtttttagg ttttgatggc actttatttg actactltaa 600
tggttatgaa gatttataaa ataaagaagt tagatttgit ggaatgcta aacagagaaat 660
aacaagagat tatcttagaa ttctaaagata ctccaggitt talggagaga ttgtagacaa 720
accgtgtgac catgatcctg agactctgga agcaatttga gaaatgcaa aagccttggc 780
tggaaatcca ggagaagagg ttctgggtga actgaaaaaa attcttgttg gtaaccctgt 840
aaatcatttg attcaacctt tctatgntct tgaatgtggt ccttatatag gtttacctgc 900
tgaatgaagt ttgaagaagt ttgacaagat cagtaaaaat gttgatggtt ttccaccaaa 960
gccagtgaat cttttggcct catlactcaa agtacaagat gattgcacaa aattggattt 1020
gggtgtgaag atcgcaaaag aggagaaaaa ccttggctta ttatagtta aaattaggaa 1080
agatttaatt aaagcaacag atagtccaga cccattgaaa cctatccag acttcatatt 1140
agattctagg gaacctgatg caactacttg tttatgtgaa ctactgaagt accaaggaga 1200
cactgtcttc ctaaaagaaa tgcagcagtg gtccattctt cctttctctg taagtgcaca 1260
tgaactcaga aaagtgggca tttcttcagg aaaaagaatt ggggtctctat tacaacagtt 1320
ggaggaacag tggaaaaaaa gtgtttacca aatggaaaaa gatgaactt tgaattacat 1380
aaaggaagacc taaaactgat ggtctactaa aagcagagca tttctggtta gactaaattt 1440
tctccccctc ctcttaatga ggttttagag actacacagc aataaaagac agtttagggg 1500
acccctctag aacaacaagg gtcttatttt gtgaattata tatttoaga actaaacaga 1560
gatccacctt tctggatctg atttatacca ctgaaatgta cagttctttt ggaatagttt 1620
cactcagaaa caactagttg gcttatctct actcttancct gttcaggctt ttanaaaaaa 1680
cgttttttgc ataggttagt actaagatct taaaagtgtg taactgtctt gaagaaaaaa 1740
cgttlatgtt ttgtttgcaa ttgaataaac aggtttacct taacaaagac tgtctatgat 1800
gtgtcaatttc tatctgaat tccaaaataa acctgtgctt aaaaaagaaa taattgacca 1860
agtaagtttg catanaatgt gataactaaa tgtgtcccca gttgtgtgca ttcataatgta 1920
caggaatttgt tctagcaagg tatgcttcag tatgtggttg atatttttct gtcacaaaga 1980
ttcttttatg caagcaagac ctgggaagat catgggaltta acttgaggtt cactattgag 2040
cctatcaant aattaaatct tgttttataa aaacaaacat ttgttttggg agulaaatal 2100
gtttatgttg tatctgacaa tgtgtattag gtgctatata caatggtaat atgctgtct 2160
ttaaagtgtt abtttattaa ttaaaaggat atggctatta ttatatatto tctaaagatt 2220
tattctctaa agatttgagt cctaaatgat ttcatcagcg cagcag 2266

<210> 87
<211> 1041
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1659474CB1

<400> 87
cagtcagat ggttcaggt tgcctagag cgcgcggcc accggcggt tctgatgggt 60
ctctggtag gtaggttgg cttctgctt gcttagagtt gccacttat gccgtgctt 120
gtcgtggtt gaacagtcg tactcatgcc tggtgcccg gccgcaccaa agccacatcg 180
cgctctgcg ccgtacactt aacaggaaac gcaccggcat tgcagaacag cttgatcggt 240
agctccttcg ctattctgag agccttttag gtgtccctat tgcataatga aacatcaaa 300
ttgtggggaga gcttggagat atttatgat atcaaggaca catctatctt aacattgaag 360
cagattttgt tattttctgc cctgaaccgg ggcagaagct tatgggtata gtttaataag 420
tgtttctag ccacattggc tgtttagtac atgggtgttt caatgcctcc attcctaacc 480
ctgagcagtt gtcagctgag cagtggcaaa ccatggagat aaacatgggt gatgaactag 540
aaattcaagt atttctgtta gactcagat ctgtctgagt attctgactt cggggaaac 600
taaatataac aagtttataa ttcaagcgt ctgaagtctc tgaagaagtt acagaaatg 660
gcactagaga agctgtctaa aaacctataa agaaagaaaa gaagaagac ccagagacat 720
atgaagtggg cagtggatcc acaagctag cagatgatgc agatgacac ccaatggaag 780
agtcagccct gcagatact aataatgcga atggcatctg ggaaggaggg ccaagaaaaa 840
agaaagagaa gaanaagcac caggaagtct aggaccagga cctgttttcc caagcagtg 900
actccagtgg ttaccaaaat gaccataaaa agaaaaaaa agaaaaagaa accaagctg 960
aagagggcca atttaccaca ctttgaatat gctcaccaaa aagaaaaggg aaaaagtaat 1020

```

WO 01/83524 PCT/US01/13862

ttcttttagtg tatttttaaac a 1041

<210> 88
 <211> 2722
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2257892CB1

<400> 88
 cgctttcttg gtaaatgttg acgtccacga tctctttcgc cggctcggcg cgggggccc 60
 atcgcacag agacgtttct cggcagacgc agctcgatcc cagataggaa aaaggaaata 120
 tgactttgat tcttcggagg tgcctcaggg actggacttt ttggaaaca agaactctgt 180
 ccacagtggt tgtggagcat cacaaacaca tcagaagccc caaatggag agaaaaaaga 240
 agagagccta actgaagga agaggagca gagcaagaaa aaaaggaaag cgaatgacttc 300
 agaaattgct tcccagaag aggtgctac tatcagtggt atgtcatctg tagaagcaaa 360
 gatggaagac aaaaaagttc agagagaaag taactaact tccggaagt tggagatctc 420
 cagaaaaaa aagataaact tcttgcggaa taacacaaa attcacgtcc aaggaaacga 480
 tcttctctgac ccaattgcta catttcagca acttgaccag gaalataaaa tcaattctcg 540
 aactcttcag acaattctag atgcagggtt ccaaatgctc acgcaaatcc aatgcaagc 600
 catccagttt abgtgcctg gtccggaaact tctggtctct gctccaaact gatctggaaa 660
 aacattagct ttttagcatt ctattttaat gcagctgaaa caaccggcaa ataaaggctt 720
 cagagccctg attatattac caacacgaga acttgccagc cagattccaa gagagttaat 780
 aaaaaattct gaggyaacag gattcagaat acacatgac cacaagcgag cagtggcagc 840
 caagaatttt ggccttaaat catctaaaaa gtttgatatt cttgtgacta ctccaaatcg 900
 actaatctat ttattaaagc aagatcccc cgggaatcgac ctagcaagtg ttgagtggtc 960
 tgtatgtgac gaatcagata aactgtttga agatggcaaa acttggttca gagaaccagt 1020
 ggtctcaatt tctctggcct gacatccca caaggtccga aggtctatgt tcaatgcaac 1080
 ttttgcattt gatgttgaac agtgggtcaa actcaacctg gacaatgtca tcaatgtgtc 1140
 cctggagcga aggaattctg cagttagaac acttagaaca gagctttctc ttgttgggac 1200
 ttgagaccga aacattcttg ccatgagaga acttgttaaa aagggtttca atccacctgt 1260
 tcttgttttt gtccagtcga ttgaagggc taagaacttt ttctatgagc tcatatgtga 1320
 aggtattaat gtgcatgtta ttcatgcaga gagaacacaa caacagagag ataacacagt 1380
 ccacagtttc agagcagaaa aaatctgggt tctgatttgt acagccttgc tagcagagag 1440
 gatgattttt aaagtgtgga acttggtgat caactatgca ttccaaacta gctcagtgga 1500
 atatatccac aggtatgtgc gaactggaag agcagggaaat aagggaaaag caattacatt 1560
 ttctactgag gatgataagc cattattaaag aagcgttgct aatgttatac agcaggtctg 1620
 gtgtcctgta ccagaataca taaaaggttt tcagaactca cttaagcaaac aaagaaaaaa 1680
 gatgattaaq aaaccatttg aaagggaagc cattagtaca actccaaat gttctttaga 1740
 aaaaagtaag gataaacaga aaagggtcac tggtagaagc agcagaaga aagtagctct 1800
 tgaagcaaaa acttaaaaaa agactttama aatactgtcc cagaaatgta attttatgat 1860
 ccaagcagga atgtattttt catggaatac ttgaagtttt acagtcacct gtaccacaaa 1920
 ttgaaataca actcaagata catgggactg gtgataaatg atctbaaact atcaagtccg 1980
 ttccaatttg taagtgcctt ttttttttcc tgttagagat agggctcttg cactgtgtcc 2040
 aggtgtgtct tgaactcctg acctcacaca atcctcctgc cttagcctcc tgaatagctg 2100
 agattacagg cacaagctgc tgcacccagc tctgtagtg acttttaaat gattatacaa 2160
 tgaataaac attcattgac atttctgtg tttgaatcca gagagatact tcttatagaa 2220
 aaacaaatgt ttatgtctaa aataaaccca aaatgtgtgt aactottaaq gacttttccc 2280
 ttcaagtgtg aaggaggtgt tgaatgaatc tctggagagc catctggaac agaaattcaa 2340
 aatgaagctc tgaatttaaa taaccccttc actgctcaact ttgtggatgg tagcagagc 2400
 tgtctaccaa gaagaaacot gctgctctct taattttaat atttctaat ttgtttgagg 2460
 cctttgtgtt tgtgaaccac aacaaagaga ggcctctttt gtggtgtgtt attccagttc 2520
 cctgggtatt taaattcttt ggtotattaa gtatccttgt atgtgatac taatacctta 2580
 gtgtgtcat aatgttgtac aagatcatga tcaagctctc cctttcttca tttttgtgaa 2640
 ttttaaccatg ttctttctgt totctttcca ttttaagatat tttatttgaa tactgtataa 2700
 caatttatcc ccccccttgg gg 2722

<210> 89
 <211> 1287
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2670307CB1

WO 01/83524

PCT/US01/13862

```

<400> 89
ccaagagctct aggttaagagt ttgttccogt ggtgcggagg gtcaaggccc acaccggaa 60
acctagcgag gtaaatgtgc gtcttgggtg tagagacgac aacttctccg cttctctggc 120
gatggcgccg tccgggagcg gtatggccca gaaacccctg gaactggcca acaactgca 180
ggagagctcag agtatcgatg aaatctacaa atacgacaaag aaacagcagc aggaatctct 240
ggcgcgaaag ccttgactaa ggatccacca ttaacttaag taactgaaaa tctcagcatt 300
ggctctgctg agatgggtga tgcattccag atcgggaggc aacttggaag tgatgggtct 360
ggtcttagga aaggtggatg gtgaaaccat gatcattatg gacagttttg ctttgctgt 420
ggagggcaact gaaacccgag taaatgctca ggctgctga latgaatata tggctgcta 480
catgaaaaat gcaaaacagg ttggccgccc tgaaaatgca atcggttggt atcatagcca 540
ccctggctat ggctgctggc ttctgggat tgatgttagt actcagatgc tcaatcagca 600
gttccaggaa ccatttggtag cagtgtgtgat tgatccaaca agaacaatat ccgcagggaa 660
agtgaatctt ggcgccttta ggacataccc aaagggctac aaacctcccg atgagggacc 720
ttctgagtac cagactattc caatbaaa aaagaagat ttgggtgac actgcaaca 780
atatctagtc ttgaagtct catatttcaa atctctttg gatgcgaat tgcttgagct 840
gttggtgaat aaatctctgg tgaatagctt gacttctctt agcttgctta ctaatgcaga 900
ctatccact ggctcaggtct ttgatttgtc tgaagaagta gacagtcag aagccagct 960
gggacggggg agtttcatgt ttggtttaga aacgcagac cgaataatcag aagacaaact 1020
tgccaaagct acaagagaca gctgtaaac taccatagaa gctatccatg gattgagtgc 1080
tcaggttatt aaggataaac tgtttaatca aattaacatc tottaaacag tctctgagaa 1140
gtactttacc tgaagacag tatgagaaaa atattcaagt aacctctaa aaccagctac 1200
cgaaatctg attagagta taaggtgctc tgaagtgcc taaatattaa tatctgttaa 1260
taagactctt taaatgaaa aaaaaaa 1287

```

<210> 90

<211> 2226

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4524210CB1

<400> 90

```

cggctcagac cggaaagcag tttccgcga gaaatagggy gcgctgtgtt ggaattgat 60
agaaaagata aagggaaccg gctcgtgtca gcttggtcta ctgactctcg tccgtttcac 120
cagggatcca gtactaagc attttttct ttttttggtt ctttgcaag tgagtgcat 180
tggtcagtg atttccatga gcactctctac cagaanaaat tgcctcgatg aggtgtgtgt 240
gaagcccgcc agccctctct aatcggttag gcttgagaaa cgtgtacct ctgcatttcc 300
gaaattaaat cagcgtgtat ggcgaagatt tctcagcat ctggtgtcaa gacactctgc 360
actattaatt cggaaagaaa aaaaaaaca aaacaccgtt ttccagcatt tctctttgtg 420
gagaactaaa caacaggaaa aatgtctatt ttccctaaga tatctttgag acctgaggtt 480
gaaaactatc ttaaggaaag ctttatgaa aaggagattg tgaactgttt aggttaacaa 540
gaagcagaaa ggaagcttga aacttgtta agcacctgt ccaatctctc atcatctaca 600
actgcaagag tgaatacaca tttagctcca gtacaacatg tgaanaaatg gttacttgat 660
gaactctaga agcagtttaa tggatbaagt gtctctatc ttcaacatcc agacattcaa 720
gactgtctac ttattctgtt tattggaccc agaaagata ttaaaaaaca acagtgtgaa 780
gcaattgttg gacccagtg ttgcaatgca gttttaagag gaggccatgt ctatgcccc 840
ggaattgtgt cagcatcaca atttatgaaa gctggagatg ttatttctgt atactctgat 900
attaaggaaa aatgtaagaa aggagccaaa gaatttgatg gaacaaagt atttcttga 960
aatggtattt ctgaactaag ccgcaagaaa atcttcaagt gattacctga actgaaggcc 1020
atgggcataa gaatgacaga accagcatat ctacgcccct cacttgacag tgaactgtcc 1080
cgttaactat ttttaacaaa tttagcatct gcttagttaa gtaactgact aaatctctaa 1140
ctctggagaga agattctaga ctbtgtgca gcacctggag ggaacaaac acacattgca 1200
gcactaatgc atgactaggg agaagttata gcactggata aaattctca caaagtagaa 1260
aaaaacaaac agaattgctt attgttaggg ctgaattcaa tcagggcatt ttgttttgat 1320
ggaacaaagg cgggttaact tgaataggtg gaggaaacag aaggagacc tccatttcta 1380
ccagaatctt ttgacgaat tctctggat gaacontgta gtggaatggg acagagacca 1440
aacatggcct gtacttggct tgtgaaagaa gtggcatcat atcagccatt acagcgaaaa 1500
ctcttcaabg cagcgttcca gctgtgaaag ccagagggtg tctgagttta tagcaaggg 1560
actataacac tggcgcaaaa tgaagaaacag gttgctggg ccttgacaaa atttcttgc 1620
cttcaacttc agcccaagga accgcagatt ttgagagaa gaaatgaggg agctgggctc 1680
tcattgtaac agttgaacaa gctgcagcga ttgatccat cggctgtgac attacgggac 1740
actgacatgg actctcttag agagggcaga agagaagaca tgttggctct ggctaaatag 1800
gactctatag gtttttttat tgcaaaattt gtaaaatgca aaagcacata ggagagggat 1860
ggatgctcag aaatgaanaa tccaaacatt tgcgtctctg ggtttttttt tttttttttt 1920
taacaaaagt gttgtcaggc caactgaatg atgatgtgtt tgcctatgaa acngaaaagg 1980

```

WO 01/83524

PCT/US01/13862

```

ctgcacagctg ttttaccagg gatccagaga catagaggaa gtaggggggtg gtatgagatt 2040
atattttcttg tttttaaagg attttttttt tttatgtatt tagtagagta taaagaaaag 2100
cagatgccta tagatgtctg gagcataatt toatttbtga tctaatgttt taatttgtaa 2160
agtggtacag toatttttaa ggttaaaaat tagtgaatct aacaaaagga ataaatttagc 2220
aatatt
2226

```

```

<210> 91
<211> 2362
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5584860CB1

```

```

<400> 91
cccggttcga cccacgcgtc cgaataaga cgcgcacgg cgcggcgcta gcctcggggc 60
ttgacgggat tgtggcgggtc ctctctccca attcgggaag tacagctacc tccggagcgt 120
ctcaagagtg cgcacctctct ggggtccaaac accatacaaa ggcagaactg ggaggatggc 180
gaccttccca ttctgttgcca gacatgtctt ggagaaaacc catatactcg aatgaccgaa 240
gaaaagtatg ggaaggaaag caaatctgtt gccaggccat tccagtggtt tggctgggtg 300
cctggagttcc gcatgcgttt caagaagact gaagtgtgcc aaacctgcag taaattgaag 360
aatgtctgtc agacctgctt cttagacctc gagtatggcc tgcacctcca ggtctgtgac 420
gcaggatgtt cttttaaaga tgacatgcga agtccagatg tcaacaaga gtactataca 480
cagaatctgg agagagagat ttctaactct gatggaacac ggcagttgg catgctgggg 540
aaagccacat ctaccagtga catgctgctc aaactggccc ggcacacacc ctactacaaa 600
aggaatcgac cccacatttg ctctctcttg gtyaaaggag agtgytaagag agggagagaa 660
tgctccataca gacatggagaa gctatagatc ccagatgacc cctctgtgta tcaagaaatt 720
aaagaccgtt attacgggaat caatgctcct gtatgtgaca agcttctaaa gcgggttcaa 780
aocattgcctc ggcctggaccc accagagggat aaaaactata cccactata tgttgggtggt 840
ctaggttgata ccattactga gacagattta agaatcatt tctaccagt cggagagagat 900
cggacgatca ctgttgtgca gagacagcag tgtgtttca tccagttttg caccagcgag 960
gctgcagaag tggctgtgta gaagtcttt aataagttga ttgtaaatgg ccgcagactg 1020
aatgtgaaat gggggaagatc ccaggccagcc agaggaaaag azaaagagaa agatggaact 1080
acagactctg ggaaccaact agaacctgtt ccaggattgc caggagctct tccctctctc 1140
cctgcagcag aagaagaagc ctctgcacac taactcaact tgcctccaaag tgggtctcaa 1200
gctgtgtgta acattgtctt gccacgcgcc cctggcattg ctccacccc acccccgagt 1260
tttggtccac acatgttcca ccaaatggga ccccccctc ctctcatgcy ggtccaggaa 1320
ccaatccact atccttctca ggaacctcag aggatgggag ctatctgtgt aaaaacagac 1380
agccctctag acctgtgtac cactctgggg ctctgtggaa gaaaggggac ttaaaactcc 1440
cagtaaatct tggaaataat atatttttcc ttcccttgta gtttccatgg tagctgaaat 1500
tgctcagatg tggcagtgca gagactgaca gccatgcttt cctatacttg ttcaagggat 1560
gtagggacgg taataagctt gccattacaa catctgggta ctgtgtgacc atgactaaat 1620
aaaacgaacg cctgttcccc ttaccctgtt gggggacacg cagatgagtg aattggaaat 1680
tccagcagag ttaccctccc aattatagtt toattttgta tatttttgg tccgggggaaa 1740
aatgcacgtg cagtataaaa acotttgacc atttttatgt ccatttgata ctttctcttt 1800
tatcatctta aaaaaagata actagtacta atcattgtag tggcctaagt gtgatttaac 1860
tcttgagctc acacccctcg aaagttaggt agaaaacagc accagacagc cccagatctt 1920
ctctttctcc tctttttctt catttattcc taagggaatc tgaccatttt acgtctctac 1980
ggcccaaaaa aagacaaaaa taaaaattcc tttttattcc tgtcaactgy atggaacac 2040
aaatttctat gagggtgta ccactgaaga aactgtgtg ctgggaagaa attactgtaa 2100
agaaattctt gtaaaaacag ttctttaaaa aactgaaatg aaaaagcatt gacgtctctg 2160
atgaagacgg tgacctctgt ctgggactct gatggtcttc agcattoacc ttctgtgtct 2220
ttcagtgctc catgtctatc cctgtctctg tttgtcttta gagtgtttgg atataactga 2280
attgtagatg gtaaaaggaa ttgtatgtgt tttttgtttt taataatta aaacggtgca 2340
atttttcaaa aaaaaaaaaa aa
2362

```

```

<210> 92
<211> 731
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5807892CB1

```

```

<400> 92

```

WO 01/83524

PCT/US01/13862

```

tagggcgga agcgaggag gctggcgag cggatcatcc gottcggag tcgaggttt 60
cgggcttgta cagcttggcg gtcggcgctg gtctggcgtt gcaggttctt tctgtgttg 120
ttctctgccc tgcoagggcc gttagagctgg tgcgtgcggg tagcggggct ctcgaggag 180
ccgcacgcgc ggggcacccat ggtccacctc actactctcc tctgcaggc ctacogtggg 240
ggccactttaa cactcagcct tgcctgggtt ggcctgggac atcggcaggt ctacccgact 300
gtggctgctc acacacagtg tcccgggat ggcgtttctg tagagcagct gggctcttat 360
gactccattgc ccaacagtca tggagaaaaa ctggttgcgc tcaacctaga caggatccgt 420
cattggattg gctgcggggc ccaactctct aagcctatgg aaaaactctt gggctcttgc 480
ggctttttcc ctctgcctcc tatgatgato acaaatgctg agagactgcg aaggaaacgg 540
gcacgtgaag tctgtttagc ttctcagaaa acagatgcag aagctacaga tacagaggtc 600
acagaaacat aatgagctg acttttagtga gcatagcagt gggaacaagg tcaaggtctc 660
tttgaacac tgcagcgatc ttaattttgt tagatttggg gttcaataaa tggagtatcc 720
tgaaaaaaa a

```

<210> 93
 <211> 2088
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3210044CB1

```

<400> 93
ctttccagaa aatcaaatga aagattcaga gtatatctcat gaattaatlt tttttcaaaa 60
ccctcaattt aatcagctgg aattacttta aaagtgtcat tctatttaac ttttgggaat 120
ggtgaatttg ctttttaata gggatgctgt attttatcat gaagatgaaa caaacctgtc 180
tttgttaatt atggcaactt catttaaccg ccgcatccag ttgttctctc tgcgggcgtc 240
agagcttctc atagctttag taggcagagc agcttttagc tttagggggc caaagtttgc 300
ctcaaaagacc cctcggcgcg tgaactgaac attgctcttg ctttcggggg tgcagctggc 360
caagctgctc cgaacagagaa aggtgaagt tatagatgtt gttcaggctt atataacag 420
aatcaaggac gtgaacccaa tgatcaatgg aattgtcaag tacagggttg aggaagcgat 480
gaagggggct catgctgtag atcaaaagct tgcagagaag cagggaagat aagccacctc 540
ggnaaaataa tggcccttcc ttggggttcc tttagacgtc aagggaagctt tccagctaca 600
aggaatgcct aattctcttg gactcatgaa ccgttggcat gcatgtgcca aaacagatgc 660
caactgttgt gcatbaatga agggagctgg tgcactctct ctgggctaaa caactgttag 720
tgagtgttgt abgtgttatg aatccagtac caagatctak ggcctgatca acaacccata 780
lgatttaacg catattgtag gtggaagttc tgggtgtgag ggcctgcacc tggcagctgc 840
ctgctcagtt attggtgtgg gctctgatat tgggtgttagc attcgaatgc ctgcttctct 900
caatgggata tttagacaca agccttctcc aggtgtgtgtt cccaacaaag gtcagtttcc 960
cttggctgtg ggaagccagc agttgtttct gtgcactggt cctatgtgct gctatgctga 1020
agacctggcc ccaatgttga aggtcatggc aggaactggg atcaaaaggt taanaactaga 1080
caaaaaggtc catttaaaag acttaaaatt ttactggatg gaacatgat gggctcact 1140
tttaatttcc aagctggccc aagatctcat tatgactcag aaaaaggttg tggttccact 1200
tgaaaactat claggagcct cagttaacca tgttaactg aagaaaalga agtactcttt 1260
tcagttgttg alcgcaatga tgtcagcaaa gggacatgat ggggaaggac ctgtgaattt 1320
tgtagatttg cttggtgacc atgggaaaca tgtcagtcct ctgtgggagt tgatcaaatg 1380
gtgctcgggt ctgtcagctc acacacatcc ttccattgga ctggctttgt tggaaagaaa 1440
gtcagatata agcaatgaga aalacaaaaa gtttaaggca gtggaagaaa gcttgcgtaa 1500
agagctgtgt gatatgctag gtgatgatgg tgtgttctta tatccctcac atcccacagt 1560
ggccacatag catcagttcc cttaaacag gcccttaaac ttgtcttaaa cagggtctct 1620
cagtgccctg ggtttgcctg tgacccaalg cccactggga ctgaalgeca aaggactccc 1680
tttaggcctc caggtgttgg ctggacccct taalgatcat ctgacctctg ctgtggccca 1740
gtactgtgag aaaaactttg ggggtgggt ctgtccagga aagttttagg aggaactctc 1800
geaaggttaa tgtgtgtgtg tgtttgtgtt cgtgtgttgg tgtttctatt aattgggtga 1860
aatcaagcac cagcagacaa gcagagaaac aactggggaa tttattgact catttagtta 1920
ttctttctac ttttatttcc ttctctaac gtgtgtctta ctzaaatgtt aatatttgtc 1980
tttgcctttt agttactgg aaaaattagg catgtaaatg gataagtcca ataaagtttc 2040
ctaaatgctg aaaaaaaa acacaaaaa aacaaaaaa aaaaaaaa 2088

```

<210> 94
 <211> 660
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature

WO 01/83524

PCT/US01/13862

<223> Inocyte ID No: 4942454CB1

```

<400> 94
ccgtcaatag cctccgcctc tccttccagt gtcgcgcgtc gtygcctgc taccocctc 60
cctccaggcc ttctccggcg aagagcgccc agtcgccac caggatgaag ttctgtgtg 120
cctacntgct tgcgtctctc gctgggaact ccaggccctc tgcctaggac ttgaacgca 180
ttctggagtc agttggctgt gaagttgaca atgaaaagat ggaactcctt cgttcccaac 240
tgagcggtaa ggacattacc gactcattg ctgctggcag ggaagagttt gcttcagtc 300
cattgctggc tggcgtgtgt gctgttctgg cagctgccc tgcgtctggt ggcctcctg 360
cagctgagcc gaagaaagaa gagaaggctg aggagaagga agaaagtgat gacgacatg 420
gcttcagcct ctccgactaa gctgtgcaa tagtcaagag tattgttttt gactgcgga 480
agcagagggg agaaaaatcg tagtcatgtt tggactttaa ctttgtttta gcttggaaag 540
tacttgaag actttctctg tggtaattct aggcgtaggt tgcgtgtctg gttggggttt 600
aclygtgaac cagagttttt ctatctccc ctatgaattt gttacotcaa gttacotgtg 660

```


【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
8 November 2001 (08.11.2001)

PCT

(10) International Publication Number
WO 01/83524 A3(51) International Patent Classification: C12N 15/12,
C07K 14/47, C12N 5/10, C07K 16/18, C12Q 1/68, G01N
33/68, 33/53, C12P 21/02CA 94577 (US); YAO, Monique, G. [US/US]; 111 Freder-
ick Court, Mountain View, CA 94043 (US); BURFORD,
Neil [US/US]; 105 Whitewood Circle, Durham, CT 06422
(US); BATRA, Sajeev [US/US]; 555 El Camino Real
#709, Sunnyvale, CA 94087 (US); POLICKY, Jennifer,
J. [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US);

(21) International Application Number: PCT/US91/13862

(22) International Filing Date: 27 April 2001 (27.04.2001)

(25) Filing Language: English

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

(26) Publication Language: English

(30) Priority Data:
60/200,184 28 April 2000 (28.04.2000) US
60/201,575 4 May 2000 (04.05.2000) US
60/202,090 4 May 2000 (04.05.2000) US
60/210,232 6 June 2000 (06.06.2000) US
60/220,553 25 July 2000 (25.07.2000) US(81) Designated States (national): AE, AG, AI, AM, AU, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, FR, GE, GR, GM, HK,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).(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): LAL, Preeti
[IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US);
YUE, Henry [US/US]; 825 Lois Avenue, Sunnyvale, CA
94087 (US); TANG, Y. Tim [US/US]; 4230 Ramnick
Court, San Jose, CA 95118 (US); LU, Dyung, Aina,
M. [US/US]; 233 Coy Drive, San Jose, CA 95123 (US);
AZIMZAI, Yalda [US/US]; 5518 Boulder Canyon Drive,
Castro Valley, CA 94552 (US); AU-YOUNG, Janice
[US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005
(US); HILLMAN, Jennifer, L. [US/US]; 230 Monroe
Drive #17, Mountain View, CA 94040 (US); BAUGHN,
Mariah, R. [US/US]; 14244 Santiago Road, San Leandro,

Published:

— with international search report
before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments(88) Date of publication of the international search report:
20 June 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.

WO 01/83524 A3

(54) Title: RNA METABOLISM PROTEINS

(57) Abstract: The invention provides human RNA metabolism proteins (RMEP) and polynucleotides which identify and encode RMEP. 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 RMEP.

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/13862
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N5/12 C07K14/47 C12N5/10 C07K16/18 C12Q1/68 G01N33/68 G01N33/53 C12P21/02		
According to International Patent Classification (IPC) or to both national classification and IPC.		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K C12Q G01N C12P		
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, BIOSIS, EPO-Internal, WPI Data		
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] ACCESSION NO: AK000598. 22 February 2000 (2000-02-22) SUGANO S., ET AL.: "Homo sapiens cDNA FLJ20591 fls, clone KAT09018." XP002187893 the whole document	1-5, 9, 11, 12, 45, 92
X	ALLMANG CHRISTINE ET AL: "The yeast exosome and human PM-ScI are related complexes of 3' 5' endonucleases." GENES & DEVELOPMENT, vol. 13, no. 16, 15 August 1999 (1999-08-15), pages 2148-2158, XP002187891 ISSN: 0890-9369 the whole document	1-5, 9, 11, 12, 45, 92
<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 "C" 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		Date of making of the international search report
24 January 2002		25. 04. 2002
Name and mailing address of the ISA European Patent Office, P.O. Box 5818 Patentlaan 2 NL 2280 HV Rijswijk Tel: (+31-70) 340-2540, Te: 01 651 epo nl Fax: (+31-70) 340-2016		Authorized officer MADDOX, A

Form PCT/ISA/210 (second sheet) (July 1999)

INTE NATIONAL SEARCH REPORT

International Application No.
PCT/US 01/13862

C.(Continuation) 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] ACCESSION NO: AW593482, 23 March 2000 (2000-03-23) "hg16h04.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2945815 3' similar to WP:B0564.1 CE20469 RIBONUCLEASE PH-LIKE PROTEIN B0564.1 ; , mRNA sequence." XP002187894 the whole document	1-5,9, 11,12, 45,92
X	DATABASE EMBL [Online] ACCESSION NO: A1652670, 5 May 1999 (1999-05-05) "wb30d08.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2307183 3' similar to WP:B0564.1 CE05175 RIBONUCLEASE PH-LIKE PROTEIN B0564.1 ; , mRNA sequence." XP002187895 the whole document	1-5,9, 11,12, 45,92
X	DATABASE EMBL [Online] ACCESSION NO: A1346595, 7 January 1999 (1999-01-07) "cp51h02.x1 NCI CGAP Co8 Homo sapiens cDNA clone IMAGE:1926579 3' similar to WP:B0564.1 CE05175 RIBONUCLEASE PH LIKE ; , mRNA sequence." XP002187896 the whole document	1-5,9, 11,12, 45,92
X	DATABASE EMBL [Online] ACCESSION NO: A1770040, 30 June 1999 (1999-06-30) "wi79c09.x1 NCI CGAP Kid12 Homo sapiens cDNA clone IMAGE:2399536 3' similar to WP:B0564.1 CE05175 RIBONUCLEASE PH-LIKE PROTEIN B0564.1 ; , mRNA" XP002187897 the whole document	1-5,9, 11,12, 45,92
P,X	DATABASE EMBL [Online] ACCESSION NO: AF281133, 10 July 2000 (2000-07-10) BROUWER R., ET AL.: "Homo sapiens exosome component Rrp41 mRNA, complete cds." XP002187898 the whole document -& "Three novel components of the human exosome." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 276, no. 9, 2 March 2001 (2001-03-02), pages 6177-6184, XP002187892 --- -/--	1-5,9, 11,12, 45,92

5

Form PCT/ISA/215 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US 01/13862

C (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 00 55173 A (HUMAN GENOME SCIENCES INC ;ROSEN CRAIG A (US); RUBEN STEVEN M (US)) 21 September 2000 (2000-09-21) See SEQ ID NOS:246 and 564 ---	1-7, 9-12,18, 25, 29-45,92
P,X	WO 00 58473 A (CURAGEN CORP ;LEACH MARTIN (US); SHIMKETS RICHARD A (US)) 5 October 2000 (2000-10-05) See SEQ ID NOS: 5333 and 5334 ---	1-7, 9-12,18, 25,26, 29-45,92
P,X	WO 00 55174 A (HUMAN GENOME SCIENCES INC ;ROSEN CRAIG A (US); RUBEN STEVEN M (US)) 21 September 2000 (2000-09-21) See SEQ ID NOS: 480 and 1420 ---	1-7,9, 10,18, 25,26, 30-42
E	WO 01 87917 A (CHEN RUI HONG ;HYSEQ INC (US); WANG JIAN RUI (US); WEHRMAN TOM (US)) 22 November 2001 (2001-11-22) see SEQ ID NO:65 page 109 ---	1-7,9, 10,18, 25,26, 30-42
A	WO 98 23744 A (INCYTE PHARMA INC ;BANDMAN OLGA (US); GOLI SURYA K (US)) 4 June 1998 (1998-06-04) the whole document ---	
A	DATABASE EMBL [Online] ACCESSION NO: AF027150, 5 November 1997 (1997-11-05) FISCHER, U., ET AL.: "Homo sapiens survival of motor neuron protein interacting protein 1 (SIPI) mRNA, complete cds." XP002188024 the whole document -& LIU QING ET AL: "The spinal muscular atrophy disease gene product, SMN, and its associated protein SIPI are in a complex with spliceosomal snRNP proteins." CELL, vol. 90, no. 6, 1997, pages 1013-1021, XP002187934 the whole document -----	

5

Form PCT/ISA/210 (continuation of second sheet) July 1999

INTERNATIONAL SEARCH REPORT	International application No. PCT/US 01/13862
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.: 20, 21, 23, 24 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically: see FURTHER INFORMATION sheet PCT/ISA/210	
3. <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.: 1-44 (partially); 45, 92 (complete)	
Remark on Protest <div style="float: right;"> <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. </div>	

International Application No. PCT/US 01/13862

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claim(s) 32-34 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
Although claim 18 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.2

Claims Nos.: 20,21,23,24

Present claims 20,21,23,24 relate to a product/compound/method defined by reference to a desirable characteristic or property, namely agonists and antagonists of the polypeptide of claim 1.
The claims cover all products/compounds/methods having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such products/compounds/methods. 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 product/compound/method by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently no search has been carried out for claims 20,21,23,24.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an international Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

International Application No. PCT/US 01/13862

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: Invention 1: Claims 1-44 all partially and 45 and 92 both completely

Isolated polypeptide and corresponding polynucleotide as defined by SEQ ID NOS:1 and 48, cells transformed with, methods for producing, antibodies to and methods for preparing them, methods for detecting target polynucleotide, compositions, methods for treating a disease condition, method for screening agonists, antagonists, binding compounds, activity modulators, and compounds for altering expression of target polynucleotide, method for assessing toxicity of a test compound, diagnostic tests, all based on abovementioned polypeptide and polynucleotide.

2. Claims: Inventions 2-47. Claims 1-44 all partially and 46-91 and 93-138 in so far as applicable

Isolated polypeptide and corresponding polynucleotide as defined by SEQ ID NOS:2 and 49 for invention 2 through to SEQ ID NOS:47 and 94 for invention 47, cells transformed with, methods for producing, antibodies to and methods for preparing them, methods for detecting target polynucleotide, compositions, methods for treating a disease condition, method for screening agonists, antagonists, binding compounds, activity modulators, and compounds for altering expression of target polynucleotide, method for assessing toxicity of a test compound, diagnostic tests, all based on abovementioned polypeptide and polynucleotide.

INTERNATIONAL SEARCH REPORT				International Application No.	
Information on patent family members				PCT/US 01/13862	
Patent document cited in search report	Publication type	Publication date	Patent family member(s)	Publication date	
WO 0055173	A	21-09-2000	AU 3395900 A	04-10-2000	
			AU 3617600 A	04-10-2000	
			AU 3617700 A	04-10-2000	
			AU 3619400 A	04-10-2000	
			AU 3619500 A	04-10-2000	
			AU 3869400 A	04-10-2000	
			EP 1168917 A2	09-01-2002	
			EP 1165588 A1	02-01-2002	
			EP 1169469 A1	09-01-2002	
			EP 1165589 A1	02-01-2002	
			EP 1159420 A1	05-12-2001	
			EP 1163358 A1	19-12-2001	
			WO 0055173 A1	21-09-2000	
			WO 0055350 A1	21-09-2000	
			WO 0055351 A1	21-09-2000	
			WO 0055180 A2	21-09-2000	
			WO 0055174 A1	21-09-2000	
			WO 0055320 A1	21-09-2000	
			US 2002039764 A1	04-04-2002	
WO 0058473	A	05-10-2000	AU 3774500 A	16-10-2000	
			EP 1165784 A2	02-01-2002	
			WO 0058473 A2	05-10-2000	
WO 0055174	A	21-09-2000	AU 3395900 A	04-10-2000	
			AU 3617600 A	04-10-2000	
			AU 3617700 A	04-10-2000	
			AU 3619400 A	04-10-2000	
			AU 3619500 A	04-10-2000	
			AU 3869400 A	04-10-2000	
			EP 1168917 A2	09-01-2002	
			EP 1165588 A1	02-01-2002	
			EP 1169469 A1	09-01-2002	
			EP 1165589 A1	02-01-2002	
			EP 1159420 A1	05-12-2001	
			EP 1163358 A1	19-12-2001	
			WO 0055173 A1	21-09-2000	
			WO 0055350 A1	21-09-2000	
			WO 0055351 A1	21-09-2000	
			WO 0055180 A2	21-09-2000	
			WO 0055174 A1	21-09-2000	
			WO 0055320 A1	21-09-2000	
			US 2002039764 A1	04-04-2002	
WO 0187917	A	22-11-2001	AU 6300501 A	26-11-2001	
			WO 0187917 A1	22-11-2001	
			AU 6300601 A	26-11-2001	
			AU 6472401 A	26-11-2001	
			AU 7487101 A	26-11-2001	
			WO 0188088 A2	22-11-2001	
			WO 0188091 A2	22-11-2001	
			WO 0188092 A2	22-11-2001	
			AU 5352001 A	30-10-2001	
WO 9823744	A	04-06-1998	US 5962226 A	05-10-1999	
			AU 7410598 A	22-06-1998	
			WO 9823744 A1	04-06-1998	

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード(参考)
A 6 1 K 45/00	A 6 1 K 45/00	4 B 0 6 5
A 6 1 K 48/00	A 6 1 K 48/00	4 C 0 8 4
A 6 1 P 3/06	A 6 1 P 3/06	4 C 0 8 5
A 6 1 P 7/00	A 6 1 P 7/00	4 H 0 4 5
A 6 1 P 25/00	A 6 1 P 25/00	
A 6 1 P 29/00	A 6 1 P 29/00	
A 6 1 P 31/10	A 6 1 P 31/10	
A 6 1 P 31/14	A 6 1 P 31/14	
A 6 1 P 31/18	A 6 1 P 31/18	
A 6 1 P 31/20	A 6 1 P 31/20	
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 37/00	A 6 1 P 37/00	
A 6 1 P 43/00	A 6 1 P 43/00	1 0 1
C 0 7 K 1/22	A 6 1 P 43/00	1 1 1
C 0 7 K 14/47	C 0 7 K 1/22	
C 0 7 K 16/18	C 0 7 K 14/47	
C 0 7 K 19/00	C 0 7 K 16/18	
C 1 2 N 1/15	C 0 7 K 19/00	
C 1 2 N 1/19	C 1 2 N 1/15	
C 1 2 N 1/21	C 1 2 N 1/19	
C 1 2 N 5/10	C 1 2 N 1/21	
C 1 2 N 15/02	C 1 2 P 21/02	C
C 1 2 P 21/02	C 1 2 P 21/08	
C 1 2 P 21/08	C 1 2 Q 1/02	
C 1 2 Q 1/02	C 1 2 Q 1/68	A
C 1 2 Q 1/68	G 0 1 N 33/15	Z
G 0 1 N 33/15	G 0 1 N 33/50	Z
G 0 1 N 33/50	G 0 1 N 33/53	D
G 0 1 N 33/53	G 0 1 N 33/53	M
G 0 1 N 33/566	G 0 1 N 33/566	
	C 1 2 N 15/00	C
	C 1 2 N 5/00	A
	A 6 1 K 37/02	

(31)優先権主張番号 60/210,232

(32)優先日 平成12年6月6日(2000.6.6)

(33)優先権主張国 米国(US)

(31)優先権主張番号 60/220,553

(32)優先日 平成12年7月25日(2000.7.25)

(33)優先権主張国 米国(US)

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CR,CU,CZ,DE,DK,DM,DZ,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,SD,SE,S

G, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW

- (72)発明者 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 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 リュ、デュング・アイナ・エム
アメリカ合衆国カリフォルニア州 9 5 1 2 3 ・ サンノゼ・コイドライブ 2 3 3
- (72)発明者 アジムザイ、ヤルダ
アメリカ合衆国カリフォルニア州 9 4 5 5 2 ・ カストロバレー・ボールダーキャニオンドライブ
5 5 1 8
- (72)発明者 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 ・ ブリスベン・ゴールデンイーグルレーン 2 3 3
- (72)発明者 ヒルマン、ジェニファー・エル
アメリカ合衆国カリフォルニア州 9 4 0 4 0 ・ マウンテンビュー・# 1 7 ・ モンロードライブ 2
3 0
- (72)発明者 ボーゲン、マライア・アール
アメリカ合衆国カリフォルニア州 9 4 5 7 7 ・ サンレアンドロ・サンティアゴロード 1 4 2 4 4
- (72)発明者 ヤオ、モニック・ジー
アメリカ合衆国カリフォルニア州 9 4 0 4 3 ・ マウンテンビュー・フレデリックコート 1 1 1
- (72)発明者 パーフォード、ニール
アメリカ合衆国コネチカット州 0 6 4 2 2 ・ ダラム・ワイルドウッドサークル 1 0 5
- (72)発明者 パトラ、サジープ
アメリカ合衆国カリフォルニア州 9 4 0 8 7 ・ サニーベイル・# 7 0 9 ・ エルカミーノリアル 5
5 5
- (72)発明者 ポリッキー、ジェニファー・ジェイ
アメリカ合衆国カリフォルニア州 9 5 1 1 8 ・ サンノゼ・ジャービスコート 1 5 1 1

F ターム(参考) 2B030 CA15 CA17 CA19

2G045 AA34 AA35 BB01 BB20 BB41 BB50 BB51 DA13 DA14 DA36

FB05 FB08 FB12 GC15

4B024 AA01 AA11 AA20 BA53 BA61 BA80 CA01 CA03 CA04 CA07

CA09 CA12 CA20 DA01 DA02 DA05 DA11 DA12 EA02 EA04

FA02 GA11 HA03 HA04 HA11 HA13 HA14 HA15 HA19 HA20

4B063 QA01 QA05 QA13 QA17 QA19 QQ02 QQ21 QQ41 QQ42 QQ43

QQ52 QQ53 QQ61 QQ79 QQ89 QR08 QR32 QR35 QR40 QR42

QR48 QR55 QR56 QR62 QR77 QR80 QR84 QS16 QS25 QS33

QS34 QS36 QX01 QX02

4B064 AG01 AG27 CA01 CA10 CA19 CA20 CC01 CC24 CE12 DA01

DA13

4B065 AA01X AA58X AA72X AA87X AA92X AA93Y AB01 AB05 AC14 BA02

BA08 CA24 CA25 CA44 CA46

4C084 AA03 AA07 AA13 AA17 BA01 BA02 BA08 BA21 BA22 BA23

CA18 DC50 NA14 ZA011 ZA021 ZA061 ZA151 ZA161 ZA181 ZA221

ZA331 ZA361 ZA451 ZA511 ZA531 ZA551 ZA591 ZA661 ZA681 ZA691

ZA751 ZA811 ZA891 ZA911 ZA941 ZA961 ZA971 ZB051 ZB111 ZB131

ZB151 ZB211 ZB212 ZB261 ZB271 ZB321 ZB331 ZB351 ZB371 ZB381

ZC021 ZC061 ZC311 ZC331 ZC351 ZC412 ZC551

4C085 AA13 AA14 AA16 BB11 CC21 DD62 EE01

4H045 AA10 AA11 AA20 AA30 BA10 BA41 CA40 DA75 DA76 DA86

EA20 EA50 FA71 FA72 FA74 GA26