

(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2004-530409

(P2004-530409A)

(43) 公表日 平成16年10月7日(2004.10.7)

(51) Int.Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A 2 G O 4 5
A 6 1 K 38/00	A 6 1 K 45/00	4 B O 2 4
A 6 1 K 38/45	A 6 1 P 1/00	4 B O 2 9
A 6 1 K 45/00	A 6 1 P 1/04	4 B O 6 3
A 6 1 P 1/00	A 6 1 P 1/06	4 B O 6 4
審査請求 未請求 予備審査請求 有 (全 279 頁) 最終頁に続く		

(21) 出願番号	特願2002-537772 (P2002-537772)	(71) 出願人	301005050
(86) (22) 出願日	平成13年10月26日 (2001.10.26)		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15年4月24日 (2003.4.24)		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/US2001/049670		
(87) 国際公開番号	W02002/034783	(74) 代理人	100089266
(87) 国際公開日	平成14年5月2日 (2002.5.2)		弁理士 大島 陽一
(31) 優先権主張番号	60/244, 017	(72) 発明者	ワレン、ブリジット、エイ
(32) 優先日	平成12年10月27日 (2000.10.27)		アメリカ合衆国カリフォルニア州95014・クーペルティノー・#2・パークウッドドライブ 10130
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/252, 855		
(32) 優先日	平成12年11月22日 (2000.11.22)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/251, 825		
(32) 優先日	平成12年12月7日 (2000.12.7)		
(33) 優先権主張国	米国 (US)		
最終頁に続く			

(54) 【発明の名称】 膜貫通タンパク質

(57) 【要約】

本発明は、ヒトの膜貫通タンパク質 (TMP)、および、TMPを同定しコードするポリヌクレオチドを提供する。本発明はまた、発現ベクター、宿主細胞、抗体、アゴニストおよびアンタゴニストをも提供する。本発明はまた、TMPの異常発現に関連する障害を診断、治療、または予防する方法をも提供する。

【特許請求の範囲】

【請求項 1】

以下の (a) 乃至 (d) からなる群から選択した単離されたポリペプチド。

(a) SEQ ID NO: 1-17 (配列番号 1 乃至 17) からなる群から選択したアミノ酸配列を含むポリペプチド

(b) SEQ ID NO: 1-17 を有する群から選択したアミノ酸配列と少なくとも 90% が同一であるような天然のアミノ酸配列を有するポリペプチド

(c) SEQ ID NO: 1-17 を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片

(d) SEQ ID NO: 1-17 を有する群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片 10

【請求項 2】

SEQ ID NO: 1-17 からなる群から選択したアミノ酸配列を有する、請求項 1 に記載の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする、単離されたポリヌクレオチド。

【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: 18-34 からなる群から選択したポリヌクレオチド配列を有する、請求項 4 に記載の単離されたポリヌクレオチド。 20

【請求項 6】

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

【請求項 7】

請求項 6 に記載の組換えポリヌクレオチドを用いて形質転換した細胞。

【請求項 8】

請求項 6 に記載の組換えポリヌクレオチドを含む遺伝形質転換体。

【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、 30

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

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

【請求項 10】

前記ポリペプチドが、SEQ ID NO: 1-17 からなる群から選択したアミノ酸配列を持つことを特徴とする、請求項 9 に記載の方法。

【請求項 11】

請求項 1 に記載のポリペプチドと特異結合するような単離された抗体。 40

【請求項 12】

以下の (a) 乃至 (e) からなる群から選択した単離されたポリヌクレオチド。

(a) SEQ ID NO: 18-34 からなる群から選択したポリヌクレオチド配列を持つポリヌクレオチド

(b) SEQ ID NO: 18-34 を有する群から選択したポリヌクレオチド配列と少なくとも 90% が同一であるような天然のポリヌクレオチド配列を有するポリヌクレオチド

(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド

(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド 50

(e) (a) ~ (d) のRNA等価物

【請求項13】

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

【請求項14】

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

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

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

【請求項15】

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

【請求項16】

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

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

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

【請求項17】

請求項1のポリペプチドと、薬剤として許容できる賦形剤とを有することを特徴とする組成物。

【請求項18】

前記ポリペプチドが、SEQ ID NO: 1-17からなる群から選択されたアミノ酸配列を含むことを特徴とする、請求項17の組成物。

【請求項19】

機能的TMPの発現低下に関連する疾患又は病状を治療する方法であって、そのような治療を必要とする患者に対して請求項17に記載の組成物を投与する過程を含むことを特徴とする方法。

【請求項20】

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

(a) 請求項1のポリペプチドを含むサンプルを化合物に曝すステップと、

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

【請求項21】

請求項20の方法で同定したアゴニスト化合物を有し、薬剤として許容できる賦形剤をも有する組成物。

【請求項22】

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

【請求項23】

請求項1のポリペプチドのアンタゴニストとしての有効性を確認するために化合物をスク

10

20

30

40

50

リーニングする方法であって、

- (a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝すステップと、
- (b) 前記サンプルにおいてアンタゴニスト活性を検出するステップとを含むことを特徴とするスクリーニング方法。

【請求項 2 4】

請求項 2 3 の方法で同定したアンタゴニスト化合物と、薬剤として許容できる賦形剤とを有する組成物。

【請求項 2 5】

機能的な T M 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、そのような治療を要する患者への請求項 2 4 の組成物の投与を含む治療方法。

10

【請求項 2 6】

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

- (a) 請求項 1 のポリペプチドを適切な条件下で少なくとも 1 つの試験化合物と混合する過程と、
- (b) 請求項 1 のポリペプチドの試験化合物との結合を検出し、それによって請求項 1 のポリペプチドに特異結合する化合物を同定する過程を含む方法。

【請求項 2 7】

請求項 1 のポリペプチドの活性をモジュレートする化合物をスクリーニングする方法であって、

- (a) 請求項 1 のポリペプチドの活性が許容される条件下で、請求項 1 のポリペプチドを少なくとも 1 つの試験化合物と混合する過程と、
- (b) 請求項 1 に記載のポリペプチドの活性を試験化合物の存在下で算定する過程と、
- (c) 試験化合物の存在下での請求項 1 のポリペプチドの活性を、試験化合物の不存在下での請求項 1 のポリペプチドの活性と比較する過程を含み、試験化合物の存在下での請求項 1 のポリペプチドの活性の変化が、請求項 1 のポリペプチドの活性をモジュレートする化合物を標示するような方法。

20

【請求項 2 8】

請求項 5 の配列を持つ標的ポリヌクレオチドの発現を改変するのに効果的な化合物をスクリーニングする方法であって、

- (a) 前記標的ポリヌクレオチドの発現に好適な条件下で、該標的ポリヌクレオチドを含むサンプルを化合物に曝露する過程と、
- (b) 前記標的ポリヌクレオチドの発現改変を検出する過程と、
- (c) 可変量の前記化合物の存在下と前記化合物の不存在下で、前記標的ポリヌクレオチドの発現を比較する過程とを含むことを特徴とする方法。

30

【請求項 2 9】

試験化合物の毒性を算定する方法であって、

- (a) 核酸を含む生物学的サンプルを前記試験化合物で処理する過程と、
- (b) 処理した前記生物学的サンプルの核酸と、請求項 1 2 のポリヌクレオチドの少なくとも 2 0 の連続するヌクレオチドを持つプローブをハイブリダイズさせる過程であって、このハイブリダイゼーションゼーションが、前記プローブと前記生物学的サンプル中の標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われ、前記標的ポリヌクレオチドが、請求項 1 2 のポリヌクレオチドまたはその断片のポリヌクレオチド配列を持つポリヌクレオチドである、前記過程と、
- (c) ハイブリダイゼーション複合体の収量を定量する過程と、
- (d) 前記処理された生物学的サンプル中のハイブリタイゼーション複合体の量を、処理されていない生物学的サンプル中のハイブリタイゼーション複合体の量と比較する過程とを含み、前記処理された生物学的サンプル中のハイブリタイゼーション複合体の量の差が、前記試験化合物の毒性を標示するような方法。

40

【請求項 3 0】

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

50

て、

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

【請求項 31】

請求項 11 の抗体であって、

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

10

【請求項 32】

請求項 11 の抗体と、許容できる賦形剤とを有する組成物。

【請求項 33】

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

【請求項 34】

前記抗体が標識されることを特徴とする、請求項 32 の組成物。

【請求項 35】

20

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

【請求項 36】

請求項 11 の抗体の特異性を持つポリクローナル抗体を調製する方法であって、

(a) 抗体反応を誘発する条件下で、SEQ ID NO: 1-17 を有する群から選択したアミノ酸配列またはその免疫原性断片を含むポリペプチドを用いて動物を免疫化する過程と、

(b) 前記動物から抗体を単離する過程と、

(c) 前記単離された抗体を前記ポリペプチドでスクリーニングし、それによって、SEQ ID NO: 1-17 からなる群から選択したアミノ酸配列を持つポリペプチドに特異結合するポリクローナル抗体を同定する過程とを含むような方法。

30

【請求項 37】

請求項 36 の方法で産生したポリクローナル抗体。

【請求項 38】

請求項 37 のポリクローナル抗体と好適なキャリアとを含む組成物。

【請求項 39】

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

(a) 抗体反応を誘発する条件下で、SEQ ID NO: 1-17 を有する群から選択したアミノ酸配列またはその免疫原性断片を含むポリペプチドを用いて動物を免疫化する過程と、

(b) 前記動物から抗体産出細胞を単離する過程と、

(c) 前記抗体産出細胞と不死化した細胞とを融合して、モノクローナル抗体を産出するハイブリドーマ細胞を形成する過程と、

(d) 前記ハイブリドーマ細胞を培養する過程と、

(e) SEQ ID NO: 1-17 からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するような前記培養モノクローナル抗体から単離する過程とを含むことを特徴とする方法。

40

【請求項 40】

請求項 39 に記載の方法で産出したモノクローナル抗体。

【請求項 41】

50

請求項 40 に記載のモノクローナル抗体と適切なキャリアとを含む組成物。

【請求項 42】

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

【請求項 43】

組換え免疫グロブリンライブラリをスクリーニングすることにより産出されることを特徴とする、請求項 11 に記載の抗体。

【請求項 44】

S E Q I D N O : 1 - 17 を有する群から選択したアミノ酸配列を有するポリペプチドをサンプル中に検出する方法であって、

10

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

(b) 特異結合を検出する過程とを含み、該特異結合が、S E Q I D N O : 1 - 17 からなる群から選択したアミノ酸配列を有するポリペプチドがサンプル中に存在することを標示することを特徴とする方法。

【請求項 45】

或るサンプルからの、S E Q I D N O : 1 - 17 からなる群から選択したアミノ酸配列を有するポリペプチドを精製する方法であって、

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

20

(b) 前記サンプルから前記抗体を分離し、S E Q I D N O : 1 - 17 からなる群から選択したアミノ酸配列を有する精製ポリペプチドを得る過程、とを含むことを特徴とする方法。

【請求項 46】

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

【請求項 47】

ポリヌクレオチドを有する或るサンプルの発現プロフィールを作製する方法であって、

(a) サンプル中のポリヌクレオチドを標識化する過程と、

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

30

(c) サンプル中のポリヌクレオチドの発現を定量する過程とを含む方法。

【請求項 48】

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

【請求項 49】

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

40

【請求項 50】

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

【請求項 51】

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

【請求項 52】

50

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

【請求項 53】

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

【請求項 54】

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

【請求項 55】

請求項 48 に記載のアレイで、基板上の固有の物理的位置の各々が複数のヌクレオチド分子を含み、任意の単一の固有の物理的位置でのその複数のヌクレオチド分子は同一の配列を有し、基板上の固有の物理的位置の各々は、基板上の別の固有の物理的位置でのヌクレオチド分子の配列とは異なる配列を有するヌクレオチド分子を含むことを特徴とするアレイ。 10

【請求項 56】

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

【請求項 57】

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

【請求項 58】

SEQ ID NO: 1 のアミノ酸配列を有する請求項 3 に記載のポリペプチド。 20

【請求項 59】

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

【請求項 60】

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

【請求項 61】

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

【請求項 62】

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

【請求項 63】

SEQ ID NO: 1 のアミノ酸配列を有する請求項 8 に記載のポリペプチド。 30

【請求項 64】

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

【請求項 65】

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

【請求項 66】

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

【請求項 67】

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

【請求項 68】

SEQ ID NO: 1 のアミノ酸配列を有する請求項 13 に記載のポリペプチド。 40

【請求項 69】

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

【請求項 70】

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

【請求項 71】

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

【請求項 72】

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

【請求項 73】

SEQ ID NO: 18 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレ 50

オチド。

【請求項 7 4】

SEQ ID NO : 1 9 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 7 5】

SEQ ID NO : 2 0 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 7 6】

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

10

【請求項 7 7】

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

【請求項 7 8】

SEQ ID NO : 2 3 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 7 9】

SEQ ID NO : 2 4 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 8 0】

SEQ ID NO : 2 5 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

20

【請求項 8 1】

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

【請求項 8 2】

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

【請求項 8 3】

SEQ ID NO : 2 8 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

30

【請求項 8 4】

SEQ ID NO : 2 9 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 8 5】

SEQ ID NO : 3 0 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 8 6】

SEQ ID NO : 3 1 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

40

【請求項 8 7】

SEQ ID NO : 3 2 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 8 8】

SEQ ID NO : 3 3 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【請求項 8 9】

SEQ ID NO : 3 4 のポリヌクレオチド配列を有する請求項 1 2 に記載のポリヌクレオチド。

【発明の詳細な説明】

50

【0001】

(技術分野)

本発明は、膜貫通タンパク質の核酸配列およびアミノ酸配列に関する。本発明はまた、これらの配列を利用した、生殖障害、発達障害、心血管障害、神経障害、胃腸障害、脂質代謝、細胞増殖異常、および自己免疫／炎症疾患の、診断・治療・予防に関する。本発明は更に、膜貫通タンパク質の、核酸配列およびアミノ酸配列の発現における外来性化合物の効果についての評価に関する。

【0002】

(発明の背景)

真核生物が原核生物と異なる点は、細胞内の、膜に囲まれる多くの区画、例えばオルガネラや小胞を持つことである。真核生物の生化学を原核生物の生化学から区別する代謝反応の多くは、これらの区画内で起こる。詳細には、多くの細胞機能は非常にストリンジェントな反応条件を必要とし、オルガネラや小胞が各反応の区画化と隔離とを可能にしなければ、細胞質の各代謝プロセスが混乱し得る。これらのオルガネラの例に、ミトコンドリア、滑面小胞体、粗面小胞体、筋小胞体、およびゴルジ体がある。小胞の例には、食胞、リソソーム、エンドソーム、ペルオキシソーム、および分泌小胞がある。オルガネラと小胞とは、1層または2層の膜で囲まれる。

【0003】

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

【0004】

形質膜タンパク質

膜貫通タンパク質 (TMP) 類は、細胞外ドメインと、膜貫通ドメインと、細胞内ドメインによって特徴付けられる。TMPドメイン群は通常、15～25の疎水性アミノ酸からなり、これらのアミノ酸は、予測では1つのα螺旋コンフォメーションをとる。TMPを分類すると、バイトピック (bitopic) (タイプIおよびタイプII) タンパク質 (膜を1回貫通する) と、ポリトピック (polytopic) (タイプIIIおよびタイプIV) (Singer, S. J. (1990) Annu. Rev. Cell Biol. 6:247-96) タンパク質とがあり、後者には、複数の膜貫通セグメントがある。シグナル伝達に関わる細胞表面受容体タンパク質として作用するTMPとしては、成長因子受容体と分化因子受容体、またショウジョウバエの *pecanex* タンパク質と *fizzled* タンパク質、LIV-1 タンパク質、NF2 タンパク質、GNS1/SUR4 真核生物の膜内在性タンパク質などの、受容体と相互作用するタンパク質が含まれる。TMPはまた、イオンまたは代謝産物のトランスポーター (例えばギャップ結合チャネル (コネキシン類) とイオンチャネル類) として作用し、また、細胞接着タンパク質 (例えばレクチン類とインテグリン類とフィブロネクチン類) として作用する。TMPは、小胞とオルガネラとを形成する分子 (例えばカベオリン (caveolin) 類) として機能し、また、細胞認識分子 (例えば分化クラスター (CD) 抗原、糖タンパク質、およびムチン類) として機能する。

【0005】

親水性分子を膜内外に輸送する作用は、チャネルタンパク質類の存在によって促進される。これらは水性ポア (aqueous pores) を形成し、ポアは脂質2重層を貫通し得る。多くのチャネルは、複数のサブユニットの集合体によって形成されるタンパク質複合体から構成される。サブユニットの少なくとも1つは膜内在性タンパク質であり、これが寄与してポアを形成する。幾つかの場合、ポアは、ただ1つまたは少数の分子種のみ

10

20

30

40

50

が選択的に通過し得るように構成される。異なるタイプの膜チャンネルがいくつかあり、大きな差異は、それらの分布と選択性とである。各タイプは、(1) アクアポリン (aquaporins、水を輸送する) (2) タンパク質伝導チャンネル (小胞体膜を通してにタンパク質を輸送する) (3) ギャップ結合 (イオンと有機小分子の隣接細胞間での拡散を促進する)、および (4) イオンチャンネル (多様な膜を通るイオンフラックスを調節する) である。

【0006】

ギャップ結合 (別名コネクソン) は、原形質膜の特化した領域であり、膜貫通チャンネル群からなる。膜貫通チャンネルは、多くの組織の隣接細胞間で細胞質を化学的また電氣的に連結させて機能する。ギャップ結合は、興奮性の各組織内での活動電位の細胞間伝播のために電氣的シナプスとして機能する。非興奮性組織では、ギャップ結合の役割は組織の恒常性にあり、また、調整された生理応答や、代謝協同、成長制御、発生と分化との調節においても働く。

10

【0007】

各コネクソンは原形質膜の脂質2重層を貫通しており、また、6つの同一サブユニットであるコネキシン群から構成される。少なくとも14種の異なるコネキシン・タンパク質があり、それぞれ類似した構造を持つが、異なる組織分布を示す。構造的には、コネキシン類には短い細胞質内N末端ドメインがあり、4つの膜貫通領域 (M1、M2、M3、およびM4) につながっている。これらの膜貫通領域は、2つの細胞外ループと1つの細胞質内ループを分離するように配置され、その後にC末端ドメインがある。C末端ドメインは細胞質内ドメインであり、多様な長さを持つ (Cx26での20残基から、Cx56での260残基までがある)。M2-M3ループと、N末端およびC末端とは、細胞質内側に向いている。保存された領域は、各膜貫通領域と、2つの細胞外ループとである。両細胞外ループ内に、3つの保存されたシステイン群があり、これらはジスルフィド結合形成に関与する。両ループのシグネチャ・パターンは、次のいずれかであり、C-[DN]-T-x-Q-P-G-C-x-(2)-V-C-Y-D または C-x(3,4)-P-C-x(3)-[LIVM]-[DEN]-C-[FY]-[LIVM]-[SA]-[KR]-P (PDOC00341, Profilescan 並びに S. Rahman および W. H. Evans, (1991) J. Cell Sci. 100:567-578) である。可変領域は細胞質内ループとC末端領域とであり、種々のコネキシン類の調節に関与し得る (Hennemann, H. 他 (1992) J. Biol. Chem. 267:17225-17233; PRINTS PR00206 connexin signature; Yeager, M. 他, (1998) Curr. O. Structr. Biol. 8:517-524を参照)。

20

30

【0008】

ギャップ結合は、心筋収縮と平滑筋収縮との同調を助け、神経伝達を加速し、細胞外シグナルの伝播を助ける。ギャップ結合の開閉は、特定の刺激に応じて可能となる (例えば pH、 Ca^{+2} 、および cAMP)。ギャップ結合の効果的なポアのサイズは約1.5 nmであり、これは、小分子 (例えば1000ダルトン未満) が自由にポアを通じて拡散できるサイズである。輸送される分子の例に、イオンや低分子代謝産物やセカンドメッセンジャー (例えば Ca^{+2} および cAMP) がある。

40

【0009】

コネキシン類は、多くの疾患に関与する。メスのコネキシン37 (Cx37) 欠損マウスが不妊である原因は、卵母細胞-顆粒膜細胞シグナル伝達経路の欠除である。Cx43欠損マウスは生後すぐに死亡し、心臓の欠陥を示すが、この欠陥は、ヒトでのある数種の肺動脈狭窄を連想させる。Cx32における突然変異は、X染色体連鎖型シャルコーマリー-トウース病を伴い、この病気は、末梢神経系の運動神経と感覚神経のニューロパシーである。Cx26は胎盤において発現し、Cx26欠損マウスでは、或るグルコース類似体の、母体循環から胎児循環への経胎盤輸送が低下する。ヒトでは、Cx26は非症候群性常染色体感音難聴の最初の感受性遺伝子として同定されている。Cx31における突然

50

変異は常染色体優性聴覚障害に連鎖し（第2細胞外ループ内のナンセンス変異またはミスセンス変異）、また、優性遺伝する皮膚障害である変異性紅斑角皮症にも連鎖する（N末端ドメインまたはM2ドメイン内のミスセンス変異）（A. M. Simon, (1999) Trends Cell Biol. 9:169-170を参照）。Cx46はレンズ線維細胞内で発現し、Cx46欠損マウスは早期発症白内障を起こすが、これはヒト核性白内障に似る（Nicholson, S. M. および R. Bruzzone (1997) Curr. Biol. 7:R340-R344を参照）。

【0010】

形質膜タンパク質（MP）は2群に分けられ、その分類基準は膜からのタンパク質の抽出方法である。膜の外にあるタンパク質すなわち膜表在性タンパク質を放出させるには、極端なイオン強度またはpHを用いるか、尿素などの、タンパク質相互作用の攪乱物質を用いる。膜の内側にあるタンパク質すなわち膜内在性タンパク質を放出させる唯一の方法は、膜の脂質2重層を界面活性剤で溶解することである。

10

【0011】

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

20

【0012】

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

【0013】

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

30

【0014】

原形質膜の膜貫通タンパク質類としてはまた、細胞表面受容体もある。細胞表面受容体が認識する物質の例は、ホルモン、例えばカテコールアミン類（例えばエピネフリン、ノルエピネフリン、およびヒスタミン）や、ペプチドホルモン類（例えばグルカゴン、インシュリン、ガストリン、セクレチン、コレシストキニン、副腎皮質刺激ホルモン、卵胞刺激ホルモン、黄体形成ホルモン、甲状腺刺激ホルモン、副甲状腺ホルモン、およびバソプレッシン）や、成長因子および分化因子（例えば上皮成長因子、線維芽細胞成長因子、トランスフォーミング成長因子、インシュリン様成長因子、血小板由来成長因子、神経成長因子、コロニー刺激因子、およびエリスロポエチン）や、サイトカイン類（例えばケモカイン類、インターロイキン類、インターフェロン類、および腫瘍壊死因子）や、小ペプチド因子（例えばボンベシン、オキシトシン、エンドセリン、アンジオテンシンII、血管作用性小腸ペプチド、およびブラジキニン）や、神経伝達物質（例えば神経ペプチドY、ニューロテンシン、ニューロメディンN、メラノコルチン類、オピオイド類（例えばエンケファリン類、エンドルフィン類、およびダイノルフィン類）、ガラニン、ソマトスタチン、およびタキキニン類）、並びに、循環系に運ばれるシグナル伝達分子（例えばアンジオテンシン、補体、カルシトニン、エンドセリン類、およびホルミルメチオニルペプチド類）がある。免疫系細胞上の細胞表面受容体は、抗原、抗体、および主要組織適合性複合体（MHC）結合ペプチドを認識する。他の細胞表面受容体はリガンド類に結合し、リガ

40

50

ンドはその細胞によって取り込まれる。この受容体介在性エンドサイトーシスは、低密度リポタンパク質 (LDL)、トランスフェリン、グルコース末端糖タンパク質、マンノース末端糖タンパク質、ガラクトース末端糖タンパク質、免疫グロブリン、ホスホビテロゲン (phosphovitellogenin) 類、フィブリン、プロテイナーゼ阻害剤複合体、プラスミノゲン活性化因子、およびトロンボスポンジンの取り込み時に機能する (Lodish, H. 他 (1995) Molecular Cell Biology, Scientific American Books, New York, NY, 723 ページ; および Mikhailenko, I. 他 (1997) J. Biol. Chem. 272: 6784-6791 を参照)。

【0015】

多くの細胞表面受容体には7つの膜貫通領域があり、またリガンドに結合する細胞外のN末端と、Gタンパク質と相互作用する細胞質内のC末端がある (Strosberg, A. D. (1991) Eur. J. Biochem. 196: 1-10)。システインリッチドメイン群は、細胞表面受容体の2つのファミリーに見られ、それはLDL受容体ファミリーと、腫瘍壊死因子受容体/神経成長因子受容体 (TNFR/NGFR) ファミリーとである。7つの連続したシステインリッチリピート群 (約40アミノ酸) はLDL受容体のN末端細胞外領域内にLDLとカルシウムとの結合部位を形成し、同様のリピートは脊椎動物の超低密度リポタンパク質受容体に見られ、また、脊椎動物低密度リポタンパク質受容体関連タンパク質1 (LRP1) (別名 α_2 マクログロブリン受容体) や、脊椎動物低密度リポタンパク質受容体関連タンパク質2 (別名 gp330 または

10

20

30

40

50

【0016】

LDL受容体は膜内在性タンパク質であり、コレステロールを血中から除去することによって脂質の取込を行う。肝臓および腸以外の殆どの細胞は、コレステロールを血中から取り込み、細胞自体では合成しない。細胞表面LDL受容体にはLDL粒子が結合し、粒子はエンドサイトーシスにより細胞内に取り込まれる (Meyers, R. A. (1995) Molecular Biology and Biotechnology, VCH Publishers, New York NY, 494-501 ページ)。LDL受容体の欠除は、家族性高コレステロール血症疾患を生じ、血漿コレステロールレベルを増加させ、最終的にはアテローム性硬化をもたらす (Stryer, L (1995) Biochemistry, W. H Freeman, New York. 691-702 ページ)。

【0017】

Gタンパク質共役受容体

Gタンパク質共役受容体 (GPCR) 類は膜内在性タンパク質のスーパーファミリーを構成し、細胞外シグナルを伝達する。GPCR類の例には、生体アミン類の受容体や、炎症の脂質媒介物、ペプチドホルモン類、および感覚シグナル媒介物の受容体がある。

【0018】

構造的には、これらの高度に保存された受容体には、7つの疎水性膜貫通 (蛇行) 領域群と、第2および第3の細胞外ループ間のシステインジスルフィド架橋群と、細胞外N末端と、細胞質内C末端とがある。3つの細胞外ループ群は、3つの細胞内ループ群と交互に現れ、7つの膜貫通領域を連結している。これらのタンパク質の、最も保存された部分は、各膜貫通領域および、第1と第2の細胞質内ループである。システインジスルフィド架橋は、第2と第3の細胞外ループを接続する。1つの保存された、酸性残基 - Arg 残基 - 芳香族残基のトリプレットが第2細胞質内ループ内にあり、Gタンパク質と相互作用し得る。GPCRコンセンサスパターンは、このスーパーファミリーに属する殆どのタンパク

質を特徴付ける (ExPASy PROSITE document PS00237 ; および Watson, S. および S. Arkininstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego, CA, 2-6ページ)。GPCRをコードする遺伝子群の突然変異と転写活性の変化に伴う神経障害に、統合失調症 (精神分裂病)、パーキンソン病、アルツハイマー病、薬物依存、および摂食障害がある。若年発達および繁殖能2 (jdf-2) 遺伝子座、別名 runty-jerky-sterile (rjs) は、HERC2における欠失と点突然変異とに関連し、HERC2は、或るグアニンヌクレオチド交換因子タンパク質をコードする遺伝子であって小胞輸送に関わる (Walkowicz, M. 他 (1999) *Mamm. Genome* 10: 870-878)。

10

【0019】

或るGPCRはFPとして知られ、プロスタグランジン $F_{2\alpha}$ (PGF $_{2\alpha}$) の受容体である。プロスタグランジン類の所属は、生理的応答と炎症性応答での、天然のパラ分泌/自己分泌媒介物の1大ファミリーである。PGF $_{2\alpha}$ の果たす役割は幾つかの組織の応答にあり、その組織は例えば生殖器官、肺、骨、および心臓であり、その応答には子宮筋収縮の刺激と、黄体の分解と、気管支収縮とがある。或るFP会合分子 (FPRP) はFPと同時に精製され、発現は、PGF $_{2\alpha}$ の生理的役割が説明された各組織でのみ起こる。FPRPは、予測では膜貫通タンパク質であり、グリコシル化した細胞外免疫グロブリン

20

【0020】

スカベンジャー受容体

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

40

【0021】

テトラスパンファミリータンパク質

膜4回貫通型スーパーファミリー (TM4SF)、別名テトラスパンファミリーは多重遺伝子族であり、タイプIII膜内在性タンパク質をコードする (Wright, M. D. および Tomlinson, M. G. (1994) *Immunol. Today* 15: 588-594)。TM4SFは、細胞膜を4回貫通する膜タンパク質からなる。TM4SFのメンバーの例は、血小板および内皮細胞の膜タンパク質と、黒色腫関連抗原、白血球表面糖タンパク質、結腸癌腫抗原、腫瘍関連抗原、および住血吸虫表面タンパク質がある (Jankowski, S. A. (1994) *Oncogene* 9: 1205-1211)。TM4SFの各メンバーは、約25~30%のアミノ酸配列同一性を

50

互いに共有する。

【0022】

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

【0023】

腫瘍抗原

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

【0024】

イオンチャネル

イオンチャネルは、事実上、身体のあるあらゆる細胞の原形質膜内に見られる。例えば、塩素イオンチャネルは多様な細胞機能を媒介し、その機能の例には、膜電位の調節と、上皮膜を通過するイオンの吸収と分泌とがある。塩素イオンチャネルが、ゴルジ体とエンドサイトーシス小胞の細胞内の膜にある場合、オルガネラのpHをも調節する（例えばGreger, R.（1988）Annu. Rev. Physiol. 50: 111-122を参照）。塩素イオンチャネル類の電気生理的および薬学的特性、例えばイオンコンダクタンス、電流-膜電位関係、およびモジュレーター類への感受性は、筋肉、神経細胞、線維芽細胞、上皮細胞、および白血球に、別々の塩素イオンチャネルがあることを示唆する。

【0025】

多くのチャネルは、1個以上のタンパク質キナーゼによるリン酸化部位を有する。これらのタンパク質キナーゼには、タンパク質キナーゼA、タンパク質キナーゼC、カゼインキナーゼII、およびチロシンキナーゼがあり、これらはすべて、細胞におけるイオンチャネル活性を調節する。膜タンパク質の不適切なリン酸化は、細胞周期の進行と細胞分化における病理的变化と相関している。細胞周期における変化は、アポトーシスの誘発や癌の誘発に関連することが示されている。細胞分化における変化は、生殖系、免疫系、および骨格筋の、疾患や障害に関連することが示されている。

【0026】

小脳顆粒神経細胞は非不活化カリウム電流を持ち、これは神経伝達物質からの受容体刺激による発火頻度を調節し、また、静止膜電位を制御する。非不活化電流を示すカリウムチャネル類の例に、etherago-go（EAG）チャネルがある。KCR1と名付けられた膜タンパク質は、ラットEAGと特異的にそのC末端領域で結合し、小脳非不活化カリウム電流を調節する。KCR1は、12の膜貫通ドメイン群と、細胞内にアミノ酸末端およびカルボキシル末端を含むと予測される。構造的特徴からは、これらの膜貫通領域はトランスポータースーパーファミリーの領域に類似してみえるが、KCR1と既知のトランスポーターとの間に相同性は見られないことから、KCR1は新規クラスのトランスポーターに属することが示唆される。KCR1は、非不活化カリウムチャネルの調節成分のようである（Hoshi, N. 他（1998）J. Biol. Chem. 273: 23080-23085）。

【0027】

プロトンポンプ

プロトンATPアーゼは大きなクラスの膜タンパク質であり、ATP加水分解エネルギーを用いて電気化学的プロトン勾配を膜の内外に発生させる。この発生した勾配は、他のイオン（ Na^+ 、 K^+ 、または Cl^- ）の膜を通しての輸送、または、オルガネラpHの維持のために使用し得る。プロトンATPアーゼ類は更に、ミトコンドリアF₁-ATPア

アーゼ、原形質膜ATPアーゼ、および液胞ATPアーゼに分類される。液胞ATPアーゼは、エンドサイトーシスとエキソサイトーシスとの各過程に関与する多様な小胞における、酸性pHを確立し維持する(Mellman, I. 他 (1986) Ann. Rev. Biochem. 55:663-700)。

【0028】

プロトンと共役する、膜12回貫通ドメイントランスポータ、例えばPEPT 1およびPEPT 2はペプチド類の、胃腸での吸収と、腎臓での再吸収とに関与し、これには電気化学的プロトン勾配を駆動力として用いる。別のタイプのペプチドトランスポータであるTAPトランスポータは、TAP 1とTAP 2から成るヘテロ二量体であり、抗原のプロセッシングに関わる。ペプチド抗原はTAPによって小胞体の膜を通して輸送され、その結果、ペプチド抗原は、細胞表面においてMHC分子と会合して発現し得る。各TAPタンパク質は複数の疎水性膜貫通セグメントと、1つの高度に保存されたATP結合カセットから構成される(Boll, M. 他 (1996) Proc. Natl. Acad. Sci. 93:2842-289)。病原微生物、例えば単純ヘルペスウイルスは、TAPが媒介するペプチド輸送の阻害因子をコードする可能性があり、その目的は、免疫の監視を逃れることである(Marusina, K. および Manaco, J. J. (1996) Curr. Opin. Hematol. 3:19-26)。

10

【0029】

ABC輸送体

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

20

【0030】

細胞間伝達に関連する膜タンパク質

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

30

40

【0031】

シナプトブレビン(synaptobrevin)類はシナプスの小胞会合膜タンパク質(VAMP)であり、初めはラット脳で発見された。これらのタンパク質は、神経細胞に限定され、一細胞の細胞膜からシナプスを通り他の細胞膜への、小胞の移動において機能すると初めは考えられていた。シナプトブレビン類は、今では、構成的な複数の小胞輸送経路に存在および機能し、多くの非神経細胞タイプの、受容体が媒介するエンドサイトーシスおよびエキソサイトーシス経路に関与することがわかった。この調節される小胞輸送経路は、シナプトブレビン分子を切断するクロストリジウム神経毒素の、高度に特異的な作用によってブロックされ得る。

50

【0032】

In vitro での研究が、多様な細胞膜に関して行われ (Galli 他 (1994) J. Cell. Biol. 125:1015-24; Link 他 (1993) J. Biol. Chem. 268:18423-6)、それらの研究は、VAMP は広く分布することを示している。これらの重要な膜輸送タンパク質は、発生期の軸索伸展にもエキソサイトーシスを通じて関与し、また、神経伝達物質と調節ペプチドとの放出や、エンドサイトーシスにも関与するようである。エンドサイトーシス性小胞輸送としては、細胞内イベント、例えば核膜、小胞体、ゴルジ体、および多様な封入体 (例えばペルオキシソームまたはリソソーム) の、融合および分裂がある。エンドサイトーシスの各プロセスは、真核細胞では、酵母、線虫 (Caenorhabditis elegans)、ショウジョウバエ、および哺乳動物に渡り、普遍的なようである。

10

【0033】

VAMP-1B は、細胞内ターゲティングに関与し、また、VAMP-1A のアイソフォームである (Isenmann, S. 他 (1998) Mol. Biol. Cell 9:1649-1660)。4種の更なるスプライス変異体 (VAMP-1C から F まで) が最近、同定された。各変異体には可変配列群があるが、その位置はきわめて末端の C 末端部のみであることから、C 末端部が、小胞ターゲティングに重要であると示唆される (Berglund, L. 他 (1999) Biochem. Biophys. Res. Commun. 264:777-780)。

【0034】

リソソーム類は、自己貪食時の細胞内物質の分解の部位であり、また、エンドサイトーシス後の細胞外分子の分解の場でもある。リソソーム酵素は、トランスゴルジ網から出芽する小胞内に詰められる。これらの小胞はエンドソームと融合して成熟リソソームを形成し、成熟リソソーム内では、エンドサイトーシスによって取り込まれた物質の加水分解性消化が起こる。リソソームはオートファゴソームと融合し、独特な区画を形成し得る。この区画の中では、オルガネラや他の細胞内成分の分解が起こる。

20

【0035】

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

30

【0036】

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

40

50

タンパク質を発見した。

【0037】

正常な胚発生と、生殖細胞成熟の制御とを変調する多数の分泌タンパク質は、それぞれの膜結合受容体と相互作用する。胚発生時の細胞運命は、アクチビン (activin) / TGF- β スーパーファミリーのメンバー、カドヘリン、IGF-2、および他のモルフォゲンによって決定される。また、生殖細胞と生殖組織との増殖、成熟、および再分化は、例えば IGF-2、インヒビン、アクチビン、およびフォリスタチン (follistatins) によって調節される (Petraglia, F. (1997) Placenta 18:3-8; Mather, J. P. 他 (1997) Proc. Soc. Exp. Biol. Med. 215:209-222)。

10

【0038】

小胞体膜タンパク質

正常に機能する真核細胞に必要な条件は、新規に合成された全てのタンパク質が正しく折りたたまれ、修飾され、特定の細胞内部位および細胞外部位に送達されることである。新規合成された膜タンパク質および分泌タンパク質は、細胞の選別・分配網に入る。これは合成時または合成直後に起こり、その後、タンパク質は細胞内外の特定の位置に送られる。この過程の最初の区画である小胞体 (ER) においてタンパク質が受ける修飾の例に、グリコシル化、ジスルフィド結合形成、およびオリゴマー形成がある。修飾されたタンパク質は次に、一連の、膜に囲まれた区画を通して輸送される。この区画の例にゴルジ体の多様な層板があり、この区画では更なる糖質修飾が起こる。区画間の輸送を生じる手段は、小胞の出芽と融合とである。一旦、分泌経路に入るとタンパク質は、細胞表面に達するのに膜を横断する必要は無い。

20

【0039】

過半数のタンパク質はER内でプロセシングされオルガネラから運び出されるが、幾つかは保持される。ER内に保持されるためのシグナルは、哺乳動物細胞では、テトラペプチド配列であるKDELからなり、その位置は、常在ER膜タンパク質のカルボキシル末端である (Munro, S. (1986) Cell 46:291-300)。KDEL配列を持つタンパク質はERを出るが、すぐに初期ゴルジ層板から回収され、ERに戻る。一方、このシグナルを欠くタンパク質は、分泌経路を進む。

【0040】

細胞の分泌経路の妨害は、幾つかのヒト疾患に関与することが示唆されている。家族性高コレステロール血症では低密度リポタンパク質受容体がER内に残り、細胞表面に移動しない (Pathak, R. K. (1988) J. Cell Biol. 106:1831-1841)。 β アミロイド前駆体タンパク質 (β APP) の輸送とプロセシングとの変異は、推定上の小胞輸送タンパク質であるプレセニリン (presenilin) と関与しており、また、早期発症アルツハイマー病に関与し得る (Levy-Lahad, E. 他 (1995) Science 269:973-977)。ERに由来するカルシウム恒常性の変化に伴う疾患の例に、心筋症、心臓肥大、筋緊張性ジストロフィー、Brody病、Smith-McCort異形成、および糖尿病がある。

30

【0041】

ミトコンドリア膜タンパク質

ミトコンドリア電子伝達系 (すなわち呼吸鎖) は、一連の3種の酵素複合体であってミトコンドリア膜内にあり、NADHから酸素への電子伝達と、この酸化作用をATPの合成に共役させる過程 (酸化的リン酸化) とに関わる。ATPはその後、主要なエネルギー源を提供することにより、細胞の、エネルギーを必要とする多くの反応を駆動する。

40

【0042】

ミトコンドリア呼吸鎖の殆どのタンパク質成分は、核がコードする遺伝子の産物であり、これらはミトコンドリア内に移入されるが、他のタンパク質はミトコンドリア遺伝子の産物である。この呼吸鎖における酵素の欠損や変容した発現に関連して多様な病態がヒトにおいて見られ、その例に、神経変性疾患、ミオパシー (筋疾患)、および癌がある。

50

【0043】

リンパ球および白血球膜タンパク質

抗原へのB細胞の応答は、正常な免疫系の必須要素である。成熟B細胞は、B細胞受容体（BCR）によってが外来抗原を認識し、これらの受容体は膜結合性の特異的抗体であって、外来抗原と結合する。この抗原／受容体複合体は細胞内に取り込まれ、抗原はタンパク質分解処理を受ける。複雑な抗原に対し有効な応答を発生するには、BCRとBCRに会合するタンパク質とT細胞応答との全てが必要である。抗原のタンパク質分解断片は、B細胞の表面の主要組織適合複合体II（MHC II）分子と複合体を形成する。B細胞の表面でこの複合体をT細胞が認識し得る。一方、マクロファージ他のリンパ様細胞は、MHC I分子に会合する抗原を、T細胞に提示する。T細胞はMHC I-抗原複合体を認識し、それによって活性化されるが、この作用は、T細胞受容体／CD3複合体との相互作用による。後者の複合体はT細胞表面多量体タンパク質で、原形質膜にある。抗原提示によって活性化したT細胞は多様なリンホカインを分泌する。リンホカインはB細胞の成熟とT細胞の増殖とを誘導し、また標的細胞を殺すマクロファージを活性化する。

【0044】

白血球は炎症応答と免疫応答とにおいて基本的な役割を持ち、白血球の例としては単球／マクロファージ、肥満細胞、多形核白血球、ナチュラルキラー細胞、好中球、好酸球、好塩基球、および骨髄系前駆体がある。白血球膜タンパク質の例に、CD抗原のメンバー、N-CAM、I-CAM、ヒト白血球抗原（HLA）クラスIおよびHLAクラスIIの遺伝子産物、免疫グロブリン、免疫グロブリン受容体、補体、補体受容体、インターフェロン、インターフェロン受容体、インターロイキン受容体、およびケモカイン受容体がある。

【0045】

異常なリンパ球と白血球との活性に伴う急性障害の例には、AIDS、免疫過敏、白血病、白血球減少症、全身性ループス、肉芽腫症、および好酸球増加症がある。

【0046】

アポトーシス関連膜タンパク質

多様なリガンド、受容体、酵素、腫瘍抑制因子、ウイルス性遺伝子産物、薬物、および無機イオンは、アポトーシスの細胞破壊の調節と実行において、重要な正または負の役割を持つ。アポトーシス経路の幾つかの特定成分はすでに同定され特徴付けられているが、関与するタンパク質間の多くの相互作用は不明確であり、この経路の主要な諸相は未知のままである。

【0047】

DNA切断と、Fasが媒介する細胞死とにおけるカルシウムレベルの関与を示す過去の研究は、アポトーシスにおいてカルシウムが必要であることを示唆した（Hewish, D. R. および L. A. Burgoyne (1973) Biochem. Biophys. Res. Comm. 52:504-510; Vignaux, F. 他 (1995) J. Exp. Med. 181:781-786; Oshimi, Y. および S. Miyazaki (1995) J. Immunol. 154:599-609)。別の研究は、細胞内カルシウム濃度の上昇が、胸腺細胞においてT細胞受容体架橋または糖質コルチコイドによってアポトーシスが引き起こされた際に見られ、また、この上昇をブロックすることによって細胞死を防げることを示している（McConkey, D. J. 他 (1989) J. Immunol. 143:1801-1806; McConkey, D. J. 他 (1989) Arch. Biochem. Biophys. 269:365-370)。したがって、膜タンパク質、例えばカルシウムチャネルおよびFas受容体は、アポトーシス応答にとって重要である。

【0048】

新規の膜貫通タンパク質、およびそれらをコードするポリヌクレオチド群の発見により、新規の組成物群を提供することで、当分野の要望に応えることができる。これら新規の組成物は、生殖障害、発達障害、心血管障害、神経障害、胃腸障害、脂質代謝障害、細胞増

殖異常、および自己免疫／炎症疾患の、診断・治療・予防において有用であり、また、膜貫通タンパク質の核酸配列およびアミノ酸配列の発現における、外来性化合物群の効果についての算定にも有用である。

【0049】

(発明の概要)

本発明は、総称して「TMP」、個別にはそれぞれ「TMP-1」、「TMP-2」、「TMP-3」、「TMP-4」、「TMP-5」、「TMP-6」、「TMP-7」、「TMP-8」、「TMP-9」、「TMP-10」、「TMP-11」、「TMP-12」、「TMP-13」、「TMP-14」、「TMP-15」、「TMP-16」、および「TMP-17」と呼ぶ膜貫通タンパク質である精製されたポリペプチドを提供する。或る実施例において本発明は、(a) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を持つ天然アミノ酸配列を有するポリペプチド、(c) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した、単離されたポリペプチドを提供する。一実施例で本発明は、SEQ ID NO: 1-17 のアミノ酸配列を持つ、単離されたポリペプチドを提供する。

10

【0050】

また、本発明は(a) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を持つ或る天然アミノ酸配列を有するポリペプチド、(c) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリヌクレオチドの免疫原性断片、からなる群から選択した或るポリペプチドをコードする、単離されたポリヌクレオチドを提供する。一実施態様では、該ポリヌクレオチドは、SEQ ID NO: 1-17 からなる群から選択した或るポリペプチドをコードする。別の実施態様では、ポリヌクレオチドはSEQ ID NO: 18-34 からなる群から選択される。

20

30

【0051】

本発明は更に、(a) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドをコードするポリヌクレオチドと機能的に連結したプロモーター配列を持つ組換えポリヌクレオチドを提供する。一実施態様で本発明は、この組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様で本発明は、この組換えポリヌクレオチドを持つ遺伝形質転換体を提供する。

40

【0052】

更に本発明は、(a) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリペプチドと、(b) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列との少なくとも90%以上の同一性を有する或る天然アミノ酸配列を持つポリペプチドと、(c) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリペプチドの生物学的活性断片と、(d) SEQ ID NO: 1-17 からなる群から選択したアミノ酸配列を持つポリペプチドの免疫原性断片、とからなる群から選択したポリペプチドを製造する方法を提供する。製造方法は、(a) 或る細胞を該ポリペプチドの発現に適した条件下で培養し、この細胞を組換えポリヌクレオチドを用

50

いて形質転換する過程と、(b) そのように発現したポリペプチドを回収する過程とからなる。この組換えポリヌクレオチドは、該ポリペプチドをコードするポリヌクレオチドに対し機能的に連結したプロモーター配列を持つ。

【0053】

本発明は更に、(a) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO: 1-17 からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドに特異結合する、単離された抗体を提供する。

10

【0054】

本発明は更に、(a) SEQ ID NO: 18-34 からなる群から選択したポリヌクレオチド配列を持つポリヌクレオチド、(b) SEQ ID NO: 18-34 からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有する或る天然ポリヌクレオチド配列を持つポリヌクレオチド、(c) (a) に相補的なポリヌクレオチド、(d) (b) に相補的なポリヌクレオチド、および(e) (a) ~ (d) のRNA等価物、からなる群から選択した、単離されたポリヌクレオチドを提供する。一実施態様で該ポリヌクレオチドは、少なくとも60の連続したヌクレオチド群からなる。

【0055】

本発明は更に、サンプル中の標的ポリヌクレオチドを検出する一方法を提供する。ここで、標的ポリヌクレオチドは、(a) SEQ ID NO: 18-34 からなる群から選択した或るポリヌクレオチド配列を持つポリヌクレオチド、(b) SEQ ID NO: 18-34 からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有する天然ポリヌクレオチド配列を持つポリヌクレオチド、(c) (a) に相補的なポリヌクレオチド、(d) (b) に相補的なポリヌクレオチド、および(e) (a) ~ (d) のRNA等価物、からなる群から選択した、或るポリヌクレオチドの配列を持つ。検出方法は、(a) サンプル中の上記標的ポリヌクレオチドに相補的な或る配列からなる少なくとも20の連続したヌクレオチド群からなる或るプローブを用いて該サンプルをハイブリダイズする過程と、(b) 該ハイブリダイゼーション複合体の有無を検出し、複合体が存在すればオプションでその量を検出する過程とからなる。該プローブと該標的ポリヌクレオチドあるいはその断片との間でハイブリダイゼーション複合体が形成されるような条件下で、プローブは、該標的ポリヌクレオチドに対し特異的にハイブリダイズする。一実施態様では、プローブは少なくとも60の連続したヌクレオチド群からなる。

20

30

【0056】

本発明はまた、サンプル中の標的ポリヌクレオチドを検出する一方法を提供する。ここで、標的ポリヌクレオチドは、(a) SEQ ID NO: 18-34 からなる群から選択したポリヌクレオチド配列を持つポリヌクレオチド、(b) SEQ ID NO: 18-34 からなる群から選択したポリヌクレオチド配列との少なくとも90%の同一性を有する天然ポリヌクレオチド配列を持つポリヌクレオチド、(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド、(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド、および(e) (a) ~ (d) のRNA等価物、からなる群から選択した、或るポリヌクレオチドの配列を持つ。検出方法は、(a) ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、(b) 増幅した標的ポリヌクレオチドまたはその断片の有無を検出し、該標的ポリヌクレオチドまたはその断片が存在すればオプションでその量を検出する過程とからなる。

40

【0057】

本発明は更に、或る有効量のポリペプチドと薬剤として許容し得る或る賦形剤とからなる、或る組成物を提供する。有効量のポリペプチドは、(a) SEQ ID NO: 1-17 からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO

50

：1-17からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO：1-17からなる群から選択した或るアミノ酸配列を含むポリペプチドの生物学的活性断片、および(d) SEQ ID NO：1-17からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択される。一実施態様では、前記組成物はSEQ ID NO：1-17からなる群から選択したアミノ酸配列を持つ。更に本発明は、機能性TMPの発現低下に関連する疾患又は病状を治療する方法であって、そのような治療を必要とする患者に対して前記組成物を投与する過程を有する方法を提供する。

【0058】

本発明はまた、(a) SEQ ID NO：1-17からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO：1-17からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO：1-17からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO：1-17からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択したポリペプチドのアゴニストとしての有効性を確認するために、或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドを有するサンプルを或る化合物に曝す過程と、(b) サンプル中のアゴニスト活性を検出する過程とからなる。別法では、本発明は、この方法で同定したアゴニスト化合物と許容される医薬用賦形剤とを有する、或る組成物を提供する。更なる別法では、本発明は、機能的TMPの発現の低下に関連した疾患やその症状の治療を必要とする患者への、この組成物の投与を含む方法を提供する。

【0059】

本発明は更に、(a) SEQ ID NO：1-17からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO：1-17からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO：1-17からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO：1-17からなる群から選択したアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択したポリペプチドのアнтаゴニストとしての有効性のために、或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドを含むサンプルを或る化合物に曝す過程と、(b) サンプル中のアンタゴニスト活性を検出する過程とからなる。別の一実施態様で本発明は、この方法で同定したアンタゴニスト化合物と薬剤として許容し得る賦形剤とを有する組成物を提供する。更なる別法で本発明は、機能的TMPの過剰な発現に関連した疾患やその症状の治療を必要とする患者への、この組成物の投与を含む方法を提供する。

【0060】

本発明は更に、(a) SEQ ID NO：1-17からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO：1-17からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO：1-17からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、または(d) SEQ ID NO：1-17からなる群から選択したアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドに特異結合する或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドを適切な条件下で少なくとも1つの試験化合物と混合する過程と、(b) この試験化合物と該ポリペプチドとの結合を検出し、それにより該ポリペプチドに特異結合する化合物を同定する過程とからなる。

【0061】

本発明は更に、(a) SEQ ID NO：1-17からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO：1-17からなる群から選択したア

ミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO: 1-17からなる群から選択した或るアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-17からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドの活性をモジュレートする或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドの活性にとり許容し得る条件下で、該ポリペプチドを少なくとも1つの試験化合物と混合する過程と、(b) 該ポリペプチドの活性をこの試験化合物の存在下で算定する過程と、(c) この試験化合物の存在下での該ポリペプチドの活性をこの試験化合物の不存在下での該ポリペプチドの活性と比較する過程とからなり、この試験化合物の存在下での該ポリペプチドの活性の変化は、該ポリペプチドの活性をモジュレートする化合物を標示する。 10

【0062】

更に本発明は、SEQ ID NO: 18-34からなる群から選択した或るポリヌクレオチド配列を持つ標的ポリヌクレオチドの発現を改変する効果につき、或る化合物をスクリーニングする一方法を提供する。この方法は、(a) この標的ポリヌクレオチドを有するサンプルを化合物に曝露する過程と、(b) この標的ポリヌクレオチドの発現の改変を検出する過程とからなる。

【0063】

本発明は更に、試験化合物の毒性の算定方法を提供する。この方法には、以下の過程がある。(a) 核酸群を有する生物学的サンプルを試験化合物で処理する過程。(b) 処理済み生物学的サンプルの核酸群をハイブリダイズする過程。この過程には、次のようなプローブを用いる。(i) SEQ ID NO: 18-34からなる群から選択した或るポリヌクレオチド配列を持つポリヌクレオチド、(ii) SEQ ID NO: 18-34からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有するポリヌクレオチド配列を持つ天然ポリヌクレオチド配列を持つポリヌクレオチド、(iii) (i)に相補的な配列を持つポリヌクレオチド、(iv) (ii)のポリヌクレオチドに相補的なポリヌクレオチド、(v) (i)~(iv)のRNA等価物、からなる群から選択した或るポリヌクレオチドの少なくとも20の連続したヌクレオチド群からなるプローブである。ハイブリダイゼーションは、上記プローブと生物学的サンプル中の標的ポリヌクレオチドとの間に特定のハイブリダイゼーション複合体が形成されるような条件下で生じる。上記標的ポリヌクレオチドは、(i) SEQ ID NO: 18-34からなる群から選択した或るポリヌクレオチド配列を持つポリヌクレオチド、(ii) SEQ ID NO: 18-34からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有する天然ポリヌクレオチド配列を持つポリヌクレオチド、(iii) (i)のポリヌクレオチドに相補的なポリヌクレオチド、(iv) (ii)のポリヌクレオチドに相補的なポリヌクレオチド、(v) (i)~(iv)のRNA等価物、からなる群から選択する。あるいは、標的ポリヌクレオチドは、上記(i)~(v)からなる群から選択したポリヌクレオチド配列の断片を持つ。毒性の算定方法には更に、以下の過程がある。 30

(c) ハイブリダイゼーション複合体の量を定量する過程と、(d) 処理した生物学的サンプルのハイブリダイゼーション複合体の量を、非処理の生物学的サンプルのハイブリダイゼーション複合体の量と比較する過程とである。処理した生物学的サンプル内のハイブリダイゼーション複合体の量の差異が、試験化合物の毒性を示す。 40

【0064】

(本発明の記載について)

本発明のタンパク質、ヌクレオチド配列および方法について説明するが、その前に、説明した特定の装置、材料および方法に本発明が限定されるものではなく、修正され得ることを理解されたい。また、ここで使用する専門用語は特定の実施態様を説明する目的で用いたものに過ぎず、特許請求の範囲にのみ限定される本発明の範囲を限定することを意図したものではないことも併せて理解されたい。

【0065】

請求の範囲および明細書中で用いている単数形の「或る」および「その（この）」の表記は、文脈から明らかにそうでないとされる場合を除いて複数のものを指す場合もあることに注意されたい。したがって、例えば「或る宿主細胞」と記されている場合にはそのような宿主細胞が複数あることもあり、「或る抗体」と記されている場合には単数または複数の抗体、および、当業者に公知の抗体の等価物などについても言及している。

【0066】

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

10

【0067】

（定義）

用語「TMP」は、天然、合成、半合成或いは組換え体などの、また全ての種（特にウシ、ヒツジ、ブタ、ネズミ、ウマおよ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TMPのアミノ酸配列を指す。

【0068】

用語「アゴニスト」は、TMPの生物学的活性を強化または模倣する分子を指す。アゴニストの例には、TMPと直接に相互作用し、あるいはTMPが関与する生物学的経路の成分に作用してTMPの活性をモジュレートする、タンパク質、核酸、糖質、小分子、または任意の他の化合物や組成物があり得る。

20

【0069】

「対立遺伝子変異体／変異配列」は、TMPをコードする遺伝子の、別の形態である。対立遺伝子変異配列群は、核酸配列における少なくとも1つの突然変異から生じることがあり、また、変容したmRNA群を生じることがあり、または、ポリペプチド群を生じることがあり、それらのポリペプチドの構造または機能は、変容することもしないこともある。或る遺伝子は、その天然型の対立遺伝子変異配列を全く持たない場合もあり、1個以上持つこともある。対立遺伝子変異配列を生じさせる通常の突然変異性変化は一般に、ヌクレオチドの自然な欠失、付加または置換による。これら各種の変化は、単独あるいは他の変化と共に、或る配列内で1回以上、生じ得る。

30

【0070】

TMPをコードする「変容した／改変された」核酸配列に含む配列は、様々なヌクレオチドの欠失、挿入、或いは置換を生じた配列であって、TMPと同一のポリペプチドを、或いは、TMPの機能特性の少なくとも1つを備えるポリペプチドを生じる配列である。この定義には、TMPをコードするポリヌクレオチド配列にとって正常な染色体の遺伝子座ではない位置での、対立遺伝子変異配列との不適當あるいは予期しないハイブリダイゼーションを含み、また、TMPをコードするポリヌクレオチドの或る特定オリゴヌクレオチドプローブを用いて容易に検出可能なあるいは検出困難な多型性を含む。コードされるタンパク質も「変容する／改変される」ことがあり、また、サイレント変化を生じた結果、機能的に等価なTMPとなるような、アミノ酸残基の欠失、挿入または置換を持ち得る。意図的なアミノ酸置換は、生物学的或いは免疫学的にTMP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、および／または両親媒性についての類似性に基づいて成され得る。例えば、負に帯電したアミノ酸にはアスパラギン酸およびグルタミン酸があり、正に帯電したアミノ酸にはリジンおよびアルギニンがある。親水性値が近似した非荷電極性側鎖を持つアミノ酸には、アスパラギンとグルタミン、セリンとトレオニンがある。親水性値が近似した非荷電側鎖を持つアミノ酸には、ロイシンとイソロイシンとバリン、グリシンとアラニン、フェニルアラニンとチロシンがある。

40

50

【0071】

用語「アミノ酸」および「アミノ酸配列」は、オリゴペプチド、ペプチド、ポリペプチド、タンパク質の配列、あるいはそれらの任意の断片を指し、天然分子または合成分子を指す。「アミノ酸配列」が或る天然タンパク質分子の配列を指す場合、「アミノ酸配列」および類似の用語は、記載したその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。

【0072】

用語「増幅」は、或る核酸配列の付加的複製物を作製する行為に関する。増幅には通常、当業者に公知のポリメラーゼ連鎖反応（PCR）技術を用いる。

【0073】

用語「アンタゴニスト」は、TMPの生物学的活性を阻害或いは減弱する分子を指す。アンタゴニストの例には、TMPと直接相互作用するか、或いはTMPが関与する生物学的経路の成分に作用してTMPの活性を調節する、抗体などタンパク質、核酸、糖質、小分子、任意の他の化合物や組成物を含み得る。

【0074】

用語「抗体」は、エピトープの決定基と結合できる、無傷の免疫グロブリン分子やそれらの断片、例えばFab、F(ab')₂およびFv断片を指す。TMPポリペプチドと結合する抗体は、免疫化抗原として、無傷のポリペプチドを用いて、または、目的の小ペプチドを持つ断片を用いて作製可能である。マウス、ラット、ウサギなどの動物を免疫化するために用いるポリペプチドまたはオリゴペプチドの由来は、RNAの翻訳の場合や、または化学合成があり得る。また、それらは望むならばキャリアタンパク質に抱合し得る。通常用いられるキャリアであってペプチドと化学結合するキャリアの例は、ウシ血清アルブミン、サイログロブリン、および、スカシガイのヘモシアニン（KLH）などがある。結合したこのペプチドは、前記動物を免疫化するために用いる。

【0075】

用語「抗原決定基」は、特定の抗体と接触する、分子の領域（すなわちエピトープ）を指す。タンパク質またはタンパク質断片を用いて宿主動物を免疫化する場合、タンパク質の多数の領域が、抗原決定基（タンパク質の特定の領域または3次元構造）に特異結合する抗体の産生を誘発し得る。或る抗原決定基は、或る抗体への結合について、無傷抗原（すなわ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。

【0076】

用語「アプタマー（aptamer）」は、核酸またはオリゴヌクレオチド分子であって、特定の分子ターゲットに結合する分子を指す。アプタマーは*in vitro*での進化プロセスに由来する（例えば、SELEX（Systematic Evolution of Ligands by EXponential Enrichmentの略、試験管内選択法）、米国特許第5,270,163号に記述）。これは、大規模な組み合わせライブラリ群から標的特異的アプタマー配列を選択するプロセスである。アプタマーの構成は二本鎖または一本鎖であり、デオキシリボヌクレオチド、リボヌクレオチド、ヌクレオチド誘導体または他のヌクレオチド様分子を含み得る。アプタマーの各ヌクレオチド成分は修飾された糖基を持つことがあり（例えば、リボヌクレオチドの2'-OH基は2'-Fまたは2'-NH₂によって置換され得る）、これは、ヌクレアーゼへの抵抗性または血中でのより長い寿命など、所望の特性を向上し得る。アプタマーを他の分子、例えば高分子量のキャリアと抱合させることにより、循環系からのこのアプタマーのクリアランスを遅らせ得る。アプタマー類は、例えば或る架橋剤の光活性化によって、それらの同種リガンド群と特異的に架橋させ得る（例えば、Brody, E. N. およびL. Gold（2000）J. Biotechnol. 74:5-13を参照）。

【0077】

用語「イントラマー（intramer）」は、*in vivo*で発現されるアプタマーを指す。例えば、ワクシニアウイルスに基づく或るRNA発現系を用いて、白血球の細胞質内で特定のRNAアプタマー類が高レベルに発現されている（Blind, M. 他（

10

20

30

40

50

1999) Proc. Natl Acad. Sci. USA 96:3606-3610)。

【0078】

用語「スピーゲルマー (spiegelmer)」は、アプタマーであって、L-DNA、L-RNAその他の左旋性ヌクレオチド誘導体または左旋性ヌクレオチド様分子を含むものを指す。左旋性のヌクレオチド群を持つアプタマー類は、通常は右旋性ヌクレオチド群を持つ基質に作用する、天然酵素類による分解に対して耐性がある。

【0079】

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

10

20

【0080】

用語「生物学的に活性」は、或る天然分子の、構造的、調節的、あるいは生化学的な機能を有するタンパク質を指す。同様に、用語「免疫学的に活性」または「免疫原性」は、天然或いは組換え体のTMP、合成TMP、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内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。

【0081】

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、「5'-AGT-3'」は、その相補配列「3'-TCA-5'」との対を形成する。

30

【0082】

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

40

【0083】

「コンセンサス配列」は、不要な塩基を除くためにDNA配列解析を繰り返し行い、XL-PCRキット (Applied Biosystems, Foster City CA) を用いて5'および/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を指し、または、GELVIEW断片構築システム (GCG, Madison, WI) またはPhrap (University of Washington, Seattle WA) などの断片構築用コンピュータプログラムを用いて1つ以上の重複する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。伸長および構築の両方を行ってコンセンサス配列を産生する配列もある。

【0084】

50

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と予測される置換を指す。すなわち、そのタンパク質の構造、特にその機能が保存され、その置換によって大きくは変わらない。下表は、タンパク質内で元のアミノ酸と置換され得るアミノ酸であり、保存的なアミノ酸置換と認められるアミノ酸を示す。

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

10

20

保存的なアミノ酸置換では通常、(a) 置換領域におけるポリペプチドのバックボーン構造、例えばβシートやα螺旋構造、(b) 置換部位における分子の電荷または疎水性、および/または(c) 側鎖の大部分、を保持する。

【0085】

用語「欠失」は、結果的に1個以上のアミノ酸残基が欠如するアミノ酸配列内の変化、あるいは1個以上のヌクレオチドが欠如するヌクレオチド配列内の変化を指す。

30

【0086】

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。例えば、アルキル基、アシル基、ヒドロキシル基またはアミノ基による水素の置換は、ポリヌクレオチドの化学修飾に含まれ得る。或るポリヌクレオチド誘導体は、天然分子の生物学的または免疫学的機能を少なくとも1つは保持している或るポリペプチドをコードする。ポリペプチド誘導体は、次のようなプロセスによって修飾されたポリペプチドである。すなわち、グリコシル化、ポリエチレングリコール化(pegylation)、あるいは任意の同様なプロセスであって、誘導元のポリペプチドの少なくとも1つの生物学的若しくは免疫学的機能を保持するプロセスによる。

【0087】

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

40

【0088】

「示差発現」は、少なくとも2つの異なったサンプルを比較して判定する、増加(上方調節)、あるいは減少(下方調節)、または欠損した、遺伝子またはタンパク質の発現を指す。このような比較は例えば、処理済サンプルと不処理サンプル、または病態サンプルと健常サンプルとの間で行われ得る。

【0089】

「エキソンシャッフリング」は、異なるコード領域(エキソン)群の組換えを意味する。或るエキソンは、コードされるタンパク質の1つの構造的または機能的ドメインを代表し

50

得るため、安定したサブストラクチャー群の新たな組み合わせによって新しいタンパク質群を構築することが可能であり、新しいタンパク質機能の進化を促進できる。

【0090】

用語「断片」は、TMPまたはTMP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) と配列が同一であるが親配列より長さが短い部分を指す。断片は、定義された配列の全長から1ヌクレオチド／アミノ酸残基を差し引いた長さよりも短い長さを有し得る。例えば或る断片は、5～1000の連続したヌクレオチドまたはアミノ酸残基を有し得る。プローブ、プライマー、抗原、治療用分子として、あるいはその他の目的のために用いられる或る断片は、少なくとも長さ5、10、15、20、25、30、40、50、60、75、100、150、250若しくは500の、連続したヌクレオチドあるいはアミノ酸残基であり得る。断片は、或る分子の特定領域から優先的に選択し得る。例えば、或るポリペプチド断片は、或る定義された配列内に見られるようなポリペプチドの最初の250または500アミノ酸（または最初の25%または50%）から選択した、或る長さの連続したアミノ酸を持ち得る。これらの長さは明らかに例として挙げているものであり、本発明の実施態様では、配列表、表および図面を含む明細書が支持する任意の長さであり得る。

10

【0091】

SEQ ID NO: 18-34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: 18-34を特異的に同定する固有のポリヌクレオチド配列の領域を持つ。SEQ ID NO: 18-34の或る断片は、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: 18-34に関連ポリヌクレオチド配列から区別する類似の方法に有用である。SEQ ID NO: 18-34の或る断片の正確な長さは、また、その断片に対応するSEQ ID NO: 18-34の領域は、その断片に意図した目的に基づき当業者が慣例的に決定し得る。

20

【0092】

SEQ ID NO: 1-17の或る断片は、SEQ ID NO: 18-34の或る断片によってコードされる。SEQ ID NO: 1-17の或る断片は、SEQ ID NO: 1-17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: 1-17の或る断片は、SEQ ID NO: 1-17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。SEQ ID NO: 1-17の或る断片の正確な長さは、また、この断片に対応するSEQ ID NO: 1-17の領域は、その断片の目的に基づいて当業者が慣例的に判定し得る。

30

【0093】

「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）と、それに続く1オープンリーディングフレームと翻訳終止コドンとを有する配列である。或る「完全長」ポリヌクレオチド配列は、或る「完全長」ポリペプチド配列をコードする。

【0094】

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

40

【0095】

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

【0096】

ポリヌクレオチド配列間の一致率を判定するには、MEGALIGN version 3.12e配列アラインメントプログラムなどに組込まれているCLUSTAL Vアルゴリズムの、デフォルトのパラメータ群を用い得る。このプログラムは、LASERGEN

50

E ソフトウェアパッケージ（一組の分子生物学的分析プログラム）（DNASTAR, Madison WI）の一部である。このCLUSTAL Vは、Higgins, D. G. および P. M. Sharp (1989) CABIOS 5:151-153、Higgins, D. G. 他 (1992) CABIOS 8:189-191に記載されている。ポリヌクレオチド配列をペアワイズアラインメントする際のデフォルトパラメータは、Ktuple=2、gap penalty=5、window=4、「diagonals saved」=4と設定される。「weighted」残基重み付け表がデフォルトで選択される。CLUSTAL Vによる一致率の報告は、アラインメントされたポリヌクレオチド配列間の「percent similarity（類似性パーセント）」としてなされる。

10

【0097】

あるいは、米国国立バイオテクノロジー情報センター（NCBI）のBasic Local Alignment Search Tool（BLAST）が一般的に用いられ、且つ、無料で入手可能な配列比較アルゴリズム一式を提供している（Altschul, S. F. 他 (1990) J. Mol. Biol. 215:403-410）。このアルゴリズムは、幾つかの情報源から入手可能であり、メリーランド州ベセスダにあるNCBIおよびインターネット（<http://www.ncbi.nlm.nih.gov/BLAST/>）からも入手可能である。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」と呼ばれるツールも利用可能であり、これは2つのヌクレオチド配列を直接のペアワイズで比較するために用いられる。「BLAST 2 Sequences」は、<http://www.ncbi.nlm.nih.gov/gorf/bl2.html>にアクセスして、対話形式で利用できる。「BLAST 2 Sequences」ツールは、blastn および blastp（以下に記載）の両方に用い得る。BLASTプログラムは、一般的には、ギャップ（gap）その他のパラメータをデフォルト設定にセットして用いる。例えば、2つのヌクレオチド配列を比較するには、デフォルトパラメータに設定した「BLAST 2 Sequences」ツールVersion 2.0.12（2000年4月21日）を用いてblastnを実行し得る。デフォルトパラメータの設定例を以下に示す。

20

30

【0098】

Matrix: BLOSUM62
 Reward for match: 1
 Penalty for mismatch: -2
 Open Gap: 5 及び Extension Gap: 2 penalties
 Gap x drop-off: 50
 Expect: 10
 Word Size: 11
 Filter: on

一致率は、ある定義された配列の全長（例えば特定のSEQ IDナンバーで定義された配列）で測定し得る。あるいは、より短い長さ、例えば、定義された、より大きな配列から得られた断片（例えば少なくとも20、30、40、50、70、100または200の連続したヌクレオチドの断片）の長さの一致率も測定し得る。ここに挙げた長さは単なる例示的なものに過ぎず、表、図および配列リストを含めた本明細書に記載された配列が支持する任意の断片長を用いて、一致率を測定し得る或る長さを説明し得ることを理解されたい。

40

【0099】

高度の同一性を示さない核酸配列が、それにもかかわらず遺伝子コードの縮重が原因で、類似のアミノ酸配列をコードする場合がある。この縮重を利用して核酸配列内で変化を生じさせ、全ての核酸配列が実質上同一のタンパク質をコードするような多数の核酸配列を

50

生成し得ることを理解されたい。

【0100】

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

【0101】

ポリペプチド配列間の一致率は、MEGALIGN version 3.12e配列アラインメントプログラムに組込まれているようなCLUSTAL Vアルゴリズムの、デフォルトのパラメータを用いて決定できる（参照先と共に上述）。CLUSTAL Vを用いて、ポリペプチド配列をペアワイズアラインメントする際のデフォルトパラメータは、Ktuple=1、gap penalty=3、window=5、「diagonals saved」=5と設定される。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一致率は、「percent similarity（類似性パーセント）」としてCLUSTAL Vによって報告される。

【0102】

あるいは、NCBI BLASTソフトウェア一式を用い得る。例えば、2つのポリペプチド配列をペアワイズで比較する場合、「BLAST 2 Sequences」ツールVersion 2.0.12（2000年4月21日）のblastpをデフォルトパラメータに設定して用い得る。デフォルトパラメータの設定例を以下に示す。

【0103】

Matrix: BLOSUM62

Open Gap: 11 及び Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

一致率は、ある定義されたポリペプチド配列の全長（例えば特定のSEQ IDナンバーで定義された配列）で測定し得る。あるいは、より短い長さ、例えば、定義された、より大きなポリペプチド配列から得た断片（例えば少なくとも15、20、30、40、50、70、または150の連続した残基の断片）の長さの同一性も測定し得る。ここに挙げた長さは単なる例示的なものに過ぎず、表、図および配列リストを含めた本明細書に記載された配列が支持する任意の断片長を用いて、一致率を測定し得る或る長さを説明し得ることを理解されたい。

【0104】

「ヒト人工染色体（HAC）」は直鎖状の微小染色体であり、約6 kb ～10 MbのサイズのDNA配列を持ち得る。また、染色体の複製、分離および維持に必要な、全てのエレメントを持つ。

【0105】

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を改変した抗体分子を指す。

【0106】

「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、或る一本鎖ポリヌクレオチドが或る相補的な一本鎖と塩基対を形成するアニーリングのプロセスである。特異的ハイブリダイゼーションは、2つの核酸配列が高い相補性を共有することの指標である。特異的ハイブリダイゼーション複合体は許容されるアニーリング条件下で形成され、1回以上の「洗浄」ステップ後もハイブリダイズされたままである。洗浄ステップ

は、ハイブリダイゼーションプロセスのストリンジェンシーを決定する際に特に重要であり、よりストリンジェントな条件では、非特異結合（すなわち完全には一致しない核酸鎖対間の結合）が減少する。核酸配列のアニーリングに関する許容条件は、当業者が慣例的に決定できる。アニーリング条件はどのハイブリダイゼーション実験でも一定であり得るが、洗浄条件は、所望のストリンジェンシーを、したがってハイブリダイゼーション特異性を得るように、実験ごとに変更し得る。許容されるアニーリング条件は、例えば、温度が68℃で、約6×SSC、約1%（w/v）のSDS、並びに約100μg/mlの、せん断して変性したサケ精子DNAの存在下である。

【0107】

一般に、ハイブリダイゼーションのストリンジェンシーは或る程度、洗浄ステップを実行する温度を基準にして表すことができる。このような洗浄温度は通常、所定のイオン強度およびpHにおける特定配列の融点（ T_m ）より約5～20℃低くなるように選択する。この T_m は、所定のイオン強度およびpHの条件下で、完全一致プローブに標的配列の50%がハイブリダイズする温度である。 T_m を計算する式および核酸ハイブリダイゼーション条件はよく知られており、Sambrook, J. 他（1989）Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NYに記載されており、特に2巻の9章を参照されたい。 10

【0108】

本発明のポリヌクレオチド間の高ストリンジェンシー条件のハイブリダイゼーションには、約0.2×SSCおよび約0.1%のSDSの存在下、68℃で1時間の洗浄条件を含む。別法では、約65℃、60℃、55℃、または42℃の温度で行い得る。SSC濃度は、約0.1%のSDS存在下で、約0.1～2×SSCの範囲で変更し得る。通常は、ブロッキング剤を用いて非特異ハイブリダイゼーションを阻止する。このようなブロッキング剤には、例えば、約100～200μg/mlの、せん断して変性サケ精子DNAがある。特定条件下で、例えばRNAとDNAとのハイブリダイゼーションでは、有機溶剤、例えば約35～50% v/vの濃度のホルムアミドを用いることもできる。洗浄条件の有用なバリエーションは、当業者には自明であろう。ハイブリダイゼーションは、特に高ストリンジェント条件下では、ヌクレオチド間の進化的な類似性を示唆し得る。このような類似性は、ヌクレオチド群、およびヌクレオチドがコードするポリペプチド群について、或る同様の役割を強く標示する。 20 30

【0109】

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の形成によって形成された、2つの核酸配列間の複合体を指す。ハイブリダイゼーション複合体は、溶解状態で形成し得る（ C_0t または R_0t 解析など）。あるいは、一方の核酸配列が溶解状態で存在し、もう一方の核酸配列が固体支持体（例えば紙、膜、フィルタ、チップ、ピンまたはガラスライド、あるいは他の適切な基板であって細胞若しくはそれらの核酸が固定される基板）に固定されているような2つの核酸配列間に形成され得る。

【0110】

用語「挿入」あるいは「付加」は、1個以上のアミノ酸残基あるいはヌクレオチドがそれぞれ追加される、アミノ酸配列あるいは核酸配列における変化を指す。 40

【0111】

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

【0112】

用語「免疫原性断片」は、例えば哺乳動物など生物に導入すると免疫応答を引き起こし得る、TMP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」にはまた、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TM 50

Pの任意のポリペプチド断片またはオリゴペプチド断片をも含む。

【0113】

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

【0114】

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。

【0115】

用語「モジュレート」または「活性を調節」は、TMPの活性の変化を指す。例えば、モジュレートによって、TMPのタンパク質活性の、或いは結合特性の、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き得る。 10

【0116】

「核酸」および「核酸配列」の語は、ヌクレオチド、オリゴヌクレオチド、ポリヌクレオチドまたはこれらの断片を指す。「核酸」および「核酸配列」の語はまた、ゲノム起源または合成起源のDNAまたはRNAであって一本鎖または二本鎖であるかあるいはセンス鎖またはアンチセンス鎖を表し得るようなDNAまたはRNAや、ペプチド核酸(PNA)や、任意のDNA様またはRNA様物質を指すこともある。

【0117】

「機能的に連結した」は、第1の核酸配列と第2の核酸配列が機能的な関係にある状態を指す。例えば、或るプロモーターが或るコード配列の転写または発現に影響を及ぼす場合には、そのプロモーターはそのコード配列と機能的に連結している。機能的に連結したDNA配列群は非常に近接するか連続的に隣接することがあり、また、2つのタンパク質コード領域を結合するために必要な場合は同一リーディングフレーム内にあり得る。 20

【0118】

「ペプチド核酸(PNA)」は、末端がリシンで終わるアミノ酸残基のペプチドのバックボーンに結合した、少なくとも約5ヌクレオチドの長さのオリゴヌクレオチドを持つ、アンチセンス分子または抗遺伝子剤を指す。末端のリシンは、この組成物に溶解性を与える。PNAは、相補的一本鎖DNAまたはRNAに優先的に結合して転写の伸長を停止させる。ポリエチレングリコール化することにより、細胞内でのPNAの寿命を延長し得る。

【0119】

TMP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、タンパク質分解性切断およびその他の当分野で既知の修飾を含み得る。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、TMPの酵素環境に依存し、細胞の種類によって異なることとなる。 30

【0120】

「プローブ」とは、同一配列或いは対立遺伝子核酸配列や、関連する核酸配列の検出に用いる、TMPやそれらの相補配列、またはそれらの断片をコードする核酸配列のことである。プローブは、単離されたオリゴヌクレオチドまたはポリヌクレオチドであって、検出可能な標識またはレポーター分子に結合したものである。典型的な標識には、放射性アイソトープ、リガンド、化学発光試薬および酵素がある。「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。プライマーは次に、DNAポリメラーゼ酵素により、標的DNA鎖に沿って伸長され得る。プライマー対は、例えばポリメラーゼ連鎖反応(PCR)による、核酸配列の増幅(および同定)に用い得る。 40

【0121】

本発明に用いるプローブおよびプライマーは通常、既知の配列の、少なくとも15の連続したヌクレオチド群からなる。特異性を高めるため、長めのプローブおよびプライマー、例えば開示した核酸配列の少なくとも20、25、30、40、50、60、70、80、90、100または150の連続したヌクレオチドからなるようなプローブおよびプライマーも用い得る。これよりもかなり長いプローブおよびプライマーもある。表、図面お 50

よび配列リストを含む本明細書が支持する、任意の長さのヌクレオチドを用い得るものと理解されたい。

【0122】

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

10

【0123】

プライマーとして用いるオリゴヌクレオチドの選択は、そのような目的のために本技術分野でよく知られているソフトウェアを用いて行う。例えばOLIGO 4.06ソフトウェアは、各100ヌクレオチドまでのPCRプライマー対の選択に有用であり、また、オリゴヌクレオチドおよび最大5,000までの大きめのポリヌクレオチドであって32キロベースまでのインプットポリヌクレオチド配列から得たものを分析するのにも有用である。類似のプライマー選択プログラムには、拡張能力のための追加機能が組み込まれている。例えば、PrimOUプライマー選択プログラム (テキサス州ダラスにあるテキサス大学南西部医療センターのゲノムセンターから一般向けに入手可能) は、メガベース配列から特定のプライマーを選択することが可能であり、したがってゲノム全体の範囲でプライマーを設計するのに有用である。Primer3プライマー選択プログラム (Whitehead Institute/MIT Center for Genome Research (マサチューセッツ州ケンブリッジ) より入手可能) を用いれば、プライマー結合部位として避けたい配列を指定できる「非プライミングライブラリ (mispriming library)」を入力できる。Primer3は特に、マイクロアレイのためのオリゴヌクレオチドの選択に有用である (後二者のプライマー選択プログラムのソースコードは、それぞれの情報源から得てユーザー固有のニーズを満たすように修正し得る)。PrimerGenプログラム (英国ケンブリッジ市の英国ヒトゲノムマッピングプロジェクトリソースセンターから一般向けに入手可能) は、多数の配列アラインメントに基づいてプライマーを設計することにより、アラインメントされた核酸配列の最大保存領域または最小保存領域のいずれかとハイブリダイズするようなプライマーの選択を可能にする。したがって、このプログラムは、独自のものであれ保存されたものであれ、オリゴヌクレオチド断片およびポリヌクレオチド断片の同定に有用である。上記選択方法のいずれかによって同定したオリゴヌクレオチドおよびポリヌクレオチドの断片は、ハイブリダイゼーション技術において、例えばPCRまたはシーケンシングプライマーとして、マイクロアレイエレメントとして、あるいは核酸のサンプルにおいて完全または部分的相補的ポリヌクレオチドを同定する特異プローブとして有用である。オリゴヌクレオチドの選択方法は、上記の方法に限定されるものではない。

20

30

40

【0124】

「組換え核酸」は、天然の配列ではなく、配列の、2つ以上の離れたセグメントを人工的に組み合わせた配列である。この人為的組合せはしばしば化学合成によって達成するが、より一般的には核酸の単離セグメントの人為的操作によって、例えばSambrookの文献 (前出) に記載されているような遺伝子工学的的手法によって達成する。組換え核酸の語は、単に核酸の一部が付加、置換または欠失により改変された核酸も含む。しばしば組換え核酸には、プロモーター配列に機能的に連結した核酸配列が含まれる。このような組

50

換え核酸を、例えばある細胞を形質転換するために使用されるベクターの一部と成し得る。

【0125】

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

【0126】

「調節エレメント」は、遺伝子の非翻訳領域に通常は由来する核酸配列であり、エンハンサー、プロモーター、イントロンおよび5'および3'の非翻訳領域(UTR)を含む。調節エレメントは、転写、翻訳またはRNA安定性を制御する宿主タンパク質またはウイルスタンパク質と相互作用する。

10

【0127】

「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る、化学的または生化学的な成分である。レポーター分子には、放射性核種、酵素、蛍光剤、化学発光剤、発色剤、基質、補助因子、阻害因子、磁気粒子およびその他の当分野で既知の成分がある。

【0128】

或るDNA配列に対する「RNA等価物」は、基準となるDNA配列と同じ直鎖のヌクレオチド配列からなるが、生じる全ての窒素性塩基のチミンがウラシルに置換され、糖鎖のバックボーンがデオキシリボースではなくリボースからなる。

20

【0129】

用語「サンプル」は、その最も広い意味で用いられている。TMP、TMPをコードする核酸群、またはその断片群を含むと推定される或るサンプルとしては、体液と、細胞や細胞から単離された染色体や細胞内小器官や膜からの抽出物と、細胞と、溶液中に存在するまたは基板に固定されたゲノムDNA、RNA、cDNAと、組織と、組織プリントなどがあり得る。

【0130】

用語「特異結合」および「特異的に結合する」は、タンパク質若しくはペプチドと、アゴニスト、抗体、アンタゴニスト、小分子、若しくは任意の天然若しくは合成の結合組成物との間の相互作用を指す。この相互作用は、タンパク質の特定の構造(例えば抗原決定基すなわちエピトープ)であって結合分子が認識する構造の有無に依存する。例えば、或る抗体がエピトープ「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」と抗体とを含む或る反応の中に、エピトープAを含むポリペプチドが、ある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。

30

【0131】

用語「実質的に精製された」は、自然の環境から採取され、単離あるいは分離された核酸配列ある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%以上除去、最も好ましくは90%以上除去されたものを指す。

【0132】

「置換」とは、1つ以上のアミノ酸残基またはヌクレオチドを、それぞれ別のアミノ酸残基またはヌクレオチドに置き換えることである。

40

【0133】

用語「基板」は、任意の好適な固体あるいは半固体の支持物を指し、膜およびフィルター、チップ、スライド、ウェハ、ファイバー、磁気または非磁気ビーズ、ゲル、チューブ、プレート、ポリマー、微小粒子、毛細管が含まれる。基板は、ウェル、溝、ピン、チャネル、孔など、様々な表面形態を有することができ、基板表面にはポリヌクレオチドやポリペプチドが結合する。

【0134】

「転写イメージ」または「発現プロフィール」は、所定条件下での所定時間における、或

50

る特定の細胞タイプまたは組織による、遺伝子発現の集合的パターンを指す。

【0135】

「形質転換」とは、外来性DNAが、或る受容細胞に導入されるプロセスを言う。形質転換は、本技術分野で知られている種々の方法に従って自然条件または人工条件下で生じ得るものであり、外来性の核酸配列を原核宿主細胞または真核宿主細胞に挿入する、任意の既知の方法を基にし得る。形質転換の方法は、形質転換する宿主細胞の種類によって選択する。限定するものではないが形質転換方法には、ウイルス感染、電気穿孔法（エレクトロポレーション）、熱ショック、リポフェクションおよび微粒子銃を用いる方法がある。「形質転換された細胞」には、導入されたDNAが、自律的に複製するプラスミドとしてある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。更に、限られた期間に一過的に導入DNA若しくは導入RNAを発現する細胞も含まれる。

10

【0136】

本明細書で用いる「遺伝形質転換体」とは任意の生物体であり、限定するものではないが動植物を含み、生物体の1個以上の細胞が、ヒトの介入によって、例えば本技術分野で公知の形質転換技術によって導入された異種核酸を有する生物である。細胞への核酸の導入は、直接または間接的に、細胞の前駆物質に導入することによって行う。これは、計画的な遺伝子操作によって、例えば微量注射法によってあるいは組換えウイルスの導入によって行う。遺伝子操作の語は、古典的な交雑育種あるいは *in vitro* 受精を指すものではなく、組換えDNA分子の導入を指す。本発明に基づいて予期される遺伝形質転換体には、バクテリア、シアノバクテリア、真菌および動植物がある。本発明の単離されたDNAは、本技術分野で知られている方法、例えば感染、形質移入、形質転換またはトランス接合によって宿主に導入することができる。本発明のDNAをこのような生物体に移入する技術は公知であり、前出の Sambrook ら（1989）等の参考文献に記載がある。

20

【0137】

或る特定の核酸配列の「変異体／変異配列」とは、核酸配列1本の或る長さ全体について、該特定核酸配列に対し少なくとも40%の配列同一性を有する核酸配列として決定された配列である。決定には、デフォルトパラメータに設定した「BLAST 2 Sequences」ツール Version 2.0.9（1999年5月7日）を用いて *blastn* を実行する。このような核酸対は、所定の長さに対して、例えば少なくとも50%、60%、70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の配列同一性を示し得る。変異配列は、例えば、「対立遺伝子」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」変異配列と記述され得る。スプライス変異配列は参照分子との顕著な同一性を有し得るが、mRNAプロセッシング中のエキソン群の選択的スプライシングによって通常、より多数またはより少数のポリヌクレオチドを有することになる。対応するポリペプチドは、追加機能ドメイン群を有するか、あるいは参照分子には存在するドメイン群が欠落していることがある。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。結果的に生じるポリペプチドは通常、相互に有意のアミノ酸同一性を持つ。多型性変異配列は、或る種の個体間における、或る特定遺伝子のポリヌクレオチド配列内での変異である。多型変異配列にはまた、ポリヌクレオチド配列の1つのヌクレオチド塩基が異なる「1塩基多型性」（SNP）も含み得る。SNPの存在は、例えば特定の個体群、病状または病状性向を示し得る。

30

40

【0138】

特定のポリペプチド配列の「変異配列」とは、ポリペプチド配列1本の或る長さ全体について、該特定ポリペプチド配列に対し少なくとも40%の配列同一性を有するポリペプチド配列として決定された配列である。決定には、デフォルトパラメータに設定した「BLAST 2 Sequences」ツール Version 2.0.9（1999年5月7日）を用いて *blastp* を実行する。このようなポリペプチド対は、そのポリペプチド

50

の一方の所定の長さに対して、例えば少なくとも50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の配列同一性を示し得る。

【0139】

(発明)

本発明は、新規のヒト膜貫通タンパク質(TMP)群、およびTMPをコードするポリヌクレオチド群の発見に基づき、また、これらの組成物を使用した、細胞増殖異常、自己免疫/炎症疾患、心血管障害、神経障害、および発達障害の診断、治療、ならびに予防の開発に基づく。

【0140】

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

【0141】

表2は、GenBankタンパク質(genpept)データベースに対するBLAST分析で同定した、本発明のポリペプチド群に相同な配列群を示す。列1および列2はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号(ポリペプチドSEQ ID NO:)と、それに対応するIncyteポリペプチド配列番号(IncyteポリペプチドID)とを示す。列3は、最も近いGenBank相同体のGenBank識別番号(GenBank ID NO:)を示す。列4は、各ポリペプチドとその相同体1つ以上との間の一致に関する確率スコアを示す。列5は、該当箇所には適当な引用を示すとともにGenBank相同体1つ以上の注釈(annotation)を示し、これらはすべて本明細書では参考文献に含まれる。

【0142】

表3は、本発明のポリペプチドの多様な構造的特徴を示す。列1および列2はそれぞれ、本発明の各ポリペプチドのポリペプチド配列識別番号(SEQ ID NO:)およびそれに対応するIncyteポリペプチド配列番号(IncyteポリペプチドID)を示す。列3は、各ポリペプチドのアミノ酸残基数を示す。列4および列5はそれぞれ、GCG配列分析ソフトウェアパッケージのMOTIFSプログラム(Genetics Computer Group, Madison WI)によって決定された、リン酸化およびグリコシル化の可能性のある部位を示す。列6は、シグネチャ配列、ドメイン、およびモチーフを含むアミノ酸残基を示す。列7は、タンパク質の構造/機能の分析のための分析方法を示し、該当箇所には更に、分析方法に利用した検索可能なデータベースを示す。

【0143】

表2および表3は共に、本発明の各々のポリペプチドの特性を要約しており、それら特性が、請求の範囲に記載されたポリペプチドが膜貫通タンパク質であることを確立している。例えば、SEQ ID NO: 2は、ラットのプロスタグランジンF2a受容体調節タンパク質(GenBank ID g1054884)との89%の同一性を有することがBasic Local Alignment Search Tool (BLAST)によって示された(表2参照)。BLAST確率スコアは0.0であり、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO: 2はまた、6つの免疫グロブリンドメイン群を有するが、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された(表3参照)。さらに、SEQ ID NO: 2はシグナルペプチドと1つの膜貫通ドメインと、1つのRGDモチーフとを

10

20

30

40

50

持ち、これはSEQ ID NO: 2がヒト膜貫通タンパク質であることを示す更に実証的な証拠を提供する。

【0144】

別の例として、SEQ ID NO: 4は、ヒトのコネキシン31.1 (GenBank ID g4336903) との56%の同一性を有することがBasic Local Alignment Search Tool (BLAST) によって示された(表2参照)。BLAST確率スコアは 5.8×10^{-68} であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO: 4はまた、1つのコネキシンドメインを有するが、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された(表3参照)。BLIMPS、MOTIFS、およびPROFILES CAN解析よりのデータは、SEQ ID NO: 4がコネキシンである、更に実証的な証拠を提供する。注意すべきは、6つの同一のコネキシン群が1つのコネクション(ギャップ結合)を構成し、これは膜貫通チャネルであって原形質膜にあり、化学的また電氣的に機能し、多くの組織内の隣接細胞間で細胞質を結合することである。SEQ ID NO: 5は長さが1554アミノ酸で、1157アミノ酸にわたってヒトMEGF7 (GenBank ID g3449306) との99%の同一性を有することがBasic Local Alignment Search Tool (BLAST) によって示された(表2参照)。BLAST確率スコアは0.0であり、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO: 5はまた、低密度リポタンパク質受容体リピート群と、低密度リポタンパク質受容体ドメイン群とを有するが、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された(表3参照)。BLIMPS分析から得たデータは、SEQ ID NO: 5がタンパク質のLDL受容体ファミリーのメンバーであることを裏づける証拠を更に提供する。

【0145】

別の例において、SEQ ID NO: 6はBasic Local Alignment Search Tool (BLAST) による同定では、マウス低密度リポタンパク質受容体関連タンパク質LRP1B/LRP-DIT (GenBank ID g8926243) との36%の同一性を有する(表2参照)。BLAST確率スコアは 1.5×10^{-40} であり、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO: 6はまた、低密度リポタンパク質受容体ドメイン群を有することが、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して判定された(表3参照)。BLIMPS分析から得たデータは、SEQ ID NO: 6が低密度リポタンパク質受容体関連分子であることを裏づける証拠を更に提供する。更に、SEQ ID NO: 14はヒトTNF誘導性タンパク質CG12-1 (GenBank ID g3978246) との59%の同一性を有することがBasic Local Alignment Search Tool (BLAST) によって示された(表2参照)。BLAST確率スコアは 1.2×10^{-94} であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。HMMER解析よりのデータは、SEQ ID NO: 14が1つの膜貫通ドメインを持つことを示す、更に実証的な証拠を提供する。SEQ ID NO: 1、SEQ ID NO: 3、SEQ ID NO: 7-13、およびSEQ ID NO: 15-17については、同様の方法で分析し、注釈を付けた。SEQ ID NO: 1-17の解析のためのアルゴリズムおよびパラメータは、表7に記述した。

【0146】

表4に示すように、本発明の完全長ポリヌクレオチド配列は、cDNA配列またはゲノムDNA由来のコード(エキソン)配列を用いて、あるいはこれら2種類の配列を任意に組み合わせる構築した。列1および列2はそれぞれ、本発明の各ポリヌクレオチドのポリヌ

10

20

30

40

50

クレオチド配列識別番号（ポリヌクレオチドSEQ ID NO:）およびそれに対応する Incyte ポリヌクレオチドコンセンサス配列番号（Incyte ポリヌクレオチドID）を示す。列3は、各ポリヌクレオチド配列の長さを塩基対で示す。列4は、ポリヌクレオチド配列の断片の内、例えば、SEQ ID NO: 18-34を同定するため、或いはSEQ ID NO: 18-34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断片を示す。列5は、cDNA配列に対応する識別番号、またはゲノムDNAから予想されたコード配列（エキソン）に、および／または、cDNAおよびゲノムDNAを共に有する配列集合に対応する識別番号を示している。これらの配列を用いて、本発明の完全長ポリヌクレオチド配列を構築した。表4の列6および列7はそれぞれ、それらの完全長配列に関連する列5のcDNA配列および／またはゲノム配列の開始ヌクレオチド（5'）位置および終了ヌクレオチド（3'）位置を示す。

10

【0147】

表4の列5の識別番号は、特に例えば Incyte cDNAとそれに対応するcDNAライブラリとを指す場合もある。例えば、6798827J1は Incyte cDNA配列の識別番号であり、COL ENOR03は、その配列が由来するcDNAライブラリの識別番号である。cDNAライブラリが示されていない Incyte cDNA（例えば、71760758V1）は、プールされているcDNAライブラリに由来する。または、列5の識別番号は、完全長ポリヌクレオチド配列の構築に用いた GenBank のcDNAまたはESTの識別番号の場合もある（例えば、g1506355）。更に、列5の識別番号は、ENSEMBL（The Sanger Centre、英国ケンブリッジ）データベースに由来した配列を示す場合もある（すなわち「ENST」の命名を含む配列）。あるいは、列5の識別番号は、NCBI RefSeq Nucleotide Sequence Records データベースに由来する場合もあり（すなわち「NM」または「NT」の命名を含む配列）、またNCBI RefSeq Protein Sequence Recordsに由来する場合もある（すなわち「NP」の命名を含む配列）。または列5の識別番号は、「エキソンスティッチング（exon-stitching）」アルゴリズムにより結び合わせたcDNAとGenscan予想エキソンとの両方からなる群を意味する場合がある。例えば、FL_XXXXXX_N1_N2_YYYYY_N3_N4 は「スティッチされた」配列であり、その内、XXXXXXは該アルゴリズムが適用される配列のクラスターの識別番号であり、YYYYYは該アルゴリズムが生成する予測の数であり、N1, 2, 3... がある場合には、解析中に手動で編集された特定のエキソン群を表す（実施例5参照）。または、列5の識別番号は「エキソンストレッチング（exon-stretching）」アルゴリズムにより結び合わせたエキソンの集合を指す場合もある。例えば、FLXXXXXX_gAAAAA_gBBBBB_1_Nは「ストレッチ」配列の識別番号である。ここでXXXXXXは Incyte プロジェクト識別番号、gAAAAAは「エキソンストレッチング」アルゴリズムを適用したヒトゲノム配列の GenBank 識別番号、gBBBBBBは一番近い GenBank タンパク質相同体の GenBank 識別番号またはNCBI RefSeq 識別番号、Nは特定のエキソンである（実施例5を参照）。或る RefSeq 配列が「エキソンスティッチング」アルゴリズムのためのタンパク質相同体として使用された場合では、「NM」、「NP」、または「NT」によって表される RefSeq 識別子が、GenBank 識別子（すなわち、gBBBBBB）の代わりに使用され得る。

20

30

40

【0148】

あるいは、接頭コードは、手動で編集された構成配列、ゲノムDNA配列から予測された構成配列、または組み合わされた配列解析方法に由来する構成配列を同定する。次の表は、構成配列の接頭コードと、接頭コードに対応する配列分析方法の例とを列記する（実施例4と5を参照）。

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

10

20

30

【0149】

場合によっては、最終コンセンサスポリヌクレオチド配列を確認するために、列5に示すような配列カバレッジと重複する Incyte cDNA カバレッジが得られたが、それに関連する Incyte cDNA 識別番号は示さなかった。

【0150】

表5は、Incyte cDNA 配列を用いて構築された完全長ポリヌクレオチド配列のための代表的な cDNA ライブラリを示している。代表的な cDNA ライブラリとは Incyte cDNA ライブラリであり、これは、最も頻繁には Incyte cDNA 配列群によって代表されるが、これらは、上記のポリヌクレオチド配列を構築および確認するために用いられた。cDNA ライブラリを作製するために用いた組織およびベクターを表5に示し、表6で説明している。

40

【0151】

本発明はまた、TMP の変異配列 (変異体) を含む。好適な TMP の変異体は、TMP アミノ酸配列に対して、少なくとも約80%、あるいは少なくとも約90%、さらには少なくとも約95%のアミノ酸配列同一性を有し、TMP の機能的または構造的特徴を少なくとも1つ含む変異体である。

【0152】

本発明はまた、TMP をコードするポリヌクレオチドを提供する。或る特定の実施例にお

50

いて、本発明は、TMPをコードするSEQ ID NO: 18-34からなる一群から選択された配列を含むポリヌクレオチド配列を提供する。配列表に示したSEQ ID NO: 18-34のポリヌクレオチド配列は、窒素系塩基のチミンがウラシルに置換され、糖鎖のバックボーンがデオキシリボースではなくリボースからなる等価RNA配列を含む。

【0153】

本発明はまた、TMP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、TMPをコードするポリヌクレオチド配列に対して少なくとも約70%のポリヌクレオチド配列同一性、或いは少なくとも約85%のポリヌクレオチド配列同一性、更には少なくとも約95%ものポリヌクレオチド配列同一性を有する。本発明の或る特定の実施態様では、SEQ ID NO: 18-34からなる群から選択されたアミノ酸配列との少なくとも約70%、或いは少なくとも約85%、または少なくとも約95%もの一致率を有するような、SEQ ID NO: 18-34からなる群から選択した或る配列を有するポリヌクレオチド配列の変異配列を含む。上記の任意のポリヌクレオチドの変異体は、TMPの機能的若しくは構造的特徴を少なくとも1つ有するアミノ酸配列をコードし得る。

【0154】

更に別の例では、本発明の或るポリヌクレオチド変異配列は、TMPをコードするポリヌクレオチド配列のスプライス変異配列である。或るスプライス変異配列はTMPをコードするポリヌクレオチド配列との顕著な配列同一性を持つ部分複数を有し得るが、mRNAプロセッシング中のエキソン群の選択的スプライシングによって生ずる、配列の数ブロックの付加または欠失により、通常、より多数またはより少数のポリヌクレオチドを有することになる。或るスプライス変異配列には、約70%未満、または約60%未満、あるいは約50%未満のポリヌクレオチド配列同一性が、TMPをコードするポリヌクレオチド配列との間で全長に渡って見られるが、このスプライス変異配列のいくつかの部分には、TMPをコードするポリヌクレオチド配列の各部との、少なくとも約70%、あるいは少なくとも約85%、または少なくとも約95%、なおまたは100%の、ポリヌクレオチド配列同一性を有することとなる。上記したスプライス変異配列は何れも、TMPの機能的或いは構造的特徴の少なくとも1つを有する或るアミノ酸配列をコードし得る。

【0155】

遺伝暗号の縮重により、TMPをコードする種々のポリヌクレオチド配列が産生され、中には、いかなる既知の天然遺伝子のポリヌクレオチド配列にも最小の類似性しか無いものもあることは、当業者には理解されよう。したがって本発明には、可能コドン選択に基づく組合せの選択によって産出し得るあらゆる可能なポリヌクレオチド配列のバリエーションを網羅し得る。これらの組合せは、天然TMPのポリヌクレオチド配列に適用される標準トリプレット遺伝暗号を基に成されるものであり、このような変異は全て明確に開示されていると考慮されたい。

【0156】

TMPとその変異配列とをコードするヌクレオチド配列は一般に、好適に選択されたストリンジェンシー条件下で天然TMPのヌクレオチド配列とハイブリダイズ可能であるが、非天然コドン群を含めるなどの実質的に異なるコドン使用を有する、TMP或いはその誘導体をコードするヌクレオチド配列を作り出すことは、有益であり得る。宿主が特定のコードンを利用する頻度に基づいて、特定の真核宿主または原核宿主に発生するペプチドの発現率を高めるようにコードンを選択し得る。コードされるアミノ酸配列を改変せずに、TMPおよ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改変する別の理由には、天然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転写物を作ることもある。

【0157】

本発明にはまた、TMPとTMP誘導体とをコードする、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製する過程をも含む。作製後、当分野で公知の試薬類を用いて、この合成配列を任意の多くの入手可能な発現ベクター及び細胞系中に挿入し得る。

更に、合成化学を用いて、TMPまたはその任意の断片をコードする配列に突然変異を誘発し得る。

【0158】

更に本発明には、種々のストリンジェンシー条件下で、請求項に記載されたポリヌクレオチド配列、特にSEQ ID NO: 18-34、およびそれらの断片とのハイブリダイズが可能なポリヌクレオチド配列を含む（例えば、Wahl, G. M. および S. L. Berger (1987) *Methods Enzymol.* 152: 399-407; Kimmel, A. R. (1987) *Methods Enzymol.* 152: 507-511を参照）。アニーリングおよび洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載した。

10

【0159】

DNAシーケンシングの方法は当分野では公知であり、本発明のいずれの実施例も、DNAシーケンシング方法を用い得る。DNAシーケンシング方法には酵素を用い得る。例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amersham, Pharmacia Biotech, Piscataway NJ) を用い得る。あるいは、例えばELONGASE増幅システム (Life Technologies, Gaithersburg MD) において見られるように、ポリメラーゼと校正エキソヌクレアーゼとを併用し得る。好適には、MICROLAB 2200液体転移システム (Hamilton, Reno, NV)、PTC 200サーマルサイクラー (MJ Research, Watertown MA) およびABI CATALYST 800サーマルサイクラー (Applied Biosystems) などの装置を用いて配列の準備を自動化する。次に、ABI 373 あるいは 377 DNAシーケンシングシステム (Applied Biosystems)、MEGABACE 1000 DNAシーケンシングシステム (Molecular Dynamics, Sunnyvale CA) または当分野で公知の他のシステムを用いてシーケンシングを行う。結果として得た配列を、当分野で周知の種々の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, 7.7ユニット、Meyers, R. A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, 856-853ページを参照）。

20

30

【0160】

当分野で周知の、PCR法をベースにした種々の方法と、部分的ヌクレオチド配列とを利用して、TMP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、上流にある配列を検出し得る。例えば、使用し得る方法の1つである制限部位PCR法は、ユニバーサルプライマーおよびネステッドプライマーを用いてクローニングベクター内のゲノムDNAからの未知の配列を増幅する方法である（例えば、Sarkar, G. (1993) *PCR Methods Applic.* 2: 318-322を参照）。別の方法に逆PCR法があり、これは多岐の方向に伸長するプライマー群を用いて、環状化した鋳型から未知の配列を増幅する方法である。鋳型は、或る既知のゲノム遺伝子座およびその周辺の配列群からなる制限酵素断片群から得る（例えば、Triglia, T. 他 (1988) *Nucleic Acids Res.* 16: 8186を参照）。第3の方法としてキャプチャPCR法があり、これにはヒトおよび酵母人工染色体DNAの既知の配列群に隣接するDNA断片群をPCR増幅する方法を含む (Lagerstrom, M. 他 (1991) *PCR Methods Applic.* 1: 111-119などを参照)。この方法では、PCRを行う前に、複数の制限酵素の消化およびライゲーション反応を用い、未知の配列領域内に組換え二本鎖配列を挿入し得る。未知の配列群を検索するために用い得る他の複数の方法も当分野で公知である (Parker, J. D. 他 (

40

50

1991) Nucleic Acids Res. 19:3055-3060などを参照)。更に、PCR、ネステッドプライマー類、およびPROMOTERFINDERライブラリ類(Clontech, Palo Alto CA)を用いてゲノムDNAをウォーキングし得る。この手順は、ライブラリ類をスクリーニングする必要がなく、イントロン/エキソン接合部の発見に有用である。全てのPCRベースの方法で、市販されているソフトウェア、例えばOLIGO 4.06プライマー分析ソフトウェア(National Biosciences, Plymouth MN)或いは別の好適なプログラムを用いて、長さが約22~30ヌクレオチド、GC含有率が約50%以上で、温度約68℃~72℃で鋳型に対してアニーリングするように、プライマー群を設計し得る。

【0161】

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

【0162】

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

【0163】

本発明の別の実施例では、TMP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TMP、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配列を産生し、これらの配列をTMPの発現に利用し得る。

【0164】

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

【0165】

本発明のヌクレオチドを、MOLECULARBREEDING (Maxygen, Santa Clara CA; 米国特許第5,837,458号; Chang, C.-C. 他(1999) Nat. Biotechnol. 17:793-797; Christians, F.C. 他(1999) Nat. Biotechnol. 17:259-264; Cramer, A. 他(1996) Nat. Biotechnol. 14:315-319)などのDNAシャッフリング技術を用いてシャッフリングし、TMP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TMPの生物学的特性を改変或いは改良することができる。DNAシャッフリング

10

20

30

40

50

は、PCRを介する遺伝子断片組換えを用いて遺伝子変異配列のライブラリを作製するプロセスである。ライブラリはその後、所望の特性を持つ遺伝子変異配列群を同定する、選択またはスクリーニングの手順を経る。続いて、これら好適な変異配列をプールし、更に反復してDNAシャッフリングおよび選択／スクリーニングを行い得る。かくして、「人工的な」育種および急速な分子の進化によって、多様な遺伝子が作られる。例えば、ランダムポイント突然変異を持つ単一の遺伝子の断片を組み換え、スクリーニングし、その後所望の特性が最適化されるまでシャッフリングし得る。あるいは、或る遺伝子の断片群を同種または異種のいずれかから得た同一遺伝子ファミリーの相同遺伝子の断片と組み換え、それによって天然に存在する複数の遺伝子の遺伝多様性を、方向付けした制御可能な方法で最大化させることができる。

10

【0166】

別の実施例によれば、TMP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、全体或いは一部を合成可能である（例えば、Caruthers, M. H. 他（1980）*Nucleic Acids Symp. Ser.* 7: 215-223; および Horn, T. 他（1980）*Nucleic Acids Symp. Ser.* 7: 225-232を参照）。別法として、化学的方法を用いてTMP自体またはその断片を合成することが可能である。例えば、種々の液相または固相技術を用いてペプチド合成を行い得る（Creighton, T.（1984）*Proteins, Structures and Molecular Properties*, WH Freeman, New York NY, 55-60ページ、Roberge, J. Y. 他（1995）*Science* 269: 202-204などを参照）。自動合成は、ABI 431Aペプチドシンセサイザ（Applied Biosystems）を用いて達成し得る。更にTMPのアミノ酸配列または任意のその一部は、直接的な合成の際の改変、および／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群との組み合わせにより、天然ポリペプチドの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

【0167】

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

30

【0168】

生物学的に活性なTMPを発現させるために、TMPをコードするヌクレオチド配列またはそれらの誘導体を好適な発現ベクターに挿入することができる。好適な発現ベクターとは、好適な宿主に挿入されたコーディング配列の転写と翻訳との制御に必要なエレメント群を持つベクターである。必要なエレメントの例には、ベクターに、またTMPをコードするポリヌクレオチド配列における調節配列、例えばエンハンサー、構成型および発現誘導型のプロモーター、5'および3'の非翻訳領域などが含まれる。このようなエレメント群は、特長および特異性が様々である。特定の開始シグナル類によって、TMPをコードする配列の、より効果的な翻訳を達成することも可能である。このようなシグナルの例には、ATG開始コドンと、コザック配列など近傍の配列が含まれる。TMPをコードする配列、およびその開始コドンや、上流の調節配列が、好適な発現ベクターに挿入された場合は、更なる転写制御シグナルや翻訳制御シグナルは必要ない場合もある。しかしながら、コーディング配列あるいはその断片のみが挿入される場合は、インフレームのATG開始コドンなど外来性の翻訳制御シグナルを発現ベクターが提示するようにすべきである。外来性の翻訳エレメント群および開始コドン群は、様々な天然物および合成物を起源とし得る。用いる特定の宿主細胞系に適したエンハンサー類を含めることにより、発現の効率を高めることが可能である（Scharf, D. 他（1994）*Results Probl. Cell Differ.* 20: 125-162などを参照）。

40

50

【0169】

当業者に周知の方法を用いて、TMPをコードする配列と、好適な転写および翻訳制御エレメントとを持つ発現ベクターを作製することが可能である。これらの方法としては、in vitro組換えDNA技術、合成技術、およびin vivo遺伝子組換え技術がある（例えば、Sambrook, J. 他. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, 4章および8章, および16-17章; および Ausubel, F. M. 他. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, 9章、13章および16章を参照）。

10

【0170】

種々の発現ベクター／宿主系を利用して、TMPをコードする配列を保持および発現し得る。限定するものではないがこのような発現ベクター／宿主系には、組換えバクテリオファージ、プラスミド、またはコスミドDNA発現ベクターで形質転換させた細菌や、酵母菌発現ベクターで形質転換させた酵母菌など微生物や、ウイルス発現ベクター（例えばバキュロウイルス）に感染した昆虫細胞系や、ウイルス発現ベクター（例えばカリフラワーモザイクウイルス、CaMVまたはタバコモザイクウイルス、TMV）または細菌発現ベクター（例えばTiプラスミドまたはpBR322プラスミド）で形質転換させた植物細胞系、動物細胞系がある（前出のSambrook、前出のAusubel、Van Heeke, G. 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) EMBO J. 6:307-311、『マグローヒル科学技術年鑑』（The McGraw Hill Yearbook of Science and Technology）(1992) McGraw Hill New York NY, 191-196ページ、Logan, J. および T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659、Harrington, J. J. 他 (1997) Nat. Genet. 15:345-355などを参照）。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. および N. Somia (1997) Nature 389:239-242などを参照）。本発明は、使用する宿主細胞によって限定されるものではない。

20

30

40

【0171】

細菌系では、TMP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多数のクローニングベクターおよび発現ベクターを選択し得る。例えば、TMPをコードするポリヌクレオチド配列の慣例的なクローニング、サブクローニング、増殖には、PBLUESCRIPT (Stratagene, La Jolla CA) またはPSPORT1プラスミド (Life Technologies) などの多機能の大腸菌ベクターを用いることができる。ベクターの多数のクローニング部位にTMPをコードする配列をライゲーションするとlacZ遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。更にこれらのベクターは、クローニングされた配列におけるin vitro転写、ジデオキシのシーケンシング、ヘルパーファージ

50

による一本鎖のレスキュー、入れ子状態の欠失の生成にも有用であろう（例えば、Van Heeke, G. および S. M. Schuster (1989) J. Biol. Chem. 264:5503-5509 を参照）。例えば、抗体の産生のためなどに多量の TMP が必要な場合は、TMP の発現をハイレベルで誘導するベクターが使用できる。例えば、強力な誘導 SP6 バクテリオファージプロモーターまたは誘導 T7 バクテリオファージプロモーターを含むベクターが使用できる。

【0172】

酵母発現系を用いて TMP を産生し得る。α 因子、アルコールオキシダーゼ、PGH プロモーターなど、構成型あるいは誘導型のプロモーターを持つ多数のベクターが、出芽酵母菌 (*Saccharomyces cerevisiae*) またはピキア酵母 (*Pichia pastoris*) 内で使用可能である。更に、このようなベクターは、発現したタンパク質の分泌か細胞内での保持かのどちらかを指示し、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ことを可能にする（例えば、Ausubel, 1995, 前出、Bitter, G. A. 他 (1987) Methods Enzymol. 153:516-544、および Scorer, C. A. 他 (1994) Bio/Technology 12:181-184 を参照）。

【0173】

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

【0174】

哺乳動物細胞においては、多数のウイルスベースの発現系を利用し得る。アデノウイルスを発現ベクターとして用いる場合、後期プロモーターと 3 連リーダー配列とを持つアデノウイルス転写／翻訳複合体に、TMP をコードする配列を結合し得る。アデノウイルスゲノムの非必須 E1 または E3 領域へ挿入することにより、宿主細胞内で TMP を発現する感染ウイルスを得ることが可能である（例えば、Logan, J. および T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659 を参照）。更に、ラウス肉腫ウイルス (RSV) エンハンサー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得る。SV40 または EBV をベースにしたベクターを用いてタンパク質を高レベルで発現させることもできる。

【0175】

また、ヒト人工染色体 (HAC) 類を用いて、或るプラスミドに含まれ得るものやプラスミドから発現し得るものより大きな、DNA の断片を送達し得る。治療のために、約 6 kb ~ 10 Mb の HACs が作製され、従来の送達方法（リポソーム、ポリカチオンアミノポリマー、またはベシクル）で送達されている（例えば、Harrington, J. J. 他 (1997) Nat Genet. 15:345-355 を参照）。

【0176】

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、細胞株における TMP の安定した発現が望ましい。例えば、TMP をコードする配列を細胞株に形質転換するために、発現ベクター類と、同じ或いは別のベクター上の選択可能マーカー遺伝子とを用い得

10

20

30

40

50

る。用いる発現ベクターは、複製のウイルス起源、および／または内在性の発現エレメント群を持ち得る。ベクターの導入後、選択培地に移す前に強化培地で約1～2日間、細胞を増殖させ得る。選択可能マーカーの目的は選択剤への抵抗性を与えることであり、選択可能マーカーの存在により、導入した配列をうまく発現するような細胞の成長および回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に適した組織培養技術を用いて増殖可能である。

【0177】

任意の数の選択系を用いて、形質転換細胞株を回収できる。限定するものではないがこのような選択系には、 tk^{-} 細胞のために用いられる単純ヘルペスウイルスのチミジンキナーゼ遺伝子と、 $ap r^{-}$ 細胞のために用いられる単純ヘルペスウイルスのアデニンホスホリボシルトランスフェラーゼ遺伝子がある (Wigler, M. 他 (1977) Cell 11:223-232、Lowy, I. 他 (1980) Cell 22:817-823などを参照)。また、選択の基礎として、代謝拮抗物質、抗生物質、あるいは除草剤への耐性を用いることができる。例えば $dhfr$ はメトトレキセートに対する耐性を与え、 neo はアミノグリコシドネオマイシンおよびG-418に対する耐性を与え、 als はクロルスルフロンに対する耐性を、 pat はホスフィノトリシンアセチルトランスフェラーゼに対する耐性を各々与える (Wigler, M. 他 (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570、Colbere-Garapin, F. 他 (1981) J. Mol. Biol. 150:1-14などを参照)。この他の選択可能な遺伝子、例えば、代謝のための、細胞の必要条件を改変する $trpB$ および $hisD$ も、文献に記載されている (Hartman, S. C. および R. C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051を参照)。可視マーカー類、例えばアントシアニンや、緑色蛍光タンパク質 (GFP; Clontech)、 β グルクロニダーゼおよびその基質である β グルクロニド、またはルシフェラーゼおよびその基質であるルシフェリンを用い得る。これらのマーカーを用いて、トランスフォーマントを同定するだけでなく、特定のベクター系に起因する一過性あるいは安定したタンパク質発現を定量し得る (Rhodes, C. A. (1995) Methods Mol. Biol. 55:121-131などを参照)。

【0178】

マーカー遺伝子発現の有無によって目的の遺伝子の存在が示唆されても、その遺伝子の存在および発現の確認が必要な場合もある。例えば、TMP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、TMPをコードする配列を含む形質転換された細胞は、マーカー遺伝子機能の欠落により同定可能である。または、1つのプロモーターの制御下で、或るマーカー遺伝子が、TMPをコードする配列とタンデムに配置され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をも標示する。

【0179】

一般に、TMPをコードする核酸配列を含み且つTMPを発現する宿主細胞は、当業者によく知られている種々の方法を用いて同定することが可能である。限定するものではないが当業者によく知られている方法には、DNA-DNA或いはDNA-RNAハイブリダイゼーションと、PCR増幅とがあり、また、核酸配列或いはタンパク質配列の検出、定量、或いはその両方を行うための、膜系、溶液ベース或いはチップベースの技術を含む、タンパク質の生物学的検定法または免疫学的検定法もある。

【0180】

特異的ポリクローナル抗体または特異的モノクローナル抗体を用いてTMPの発現の検出と計測とを行うための免疫学的方法は、当分野で既知である。このような技術の例としては、酵素結合イムノソルベントアッセイ (ELISA)、ラジオイムノアッセイ (RIA)、蛍光活性化細胞選別 (FACS) などが挙げられる。TMP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群を用いた、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、競合する結合アッセイを用いることもできる。これらのアッセイおよびこれ以外のアッセイは、当分野で公知である (Hampton, R. 他 (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Sect. IV、Coligan, J. E. 他 (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY、Pound, J. D. (1998) Immunochemical Protocols, Humana Press, Totowa NJなどを参照)。

10

【0181】

多岐にわたる標識方法および抱合方法が当業者に知られており、様々な核酸アッセイおよびアミノ酸アッセイにこれらの方法を用い得る。TMPをコードするポリヌクレオチドに関連する配列を検出するための、標識されたハイブリダイゼーションプローブ或いはPCRプローブを生成する方法には、オリゴ標識化、ニックトランスレーション、エンドラベリング (末端標識化)、または、標識したヌクレオチドを用いるPCR増幅が含まれる。別法として、TMPをコードする配列、またはその任意の断片を、mRNAプローブを生成するためのベクターにクローニングすることも可能である。このようなベクターは、当分野において知られており、市販もされており、T7、T3またはSP6などの好適なRNAポリメラーゼおよび標識されたヌクレオチドを加えて、in vitroでRNAプローブの合成に用いることができる。このような方法は、例えばAmersham Pharmacia Biotech、Promega (Madison WI)、およびU. S. Biochemicalなどから市販されている種々のキットを用いて実行することができる。検出を容易にするために用い得る好適なレポーター分子あるいは標識には、基質、補助因子、インヒビター、磁気粒子のほか、放射性核種、酵素、蛍光剤、化学発光剤、発色剤などがある。

20

【0182】

TMP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および回収に好適な条件下で培養される。形質転換細胞から製造されたタンパク質が分泌されるか細胞内に留まるかは、使用される配列、ベクター、あるいはその両者に依存する。TMPをコードするポリヌクレオチドを持つ発現ベクターは、原核細胞膜または真核細胞膜を通してのTMPの分泌を誘導するシグナル配列群を持つように設計できることは、当業者には理解されよう。

30

【0183】

更に、宿主細胞株の選択は、挿入した配列の発現をモジュレートする能力、または発現したタンパク質を所望の形に処理する能力に関して行い得る。限定するものではないがこのようなポリペプチドの修飾には、アセチル化、カルボキシル化、グリコシル化、リン酸化、脂質化およびアシル化がある。タンパク質の「プレプロ」または「プロ」形を切断する翻訳後のプロセッシングを利用して、タンパク質ターゲティング、折りたたみ、および/または活性を特定することが可能である。翻訳後の活性のための、特定の細胞装置および特徴的な機序を持つ、種々の宿主細胞 (例えばCHO、HeLa、MDCK、MEK293、WI38など) は、American Type Culture Collection (ATCC, Manassas, VA) から入手可能であり、外来タンパク質の正しい修飾および処理を確実にするように選択し得る。

40

【0184】

本発明の別の実施例では、TMPをコードする、自然のあるいは修飾された、または組換えの核酸配列を、或る異種配列に結合させることができ、この結果、上記した任意の宿主系の、或る融合タンパク質の翻訳を生じる。例えば、市販されている抗体を用いて認識可能な異種部分を含むキメラTMPタンパク質は、TMP活性阻害剤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および異種ペプチド部

50

分も、市販されている親和性マトリックスを用いて融合タンパク質の精製を促進し得る。限定されるものではないがこのような部分には、グルタチオンSトランスフェラーゼ (GST)、マルトース結合タンパク質 (MBP)、チオレドキシン (Trx)、カルモジュリン結合ペプチド (CBP)、6-His、FLAG、c-myc、赤血球凝集素 (HA) がある。GSTは固定化グルタチオン上で、MBPはマルトース上で、Trxはフェニルアルシンオキシド上で、CBPはカルモジュリン上で、そして6-Hisは金属キレート樹脂上で、同族の融合タンパク質の精製を可能にする。FLAG、c-mycおよび赤血球凝集素 (HA) は、これらのエピトープ標識を特異的に認識する市販のモノクローナル抗体およ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精製を可能にする。また、TMP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、TMPが、精製の後に異種部分から切断され得る。融合タンパク質の発現および精製方法は、前出のAusubel (1995) 10章に記載されている。市販されている種々のキットを用いて融合タンパク質の発現および精製を促進することもできる。

【0185】

本発明の別の実施例では、TNTウサギ網状赤血球可溶化液またはコムギ胚芽抽出系 (Promega) を用いて in vitro で、放射能標識したTMPの合成が可能である。これらの系は、T7、T3またはSP6プロモーターと機能的に連結したタンパク質コード配列の、転写と翻訳とを結合する。翻訳は、例えば³⁵Sメチオニンのような放射能標識したアミノ酸前駆体の存在下で起こる。

【0186】

本発明のTMPまたはその断片を用いて、TMP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、TMP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。

【0187】

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TMPの天然リガンド、または天然の基質、構造的または機能的な擬態物質、または自然結合パートナーに密接に関連している (Coligan, J. E. 他 (1991) Current Protocols in Immunology 1 (2) の5章等を参照)。同様に、この化合物は、TMPが結合する天然受容体に、或いは例えばリガンド結合部位など少なくともこの受容体の或る断片に、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TMPを発現する好適な細胞の作製が含まれる。好適な細胞には、哺乳動物、酵母、ショウジョウバエ、または大腸菌からの細胞が含まれる。TMPを発現する細胞またはTMPを含有する細胞膜分画を試験化合物と接触させて、TMPまたはこの化合物のどちらかの結合、刺激、または活性の阻害を分析する。

【0188】

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

【0189】

本発明のTMPまたはその断片を用いて、TMPの活性をモジュレート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などが含まれる。一実施例では、TMPが少なくとも1つの試験化合物と結合する、TMP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TMPの活性を、試験化合物不在下でのTMPの活性と比較する。試験化合物の存在下でのTMPの活性の変化は、TMPの活性をモジュレートする化合物の存在を示唆する。別法では、試験化合物を、TMPの活性に適した条件下で、TMPを有する *in vitro* 系すなわち無細胞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においても、TMPの活性をモジュレートする試験化合物は間接的にモジュレートする場合があります、その際は試験化合物と直接接触する必要がない。少なくとも1つ、または複数の試験化合物をスクリーニングできる。

【0190】

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、TMP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号および第5,767,337号などを参照）。例えば129/SvJ細胞株などのマウスES細胞は初期のマウス胚に由来し、培地で増殖させることができる。このES細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（neo; Capecchi, M. R. (1989) *Science* 244:1288-1292）などのマーカー遺伝子で破壊した、目的の遺伝子を持つ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マウス株などから採取したマウス細胞胚盤胞に微量注入する。これらの胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を特定し、これらを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、潜在的治療薬や毒性薬剤で検査することができる。

10

20

30

40

50

【0191】

TMP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も可能である。ヒトES細胞は、内胚葉、中胚葉および外胚葉の細胞タイプを含む、少なくとも8つの別々の細胞系統に分化する可能性を有する。これらの細胞系統は、例えば神経細胞、造血系統および心筋細胞に分化する（Thomson, J. A. 他 (1998) *Science* 282:1145-1147）。

【0192】

TMP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）を作製することも可能である。ノックイン技術を用いて、TMP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。形質転換した細胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について試験し、潜在的医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばTMPを乳汁内に分泌するなど、TMP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）。

【0193】

（治療）

TMPの各領域と膜貫通タンパク質類との間には、例えば配列およびモチーフの文脈における、化学的および構造的類似性が存在する。さらに、TMPの発現は、脳、前立腺、平滑筋、心血管、下垂体、胃腸、肺、膵臓、および小腸の各組織に密接に関連している。従って、TMPは、生殖障害、発達障害、心血管障害、神経障害、胃腸障害、脂質代謝障害

、細胞増殖異常、および自己免疫／炎症性の疾患においてある役割を果たすと考えられる。T M Pの発現または活性の増大に関連する疾患の治療においては、T M Pの発現または活性を低下させることが望ましい。また、T M Pの発現または活性の低下に関連する疾患の治療においては、T M Pの発現または活性を増大させることが望ましい。

【0194】

したがって、一実施例において、T M Pの発現または活性の低下に関連した疾患の治療または予防のために、患者にT M Pまたはその断片や誘導体を投与することが可能である。限定するものではないがこのような疾患の例としては、生殖障害として、プロラクチン産生の障害、不妊（例えば卵管疾患、排卵不良、子宮内膜症、性周期の途絶、月経周期の途絶、多嚢胞性卵巣症候群、卵巣過剰刺激症候群、子宮内膜腫瘍や卵巣腫瘍、子宮筋腫、自己免疫障害、異所性妊娠、奇形発生、乳癌、乳房線維嚢胞病、乳漏症、精子形成の途絶、異常精子生理、精巣癌、前立腺癌、良性前立腺肥大、前立腺炎、ペーロニー病、インポテンス、男性乳癌、女性化乳房、高ゴナドトロピン性腺機能低下症、低ゴナドトロピン性腺機能低下症、仮性半陰陽、無精子症、早発卵巣不全、アクロシン欠損症、晩発思春期、逆行性射精、無射精、血管芽腫、嚢胞・クロム親和細胞腫（c y s t s p h a e o c h r o m o c y t o m a s）、傍神経節腫、精巣上体の嚢胞腺腫、および内リンパ嚢腫瘍）が、発達障害として、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／ベッカー筋ジストロフィ、癲癇、性腺发育障害、W A G R症候群（ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神遅滞）、スミス・マジェニス症候群、骨髄異形成症候群、遺伝性粘液上皮異形成、遺伝性角皮症、遺伝性ニューロパシー（例えばシャルコー・マリー・トゥース病および神経線維腫症）、甲状腺機能低下症、水頭症、発作障害（例えばシデナム舞蹈病と脳性麻痺）二分脊椎、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感音難聴が含まれる他、心血管障害も含まれ、その中には動静脈瘻、アテローム性動脈硬化症、高血圧、脈管炎、レイノー病、動脈瘤、動脈解離、静脈瘤、血栓静脈炎および静脈血栓、血管系腫瘍、血栓溶解の合併症、バルーン血管形成術、血管置換術、冠動脈バイパス移植、うっ血性心不全、虚血性心疾患、狭心症、心筋梗塞、高血圧性心疾患、変性弁膜性心疾患、石灰化大動脈弁狭窄症、先天性2尖大動脈弁、僧帽弁輪部石灰化、僧帽弁逸脱、リウマチ熱、リウマチ性心疾患、感染性心内膜炎、非細菌性血栓性心内膜炎、全身性紅斑性狼瘡の心内膜炎、カルチノイド心疾患、心筋症、心筋炎、心膜炎、腫瘍性心疾患、先天性心臓疾患、心臓移植の合併症、先天性肺異常、肺拡張不全、肺うっ血および肺水腫、肺動脈塞栓症、肺出血、肺梗塞、肺高血圧症、血管硬化症、閉塞性肺疾患、拘束性肺疾患、慢性閉塞性肺疾患、肺気腫、慢性気管支炎、気管支喘息、気管支拡張症、細菌性肺炎、ウイルス性肺炎およびマイコプラズマ肺炎、肺膿瘍、肺結核、びまん性間質性疾患、塵肺症、サルコイド症、特発性肺線維症、剥離性間質性肺炎、過敏性肺炎、肺好酸球増加、閉塞性細気管支炎—器質化肺炎（b r o n c h i o l i t i s o b l i t e r a n s - o r g a n i z i n g p n e u m o n i a）、びまん性肺出血症候群、グッドパスチャー症候群、特発性肺血鉄症、膠原血管病併発性肺疾患、肺胞たんぱく症、肺腫瘍、炎症性および非炎症性胸水、気胸症、胸膜腫瘍、薬物性肺疾患、放射線による肺疾患、および肺移植の合併症などが含まれる。神経障害として、癲癇、虚血性脳血管障害、脳卒中、脳腫瘍、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病その他の錐体外路障害、筋萎縮性側索硬化その他の運動ニューロン障害、進行性神経性筋萎縮症、網膜色素変性症（色素性網膜炎）、遺伝性運動失調、多発性硬化症その他の脱髄疾患、細菌性およびウイルス性髄膜炎、脳膿瘍、硬膜下膿瘍、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎および神経根炎、ウイルス性中枢神経系疾患と、クールー、クロイツフェルト－ヤコブ病およびゲルストマン－シュトロイスラー－シャインカー症候群を含むプリオン病と、致死性家族性不眠症、神経系の栄養病および代謝病、神経線維腫症、結節硬化症、小脳網膜性血管腫症（c e r e b e l l o r e t i n a l h e m a n g i o b l a s t o m a t o s i s）、脳3叉神経血管症候群、精神遅滞、ダウン症候群を含むその他の中枢神経系発達障害、脳性麻痺、神経骨格異常、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィーその他の神経筋疾患、末梢神経疾患、皮膚筋炎および多発性筋炎

と、遺伝性、代謝性、内分泌性および中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分障害、不安障害および統合失調症（精神分裂病）を含む精神障害と、季節性感情障害（SAD）と、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、遅発性ジスキネジア、ジストニー、パラノイド精神病、帯状疱疹後神経痛、トゥーレット病、進行性核上麻痺、皮質基底核変性症、家族性前頭側頭痴呆が含まれ、また胃腸障害も含まれ、その中には嚥下障害、消化性食道炎、食道痙攣、食道狭窄、食道癌、消化不良、消化障害、胃炎、胃癌、食欲不振、悪心、嘔吐、胃不全麻痺、洞または幽門の浮腫、腹部アンギナ、胸焼け、胃腸炎、イレウス、腸管感染、消化性潰瘍、胆石症、胆嚢炎、胆汁うっ滞、膵臓炎、膵臓癌、胆道疾患、肝炎、高ビリルビン血症、硬変症、肝臓の受動的うっ血、ヘパトーム、感染性大腸炎、潰瘍性大腸炎、潰瘍性直腸炎、クローン病、ホウィップル病、マロリー-ワイス症候群、結腸癌、結腸閉塞、過敏性腸症候群、短小腸症候群、下痢、便秘、胃腸出血、後天性免疫不全症候群（AIDS）腸症、黄疸、肝性脳症、肝腎症候群、脂肪肝、血色素症、ウィルソン病、 α_1 アンチトリプシン欠損症、ライ症候群、原発性硬化性胆管炎、肝梗塞、門脈循環閉塞および血栓、小葉中心壊死、肝臓紫斑病、肝静脈血栓、肝静脈閉塞症、子癇前症、子癇、妊娠性急性脂肪肝、妊娠性肝臓内胆汁うっ滞と、結節性過形成および腺腫、癌腫を含む肝癌が含まれ、脂質代謝障害として、脂肪肝、胆汁うっ滞、原発性胆汁性肝硬変、カルニチン欠乏症、カルニチンパルミトイルトランスフェラーゼ欠乏症、ミオアデニル酸デアミナーゼ欠損症、高トリグリセリド血症、ファブリー病などの脂質貯蔵病、ゴーシェ病、ニーマン-ピック病、異染性白質ジストロフィー、副腎白質ジストロフィー、GM₂ ガングリオシド蓄積症、セロイドリポフスチン症、無 β -リポ蛋白血症、タンジアー病、高リポ蛋白血症、糖尿病、脂肪異栄養症、脂肪腫症、急性皮下脂肪組織炎、播種性脂肪壊死、有痛脂肪症、リポイド副腎過形成、微小変化型ネフローゼ、脂肪腫、アテローム性動脈硬化症、高コレステロール血症、高トリグリセリド血症を伴った高コレステロール血症、原発性低 α リポ蛋白血症、甲状腺機能低下症、腎疾患、肝疾患、レシチン：コレステロールアシルトランスフェラーゼ欠損症、脳髄黄色腫症、シトステロール血症、低コレステロール血症、テイ-サックス病、サンドホフ病、高脂血症、高脂質血症、脂質筋障害、肥満症が、細胞増殖障害として、光線性角化症、動脈硬化症、アテローム性硬化症、滑液包炎、硬変、肝炎、混合性結合組織病（MCTD）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症と、腺癌、白血病、リンパ腫、黒色腫、骨髄腫、肉腫、奇形癌を含む癌、具体的には副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌等が含まれ、また自己免疫／炎症疾患も含まれ、その中には後天性免疫不全症候群（AIDS）、アジソン病、成人呼吸窮迫症候群、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー（APECED）、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球傷害因子性偶発性リンパ球減少症（episodic lymphopenia with lymphocytotoxicity）、新生児溶血性疾患（胎児赤芽球症）、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性腸症候群、多発性硬化症、重症筋無力症、心筋または心膜の炎症、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性紅斑性狼瘡、全身性強皮症（systemic sclerosis）、血小板減少性紫斑病、潰瘍性大腸炎、ブドウ膜炎、ウェルナー症候群、癌の合併症、血液透析、体外循環、ウイルス性感染症、細菌性感染症、真菌性感染症、寄生虫感染症、原虫感染症、蠕虫の感染症および外傷が含まれる。

【0195】

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TMP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、TMP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。

【0196】

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

【0197】

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TMP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、TMP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。

【0198】

更なる実施例では、TMP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TMP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、上記した生殖障害、発達障害、心血管障害、神経障害、胃腸障害、脂質代謝障害、細胞増殖異常、および自己免疫／炎症性の疾患が含まれる。一実施態様では、TMP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TMP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機構或いは送達機構として間接的に用いられ得る。 10

【0199】

別の実施例では、TMP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して、限定するものではないが上記した疾患を含む、TMP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も可能である。 20

【0200】

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

【0201】

TMPのアンタゴニストは、本技術分野で一般的に知られている方法を用いて製造し得る。具体的には、精製されたTMPを用いて抗体を作るか、治療薬のライブラリをスクリーニングして、TMPと特異結合するものを同定することが可能である。TMPへの抗体も、本技術分野で公知の方法を用いて製造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、一本鎖抗体、Fab断片、およびFab発現ライブラリによって作られた断片が含まれ得る。中和抗体（すなわち二量体の形成を阻害する抗体）は通常、治療用に好適である。 30

【0202】

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

【0203】

TMPに対する抗体を誘導するために用いるオリゴペプチド、ペプチドまたは断片は、少なくとも約5のアミノ酸からなり、一般的には少なくとも約10のアミノ酸からなるアミノ酸配列を有するものが好ましい。これらのオリゴペプチド、ペプチドまたは断片は、天然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。TMPアミノ酸の短い 50

ストレッチは、K L Hなどの別のタンパク質の配列と融合し、キメラ分子に対する抗体が産生され得る。

【0204】

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

【0205】

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、好適な抗原特異性およ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. USA 81: 6851-6855、Neuberger, M. S. 他(1984) Nature 312: 604-608、Takeda, S. 他(1985) Nature 314: 452-454などを参照)。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、TMP特異性一本鎖抗体を生成する。関連した特異性を有するがイディオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリからチェーンシャッフリングによって産生することもできる(例えばBurton D. R. (1991) Proc. Natl. Acad. Sci. USA 88: 10134-10137を参照)。

【0206】

抗体類の産生は、リンパ球集団における*in vivo*産生の誘導によって、あるいは免疫グロブリンライブラリ群のスクリーニングによって、または文献に開示されているような、高度に特異的な結合試薬のパネル群のスクリーニングによっても行い得る(例えばOrlandi, R. 他(1989) Proc. Natl. Acad. Sci. USA 86: 3833-3837、Winter, G. 他(1991) Nature 349: 293-299を参照)。

【0207】

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

【0208】

種々の免疫学的検定(イムノアッセイ)を用いてスクリーニングし、所望の特異性を有する抗体を同定することができる。確立された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる免疫放射線アッセイまたは競合結合アッセイに対する数々のプロトコルは、本技術分野において公知である。このようなイムノアッセイは通常、TMPとその特異抗体との間の複合体形成の計測を含む。2つの非干渉性TMPエピトープに対して反応性を持つ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)。

【0209】

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、TMPに対する抗体の親和性を評価し得る。親和性を結合定数K_aで表すが、このK_aは、

平衡状態下でのTMP-抗体複合体のモル濃度を遊離抗体と遊離抗原とのモル濃度で除して得られる値である。ポリクローナル抗体類は多様なTMPエピトープに対する親和性が不均一であり、或るポリクローナル抗体試薬に関して判定した K_a は、TMP抗体群の平均親和性または結合活性を表す。或る特定TMPエピトープに単一特異的なモノクローナル抗体類の、試薬の K_a は、親和性の真の測定値を表す。 K_a 値が $10^9 \sim 10^{12} \text{ L/mol}$ の高親和性抗体試薬は、TMP-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 $10^6 \sim 10^7 \text{ liter/mol}$ の低親和性抗体試薬は、TMP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) およ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. および A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY)。

【0210】

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

【0211】

本発明の別の実施例では、TMPをコードするポリヌクレオチド、またはその任意の断片や相補配列を、治療目的で使うことができる。ある実施態様では、TMPをコードする遺伝子のコーディング領域や調節領域に相補的な配列やアンチセンス分子 (DNA、RNA、PNA、または修飾したオリゴヌクレオチド) を設計して遺伝子発現を変更することができる。このような技術は当分野では周知であり、アンチセンスオリゴヌクレオチドまたはより大きな断片を、TMPをコードする配列の制御領域、またはコード領域に沿った、さまざまな位置から設計可能である (例えばAgrawal, S., 編集 (1996) Antisense Therapeutics, Humana Press, Totawa NJを参照)。

【0212】

治療に用いる場合、アンチセンス配列を適切な標的細胞に導入するのに好適な、任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を作製する発現プラスミドの形で、細胞内に送達し得る (例えばSlater, J. E. 他 (1998) J. Allergy Clin. Immunol. 102 (3): 469-475 および Scanlon, K. J. 他 (1995) 9 (13): 1288-1296を参照)。アンチセンス配列はまた、例えばレトロウイルスやアデノ随伴ウイルスなどのウイルスベクターを用いて細胞内に導入することもできる (例えばMiller, A. D. (1990) Blood 76: 271、前出のAusubel、Uckert, W. および 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、リポソーム系、人工的なウイルスエンベロープおよび当分野で公知のその他のシステムが含まれる (Rossi, J. J. (1995) Br. Med. Bull. 51 (1): 217-225; Boado, R. J. 他 (1998) J. Pharm. Sci. 87 (11): 1308-1315、Morris, M. C. 他 (1997) Nucleic Acids Res. 25 (14): 2730-2736などを参照)。

【0213】

本発明の別の実施例では、TMPをコードするポリヌクレオチドを、体細胞若しくは生殖細胞遺伝子治療に用いることが可能である。遺伝子治療を行うことにより、(i) 遺伝子欠損症（例えばX染色体連鎖遺伝（Cavazzana-Calvo, M. 他（2000）*Science* 288:669-672）により特徴付けられる重症複合型免疫不全症（SCID）-X1の場合）、先天性アデノシンデアミナーゼ（ADA）欠損症に関連する重症複合型免疫不全症候群（Blaese, R. M. 他（1995）*Science* 270:475-480、Bordignon, C. 他（1995）*Science* 270:470-475）、囊胞性繊維症（Zabner, J. 他（1993）*Cell* 75:207-216；Crystal, R. G. 他（1995）*Hum. Gene Therapy* 6:643-666、Crystal, R. G. 他（1995）*Hum. Gene Therapy* 6:667-703）、サラセミア（thalassaemia）、家族性高コレステロール血症、第VIII因子若しくは第IX因子欠損に起因する血友病（Crystal, R. G.（1995）*Science* 270:404-410、Verma, I. M. および Somia, N.（1997）*Nature* 389:239-242）を治療し、(ii) 条件的致死性遺伝子産物を発現させ（例えば制御不能な細胞増殖に起因する癌の場合）、(iii) 細胞内の寄生生物（例えばヒト免疫不全ウイルス（HIV）（Baltimore, D.（1988）*Nature* 335:395-396、Poeschla, E. 他（1996）*Proc. Natl. Acad. Sci. USA* 93:11395-11399）などヒトレトロウイルス、B型若しくはC型肝炎ウイルス（HBV、HCV）、Candida albicans および Paracoccidioides brasiliensis 等の寄生真菌、並びに熱帯熱マラリア原虫およびクルーズトリパノソーマ等の寄生原虫に対する防御機能を有するタンパク質を発現させることができる。TMP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TMP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。

【0214】

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

【0215】

TMPの発現に影響を及ぼし得る発現ベクターには、限定するものではないが、PCDNA 3.1、EPITAG、PRCCMV2、PREP、PVAX、PCR2-TOPO TA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。TMPを発現させるために、(i) 構成的に活性なプロモーター（その由来は例えば、サイトメガロウイルス（CMV）、ラウス肉腫ウイルス（RSV）、SV40ウイルス、チミジンキナーゼ（TK）、若しくはβ-アクチンの遺伝子等）、(ii) 誘導性プロモーター（例えば、市販のT-REXプラスミド（Invitrogen）に含まれるテトラサイクリン調節性プロモーター（Gossen, M. および H. Bujard（1992）*Proc. Natl. Acad. Sci. U. S. A.* 89:5547-555

1 ; G o s s e n , M . 他 (1 9 9 5) S c i e n c e 2 6 8 : 1 7 6 6 - 1 7 6 9 ; R o s s i , F . M . V . および H . M . B l a u (1 9 9 8) C u r r . O p i n . B i o t e c h n o l . 9 : 4 5 1 - 4 5 6)) 、エクジソン誘導性プロモーター (市販のプラスミド P V G R X R および P I N D に含まれる ; I n v i t r o g e n) 、 F K 5 0 6 / ラパマイシン誘導性プロモーター、または R U 4 8 6 / ミフェプリストン誘導性プロモーター (R o s s i , F . M . V . および H . M . B l a u , 前出) 、または (i i i) 正常な個体に由来する、TMP をコードする内在性遺伝子の天然プロモーター、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。

【0216】

市販のリボソーム形質転換キット (例えば I n v i t r o g e n 社の P e r F e c t L i p i d T r a n s f e c t i o n K i t) を用いれば、当業者は実験の各パラメータを最適化する努力をさほど要さず、ポリヌクレオチド群を、培養中の標的細胞群に送達し得る。別法では、リン酸カルシウム法 (G r a h a m , F . L . および A . J . E b (1 9 7 3) V i r o l o g y 5 2 : 4 5 6 - 4 6 7) 若しくは電気穿孔法 (N e u m a n n , B . 他 (1 9 8 2) E M B O J . 1 : 8 4 1 - 8 4 5) を用いて形質転換を行う。初代培養細胞に DNA を導入するためには、標準化された哺乳動物の形質移入プロトコルの修正が必要である。

【0217】

本発明の別の実施例では、TMP の発現に関連する遺伝子欠損によって起こる疾患や障害を、(i) レトロウイルス長末端反復配列 (L T R) プロモーター若しくは或る独立プロモーターのコントロール下で TMP をコードするポリヌクレオチドと、(i i) 好適な RNA パッケージングシグナル群と、(i i i) 追加のレトロウイルス・シス作用性 RNA 配列と、効率的なベクター増殖に必要なコーディング配列とを伴う R e v 応答性エレメント (R R E) 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 P F B および P F B N E O) は S t r a t a g e n e 社から市販されており、刊行データ (R i v i e r e , I . 他 (1 9 9 5) P r o c . N a t l . A c a d . S c i . U S A 9 2 : 6 7 3 3 - 6 7 3 7) に基づく。上記データを引用することを以て本明細書の一部とする。このベクターは、好適なベクター産生細胞株 (V P C L) において増殖される。V P C L は、各標的細胞上の受容体への親和性を持つエンベロップ遺伝子を、または V S V g など汎親和性エンベロップタンパク質を発現する (A r m e n t a n o , D . 他 (1 9 8 7) J . V i r o l . 6 1 : 1 6 4 7 - 1 6 5 0 、 B e n d e r , M . A . 他 (1 9 8 7) J . V i r o l . 6 1 : 1 6 3 9 - 1 6 4 6 、 A d a m , M . A . および A . D . M i l l e r (1 9 8 8) J . V i r o l . 6 2 : 3 8 0 2 - 3 8 0 6 、 D u l l , T . 他 (1 9 9 8) J . V i r o l . 7 2 : 8 4 6 3 - 8 4 7 1 、 Z u f f e r e y , R . 他 (1 9 9 8) J . V i r o l . 7 2 : 9 8 7 3 - 9 8 8 0) 。 R i g g に付与された米国特許第 5 , 9 1 0 , 4 3 4 号 (「M e t h o d f o r o b t a i n i n g r e t r o v i r u s p a c k a g i n g c e l l l i n e s p r o d u c i n g h i g h t r a n s d u c i n g e f f i c i e n c y r e t r o v i r a l s u p e r n a t a n t」) において、レトロウイルスパッケージング細胞株群を得る方法が開示されており、引用を以て本明細書の一部とする。レトロウイルスベクター類の増殖や、細胞集団 (例えば C D 4 ⁺ T 細胞群) の形質導入、および形質導入した細胞群の患者への戻しは、遺伝子治療の分野では当業者に周知の手法であり、多数の文献に記載されている (R a n g a , U . 他 (1 9 9 7) J . V i r o l . 7 1 : 7 0 2 0 - 7 0 2 9 、 B a u e r , G . 他 (1 9 9 7) B l o o d 8 9 : 2 2 5 9 - 2 2 6 7 、 B o n y h a d i , M . L . (1 9 9 7) J . V i r o l . 7 1 : 4 7 0 7 - 4 7 1 6 、 R a n g a , U . 他 (1 9 9 8) P r o c . N a t l . A c a d . S c i . U . S . A . 9 5 : 1 2 0 1 - 1 2 0 6 、 S u , L . (1 9 9 7) B l o o d 8 9 : 2 2 8 3 - 2 2 9 0) 。

【0218】

別法では、アデノウイルス系遺伝子治療の或る送達系を用いて、TMP の発現に関連する

1つ以上の遺伝子異常を有する細胞群に、TMPをコードするポリヌクレオチド群を送達する。アデノウイルス系ベクター類の作製およびパッケージングについては、当業者に周知である。複製不全型アデノウイルスベクター類は、種々の免疫調節タンパク質をコードする遺伝子群を、無損傷の臍島内にインポートする目的で多様に利用し得ることが証明された(Csete, M. E. 他 (1995) Transplantation 27: 263-268)。潜在的に有用なアデノウイルスベクター類はArmentanoに付与された米国特許第5,707,618号(「Adenovirus vectors for gene therapy」)に記載されており、引用することを以て本明細書の一部とする。アデノウイルスベクター類については、Antinozzi, P. A. 他 (1999) Annu. Rev. Nutr. 19: 511-544 並びに、Verma, I. M. および N. Somia (1997) Nature 18: 389: 239-242も参照されたい。両文献は、引用を以て本明細書の一部とする。

【0219】

別法では、ヘルペス系遺伝子治療の送達系を用いて、TMPの発現に関連する1つ或いは複数の遺伝子異常を有する標的細胞に、TMPをコードするポリヌクレオチドを送達する。単純ヘルペスウイルス(HSV)系のベクターは、HSV親和性の中枢神経細胞にTMPを導入する際に特に有用である。ヘルペス系ベクター類の作製およびパッケージングは、当業者に公知である。或る複製適格性単純ヘルペスウイルス(HSV)I型系のベクターが、或るレポーター遺伝子の、霊長類の眼への送達に用いられている(Liu, X. 他 (1999) Exp. Eye Res. 169: 385-395)。HSV-1ウイルスベクターの作製についても、DeLucaに付与された米国特許第5,804,413号(「Herpes simplex virus strains for gene transfer」)に詳細に開示されており、該特許の引用を以て本明細書の一部とする。米国特許第5,804,413号は、ヒト遺伝子治療などの目的で、好適なプロモーターの制御下において細胞に導入される少なくとも1つの外来遺伝子を有するゲノムを持つ、組換えHSVd92の利用について記載する。上記特許はまた、ICP4、ICP27、およびICP22を欠失した組換えHSV系統の、作製および使用について開示している。HSVベクター類については、Goins, W. F. 他 (1999) J. Virol. 73: 519-532 および Xu, H. 他 (1994) Dev. Biol. 163: 152-161も参照されたい。両文献は、引用を以て本明細書の一部とする。クローン化したヘルペスウイルス配列群の操作、大ヘルペスウイルスゲノム(large herpesvirus genomes)の様々なセグメントを持つ複数のプラスミドを形質移入した後の組換えウイルスの産生、ヘルペスウイルスの成長および増殖、並びに細胞群へのヘルペスウイルスの感染は、当業者に公知の技術である。

【0220】

別法では、或るαウイルス(正の一本鎖RNAウイルス)ベクターを用いて、TMPをコードするポリヌクレオチド群を標的細胞群に送達する。プロトタイプのαウイルスであるセムリキ森林熱ウイルス(SFV)の生物学的研究が広範に行われており、遺伝子導入ベクター類は、SFVゲノムに基づく(Garoff, H. および K. -J. Li (1998) Curr. Opin. Biotechnol. 9: 464-469)。αウイルスRNAの複製中に、ウイルスのキャプシッドタンパク質群を通常はコードする、サブゲノムRNAが産生される。このサブゲノムRNAは、完全長ゲノムRNAよりも高レベルに複製するので、酵素活性(例えばプロテアーゼおよびポリメラーゼ)を持つウイルスタンパク質に比べてキャプシッドタンパク質群が過剰産生される。同様に、TMPをコードする配列をαウイルスゲノムのキャプシッ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TMPをコードするRNAが産生され、高いレベルでTMPが合成される。通常、αウイルスの感染は、数日以内の細胞溶解を伴う。一方、シンドビスウイルス(SIN)の或る変異体を有するハムスター正常腎臓細胞(BHK-21)群が持続的な感染を確立する能力は、αウイルス類の溶解複製を、遺伝子治療に応用し得るように好適に改変可能であることを示唆する(Dryga, S. A. 他 (

1997) *Virology* 228:74-83)。様々な宿主にαウイルスを導入できることから、様々なタイプの細胞にTMPを導入することできる。或る集団における或るサブセットの細胞群の特異的形質導入は、形質導入前に細胞の選別を必要とし得る。αウイルスの感染性cDNAクローンの処置方法、αウイルスのcDNAおよびRNAの形質移入方法およびαウイルスの感染方法は、当業者に公知である。

【0221】

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

10

【0222】

リボザイム類は、酵素性RNA分子である。RNAの特異的切断を触媒するためにリボザイムを用い得る。リボザイム作用のメカニズムは、相補的標的RNAへのリボザイム分子の配列特異性ハイブリダイゼーションと、その後起こる内ヌクレオチド鎖切断に参与している。例えば、人工的に作製されたハンマーヘッド型リボザイム分子が、TMPをコードする配列の、内ヌクレオチド鎖分解性の切断を特異的且つ効果的に触媒できる可能性がある。

20

【0223】

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

30

【0224】

本発明の相補リボ核酸分子およびリボザイムは、核酸分子合成のために当分野でよく知られている任意の方法を用いて作製し得る。作製方法には、固相ホスホラミダイト化学合成など、オリゴヌクレオチドを化学的に合成する方法がある。或いは、TMPをコードするDNA配列の*in vitro*および*in vivo*転写によってRNA分子を産出し得る。このようなDNA配列は、T7やSP6などの好適なRNAポリメラーゼプロモーターを用いて多様なベクター内に組み込むことが可能である。あるいは、相補的RNAを構成的あるいは誘導的に合成するこれらのcDNA構成物を、細胞株、細胞または組織内に導入することができる。

40

【0225】

細胞内の安定性を高め、半減期を長くするためにRNA分子を修飾し得る。限定するものではないが可能な修飾には、分子の5'末端、3'末端、あるいはその両方において隣接配列群を追加したり、分子の主鎖内においてホスホジエステラーゼ結合ではなくホスホロチオネートまたは2'-O-メチルを使用したりすることが含まれる。この概念は、本来はPNA群の産出におけるものであるが、これら全ての分子に拡大することができる。そのためには、内在性エンドヌクレアーゼによって容易には認識されない、アデニン、シチジン、グアニン、チミン、およびウリジンにアセチル、メチル、チオールおよび同様の修飾をしたものや、非従来型塩基、例えばイノシン、クエオシン(queosine)、ワイブトシン(wybutosine)を加える。

50

【0226】

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

10

【0227】

或る特定ポリヌクレオチドの発現を改変する際の有効性に対して、少なくとも1個または複数個の試験化合物をスクリーニングし得る。試験化合物は、当分野で広く知られている任意の方法により得られる。このような方法には、ポリヌクレオチドの発現を改変する場合と、既存の、商用または専用の、天然または非天然の化合物のライブラリから選択する場合と、標的ポリヌクレオチドの化学的および／または構造的特性に基づく化合物を合理的にデザインする場合と、組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合と、に有効であることが知られる化合物の化学修飾がある。TMPをコードするポリヌクレオチドを有するサンプルを、このようにして得た試験化合物の少なくとも1つに曝露する。サンプルは例えば、無傷細胞、透過化処理した細胞、in vitro無細胞系すなわち再構成生化学系があり得る。TMPをコードするポリヌクレオチドの発現における改変は、当分野で周知の任意の方法でアッセイする。通常は、TMP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量を定量することも可能であり、それにより、1つ若しくは複数の試験化合物に曝露される、または曝露されないポリヌクレオチドの発現の比較のための基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現における変化の検出は、ポリヌクレオチドの発現を改変する際に試験化合物が有効であることを示す。或る特定ポリヌクレオチドの改変発現に有効な化合物に対して、例えば分裂酵母 (Schizosaccharomyces pombe) 遺伝子発現系 (Atkins, D. 他 (1999) 米国特許第5, 932, 435号、Arndt, G. M. 他 (2000) Nucleic Acids Res. 28:E15)、またはHeLa細胞等のヒト細胞株 (Clarke, M. L. 他 (2000) Biochem. Biophys. Res. Commun. 268:8-13) を用いてスクリーニングを実行できる。本発明の或る特定の実施例は、或る特定ポリヌクレオチド配列に対するアンチセンス活性について、オリゴヌクレオチド類 (デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾したオリゴヌクレオチド) の組み合わせライブラリをスクリーニングする過程に関する (Bruce, T. W. 他 (1997) 米国特許第5, 686, 242号、Bruce, T. W. 他 (2000) 米国特許第6, 022, 691号)。

20

30

40

【0228】

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

【0229】

50

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

【0230】

本発明の或る更なる実施例は、薬剤として許容できる賦形剤と共に処方される活性成分を一般に持つ、或る組成物の投与に関する。賦形剤には例えば、糖、でんぷん、セルロース、ゴムおよびタンパク質がある。様々な処方が広く知られており、詳細は Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA) の最新版に記載されている。このような組成物は、TMP、TMPの抗体、擬態物質、アゴニスト、アンタゴニスト、またはTMPのインヒビターなどからなる。

10

【0231】

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

【0232】

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

20

【0233】

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

【0234】

TMPまたはその断片を含む高分子を直接細胞内に輸送するべく、特殊な形態に組成物が調製されるのが好ましい。例えば、細胞不透過性高分子を含むリポソーム製剤は、その高分子の細胞融合と細胞内送達とを促進し得る。別法では、TMPまたはその断片を、HIV Tat-1タンパク質の陽イオンN末端部に結合することもできる。このようにして生成された融合タンパク質類は、或るマウスモデル系の、脳を含む全ての組織の細胞群に形質導入することがわかっている (Schwarze, S. R. 他 (1999) Science 285:1569-1572)。

30

【0235】

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

40

【0236】

医学的に効果的な薬用量とは、症状や容態を回復させる、たとえばTMPまたはその断片、TMPの抗体、TMPのアゴニストまたはアンタゴニスト、インヒビターなど活性処方成分の量を指す。薬用有効度および毒性は、細胞培養または動物実験における標準的な薬学手法によって、例えばED₅₀（集団の50%の医薬的有效量）またはLD₅₀（集団の50%の致死量）を測定するなどして決定することができる。毒性効果の、治療効果に対する投与量比が治療指数であり、LD₅₀/ED₅₀比として表すことができる。高い治療指数を示すような組成物が望ましい。細胞培養アッセイと動物実験とから得られたデータは、ヒトに用いる投与量の範囲での調剤に用いられる。このような組成物に含まれる投与量は、毒性を殆どあるいは全く含まず、ED₅₀を含むような血中濃度の範囲にある

50

ことが好ましい。用いられる投与形態、患者の感受性および投与の経路によって、投与量はこの範囲内で様々に変わる。

【0237】

正確な投与量は、治療が必要な被験者に関する要素を考慮して、現場の医者が決定することになる。十分なレベルの活性成分を与え、あるいは所望の効果を維持すべく、用法および用量を調整する。被験者に関する要素としては、疾患の重症度、患者の一般的な健康状態、患者の年齢、体重および性別、投与の時間および頻度、薬剤の配合、反応感受性および治療に対する応答などを考慮し得る。作用期間が長い組成物は、特定の製剤の半減期およびクリアランス率によって3～4日毎に1度、1週間に1度、あるいは2週間に1度の間隔で投与し得る。

10

【0238】

通常の投与量は、投与の経路にもよるが約0.1～100,000 μ gであり、合計で約1gまでとする。特定の投与量および送達方法に関するガイダンスは文献に記載されており、現場の医者は通常それを利用することができる。当業者は、タンパク質またはそれらのインヒビター類についての処方とは異なる、ヌクレオチドについての処方を利用することになる。同様に、ポリヌクレオチドまたはポリペプチドの送達は、特定の細胞、症状、部位などに特異的なものとなる。

【0239】

(診断)

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

20

【0240】

TMPを測定するためのELISA,RIA,およびFACSなど、種々のプロトコルは、当分野では周知であり、変容した或いは異常なレベルのTMPの発現を診断するための基礎を提供する。正常あるいは標準的なTMPの発現の値は、複合体の形成に適した条件下で、正常な哺乳動物、例えばヒトなどの被験者から採取した体液または細胞抽出物と、TMPに対する抗体とを混合することによって確定する。標準複合体形成量は、種々の方法、例えば測光法で定量できる。被験者、対照、および、生検組織からの疾患サンプルでの、TMPの発現の量を標準値と比較する。標準値と被験者の値との偏差が、疾患を診断するパラメータを確定する。

30

【0241】

本発明の別の実施態様によれば、TMP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。用い得るポリヌクレオチドの例には、オリゴヌクレオチド配列、相補的RNAおよびDNA分子、そしてPNAが含まれる。これらのポリヌクレオチドを用いて、TMPの発現が疾患と相関し得る生検組織における遺伝子発現を検出し定量し得る。この診断アッセイを用いて、TMPの存在の有無、更には過剰な発現を判定し、治療時のTMPレベルの調節を監視する。

40

【0242】

或る実施形態では、TMPまたは近縁の分子をコードする、ゲノム配列などポリヌクレオチド配列を検出可能なPCRプローブ類とのハイブリダイゼーションを、TMPをコードする核酸配列を同定するために用いることができる。例えば5'調節領域など高度に特異的な領域か、それとも例えば保存されたモチーフなどやや特異性の低い領域から作られているかという、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

50

ンシーとにより、プローブが、TMPをコードする天然配列のみを同定するのか、或いは対立遺伝子変異体や関連配列を同定するかが決まる。

【0243】

プローブはまた、関連する配列の検出に利用でき、また、TMPをコードする任意の配列との少なくとも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、プローブはSEQ ID NO: 18-34の配列、或いはTMP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。

【0244】

TMP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、TMPをコードまたはTMP誘導体をコードするポリヌクレオチド配列をmRNAプローブの作製のためのベクターにクローニングする方法がある。mRNAプローブ作製用ベクター類は、当業者に知られており、市販されており、好適なRNAポリメラーゼと好適な標識されたヌクレオチドと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得る。ハイブリダイゼーションプローブ類は、種々のレポーターの集団によって標識され得る。レポーター集団の例としては、³²Pまたは³⁵Sなどの放射性核種が、あるいはアビジン／ビオチン結合系を介してプローブに結合されたアルカリホスファターゼなど酵素標識が挙げられる。

【0245】

TMPをコードするポリヌクレオチド配列は、TMPの発現に関係する疾患の診断の為に用い得る。限定するものではないがこのような疾患の例としては、生殖障害として、プロラクチン産生の障害、不妊（例えば卵管疾患、排卵不良、子宮内膜症、性周期の途絶、月経周期の途絶、多嚢胞性卵巣症候群、卵巣過剰刺激症候群、子宮内膜腫瘍や卵巣腫瘍、子宮筋腫、自己免疫障害、異所性妊娠、奇形発生、乳癌、乳房線維嚢胞病、乳漏症、精子形成の途絶、異常精子生理、精巣癌、前立腺癌、良性前立腺肥大、前立腺炎、ペーロニー病、インポテンス、男性乳癌、女性化乳房、高ゴナドトロピン性腺機能低下症、低ゴナドトロピン性腺機能低下症、仮性半陰陽、無精子症、早発卵巣不全、アクロシン欠損症、晩発思春期、逆行性射精、無射精、血管芽腫、嚢胞・クロム親和細胞腫（cystosphaeochromocytomas）、傍神経節腫、精巣上体の嚢胞腺腫、および内リンパ嚢腫瘍）が、発達障害として、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／ベッカー筋ジストロフィー、癲癇、性腺發育障害、WAGR症候群（ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神遅滞）、スミス・マジエニス症候群、骨髄異形成症候群、遺伝性粘液上皮異形成、遺伝性角皮症、遺伝性ニューロパシー（例えばシャルコー・マリー・トゥース病および神経線維腫症）、甲状腺機能低下症、水頭症、発作障害（例えばシデナム舞蹈病と脳性麻痺）二分脊椎、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感音難聴が含まれ、また心血管障害も含まれ、その中には動静脈瘤、アテローム性動脈硬化症、高血圧、脈管炎、レイノー病、動脈瘤、動脈解離、静脈瘤、血栓静脈炎および静脈血栓、血管系腫瘍、血栓溶解の合併症、バルーン血管形成術、血管置換術、冠動脈バイパス移植、うっ血性心不全、虚血性心疾患、狭心症、心筋梗塞、高血圧性心疾患、変性弁膜性心疾患、石灰化大動脈弁狭窄症、先天性2尖大動脈弁、僧帽弁輪部石灰化、僧帽弁逸脱、リウマチ熱、リウマチ性心疾患、感染性心内膜炎、非細菌性血栓性心内膜炎、全身性紅斑性狼瘡の心内膜炎、カルチノイド心疾患、心筋症、心筋炎、心膜炎、腫瘍性心疾患、先天性心臓疾患、心臓移植の合併症、先天性肺異常、肺拡張不全、肺うっ血および肺水腫、肺動脈塞栓症、肺出血、肺梗塞、肺高血圧症、血管硬化症、閉塞性肺疾患、拘束性肺疾患、慢性閉塞性肺疾患、肺気腫、慢性気管支炎、気管支喘息、気管支拡張症、細菌性肺炎、ウイルス性肺炎およびマイコプラズマ肺炎、肺膿瘍、肺結核、びまん性間質性疾患、塵肺症、サルコイド症、特発性肺線維症、剥離性間質性肺炎、過敏性肺炎、肺好酸球増加、閉塞性細気管支炎—器質化肺炎（bronchiolitis obliterans—organizing pneumonia）、びまん性肺出血症候群、グッドパスチャー症候群、特発性肺血鉄症、膠原血管病併発性肺疾患、肺胞タン

パク症、肺腫瘍、炎症性および非炎症性胸水、気胸症、胸膜腫瘍、薬物性肺疾患、放射線による肺疾患、および肺移植の合併症などが含まれる。神経障害として、癲癇、虚血性脳血管障害、脳卒中、脳腫瘍、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病その他の錐体外路障害、筋萎縮性側索硬化その他の運動ニューロン障害、進行性神経性筋萎縮症、網膜色素変性症（色素性網膜炎）、遺伝性運動失調、多発性硬化症その他の脱髄疾患、細菌性およびウイルス性髄膜炎、脳膿瘍、硬膜下膿瘍、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎および神経根炎、ウイルス性中枢神経系疾患と、クールー、クロイツフェルト-ヤコブ病およびゲルストマン-シュトロイスラー-シャインカー症候群を含むプリオン病と、致死性家族性不眠症、神経系の栄養病および代謝病、神経線維腫症、結節硬化症、小脳網膜性血管腫症（cerebelloretinal hemangioblastomatosis）、脳3叉神経血管症候群、精神遅滞、ダウン症候群を含むその他の中枢神経系発達障害、脳性麻痺、神経骨格異常、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィーその他の神経筋疾患、末梢神経疾患、皮膚筋炎および多発性筋炎と、遺伝性、代謝性、内分泌性および中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分障害、不安障害および統合失調症（精神分裂病）を含む精神障害と、季節性感情障害（SAD）と、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、遅発性ジスキネジア、ジストニー、パラノイド精神病、帯状疱疹後神経痛、トゥーレット病、進行性核上麻痺、皮質基底核変性症、家族性前頭側頭痴呆が含まれ、また胃腸障害も含まれ、その中には嚥下障害、消化性食道炎、食道痙攣、食道狭窄、食道癌、消化不良、消化障害、胃炎、胃癌、食欲不振、悪心、嘔吐、胃不全麻痺、洞または幽門の浮腫、腹部アンギナ、胸焼け、胃腸炎、イレウス、腸管感染、消化性潰瘍、胆石症、胆嚢炎、胆汁うっ滞、膵臓炎、膵臓癌、胆道疾患、肝炎、高ビリルビン血症、硬変症、肝臓の受動的うっ血、ヘパトーム、感染性大腸炎、潰瘍性大腸炎、潰瘍性直腸炎、クローン病、ホウィツブル病、マロリー-ワイス症候群、結腸癌、結腸閉塞、過敏性腸症候群、短小腸症候群、下痢、便秘、胃腸出血、後天性免疫不全症候群（AIDS）腸症、黄疸、肝性脳症、肝腎症候群、脂肪肝、血色素症、ウィルソン病、 α_1 アンチトリプシン欠損症、ライ症候群、原発性硬化性胆管炎、肝梗塞、門脈循環閉塞および血栓、小葉中心壊死、肝臓紫斑病、肝静脈血栓、肝静脈閉塞症、子癇前症、子癇、妊娠性急性脂肪肝、妊娠性肝臓内胆汁うっ滞と、結節性過形成および腺腫、癌腫を含む肝癌が含まれ、脂質代謝障害として、脂肪肝、胆汁うっ滞、原発性胆汁性肝硬変、カルニチン欠乏症、カルニチンパルミトイルトランスフェラーゼ欠乏症、ミオアデニル酸デアミナーゼ欠損症、高トリグリセリド血症と、ファブリー病などの脂質貯蔵病、ゴーシェ病、ニーマン-ピック病、異染性白質ジストロフィー、副腎白質ジストロフィー、GM₂ガングリオシド蓄積症、セロイドリポフスチン症、無 β -リポ蛋白血症、タンジアー病、高リポ蛋白血症、糖尿病、脂肪異栄養症、脂肪腫症、急性皮下脂肪組織炎、播種性脂肪壊死、有痛脂肪症、リポイド副腎過形成、微小変化型ネフローゼ、脂肪腫、アテローム性動脈硬化症、高コレステロール血症、高トリグリセリド血症を伴った高コレステロール血症、原発性低 α リポ蛋白血症、甲状腺機能低下症、腎疾患、肝疾患、レシチン：コレステロールアシルトランスフェラーゼ欠損症、脳腱黄色腫症、シトステロール血症、低コレステロール血症、テイ-サックス病、サンドホフ病、高脂血症、高脂質血症、脂質筋障害、肥満症が、細胞増殖障害として、光線性角化症、動脈硬化症、アテローム性硬化症、滑液包炎、硬変、肝炎、混合性結合組織病（MCTD）、骨髓線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症と、腺癌、白血病、リンパ腫、黒色腫、骨髄腫、肉腫、奇形癌を含む癌、具体的には副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌等が含まれ、また自己免疫/炎症疾患も含まれ、その中には後天性免疫不全症候群（AIDS）、アジソン病、成人呼吸窮迫症候群、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー（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、リンパ球傷害因子

性偶発性リンパ球減少症 (episodic lymphopenia with lymphocytotoxins)、新生児溶血性疾患 (胎児赤芽球症)、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性腸症候群、多発性硬化症、重症筋無力症、心筋または心膜の炎症、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性紅斑性狼瘡、全身性強皮症 (systemic sclerosis)、血小板減少性紫斑病、潰瘍性大腸炎、ブドウ膜炎、ウェルナー症候群、癌の合併症、血液透析、体外循環、ウイルス感染症、細菌性感染症、真菌性感染症、寄生虫感染症、原虫感染症、蠕虫の感染症および外傷が含まれる。TMPをコードするポリヌクレオチド配列は、サザーン法やノーザン法、ドット
10
ブロット法、或いはその他の膜系の技術、PCR法、ディップスティック (dipstick)、ピン (pin)、およびマルチフォーマットのELISA式アッセイ、および、変容したTMP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可能である。このような定性方法または定量方法は、当分野で公知である。

【0246】

或る特定の形態では、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、TMPをコードするヌクレオチド配列が有用であり得る。TMPをコードするヌクレオチド配列を、標準的な方法で標識化し、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。好適なインキュ
20
ベーション期間が経過したらサンプルを洗浄し、シグナルを定量して標準値と比較する。患者のサンプルのシグナルの量が、対照サンプルと較べて著しく変化した場合は、サンプル内のTMPをコードするヌクレオチド配列の変異レベルにより、関連する疾患の存在が明らかになる。このようなアッセイは、動物実験、臨床試験における特定の治療効果を推定するため、あるいは個々の患者の治療をモニターするために用いることもできる。

【0247】

TMPの発現に関連する疾患の診断の基準を提供するために、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あるいは増幅に好適な条件の下、動物あるいはヒトのいずれかの正常な被験者から抽出された体液あるいは細胞と、TMPをコードする配列あるいはその断片とを混合することにより達成され得る。実質
30
的に精製されたポリヌクレオチドを既知量で用いて行った実験から得た値を正常な対象から得た値と比較することにより、標準ハイブリダイゼーションを定量することができる。このようにして得た標準値は、疾患の徴候を示す患者から得たサンプルから得た値と比較することができる。標準値からの偏差を用いて疾患の存在を確定する。

【0248】

疾患の存在が確定されて治療プロトコルが開始されると、患者の発現レベルが正常な被検者に観察されるレベルに近づき始めたかどうかを測定するため、ハイブリダイゼーションアッセイを定期的に繰り返し得る。連続アッセイから得られた結果を用いて、数日から数
40
ヶ月の期間にわたる治療の効果を示し得る。

【0249】

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

【0250】

TMPをコードする配列から設計されたオリゴヌクレオチドのさらなる診断への利用には、PCRの利用が含まれ得る。これらのオリゴマーは、化学的に合成するか、酵素により生産するか、あるいは in vitro で産出し得る。オリゴマーは、好ましくはTMP
50
をコードするポリヌクレオチドの断片を、或いはTMPをコードするポリヌクレオチドに

対し相補的なポリヌクレオチドの断片を含み、最適な条件下で、特定の遺伝子や条件を識別するために利用される。また、オリゴマーは、やや緩いストリンジェンシー条件下で、近縁のDNAあるいはRNA配列の検出、定量、あるいはその両方のため用いることが可能である。

【0251】

或る特定の態様においては、TMP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、多くの場合にヒトの先天性または後天性遺伝病の原因となるような置換、挿入および欠失である。限定するものではないがSNPの検出方法には、SSCP (single-stranded conformation polymorphism) および蛍光SSCP (fSSCP) がある。SSCPでは、TMP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と、ポリメラーゼ連鎖反応法 (PCR) とを用いたDNAの増幅を行う。DNAは例えば、病変組織または正常組織、生検サンプル、体液その他に由来し得る。DNA内のSNPは、一本鎖形のPCR産物の2次および3次構造に差異を生じさせる。差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。fSSCPでは、オリゴヌクレオチドプライマーを蛍光性に標識する。それによってDNAシーケンシング機などの高処理機器でアンプリマー (amplimer) の検出が可能になる。更に、インシリコSNP (in silico SNP, isSNP) と呼ばれる配列データベース分析法は、一般的なコンセンサス配列に配列されるような個々の重畳するDNA断片の配列を比較することにより、多型性を同定し得る。これらのコンピュータベースの方法は、DNAの実験室での調製に、または統計モデルとDNA配列クロマトグラムの自動分析とを用いたシーケンシングのエラーに起因する配列変異を、フィルタリングして除去する。別の態様では、例えば高処理MASSARRAYシステム (Sequenom, San Diego CA) を用いた質量分析によりSNPを検出し、特徴付ける。

10

20

【0252】

TMPの発現を定量するために用い得る方法には、ヌクレオチドの放射標識またはビオチン標識、対照核酸の共増幅 (coamplification)、および標準曲線から得た結果の補間もある (例えば、Melby, P. C. 他 (1993) J. Immunol. Methods, 159: 235-244; Duplaa, C. 他 (1993) Anal. Biochem. 212: 229-236を参照)。複数サンプルの定量速度を加速するには、目的のオリゴマーまたはポリヌクレオチドを種々の希釈液中に置き、分光光度法または比色反応によって定量が迅速になるような高処理フォーマットでのアッセイを行い得る。

30

【0253】

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エレメントとして用いることができる。多数の遺伝子の相対発現レベルを同時にモニターする転写イメージング技術にマイクロアレイを用いることが可能である。これについては、以下に記載する。マイクロアレイはまた、遺伝変異体、突然変異および多型性の同定に用いることができる。この情報を用いることで、遺伝子機能を決定し、疾患の遺伝的根拠を理解し、疾患を診断し、遺伝子発現の機能としての疾病の進行/後退をモニターし、疾病治療における薬剤の活性を開発およびモニターすることができる。特に、患者にとって最適且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開発することができる。例えば、患者の薬理ゲノムプロファイルに基づき、患者に対して高度に効果的で副作用を殆ど示さない治療薬を選択することができる。

40

【0254】

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

50

【0255】

或る実施態様は、或る組織または細胞タイプの転写イメージを生成する、本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞タイプによる、遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所与の条件下で所与の時間に発現した遺伝子の数および相対存在量を定量することにより分析し得る（Sei l h a m e r 他、米国特許第5, 840, 484号「Comparative Gene Transc r i p t A n a l y s i s」を参照。該特許は特に引用することを以って本明細書の一部となす）。したがって、特定の組織または細胞タイプの転写物または逆転写物全体に、本発明のポリヌクレオチドまたはそれらの相補体をハイブリダイズすることにより、転写イメージを生成し得る。或る実施例では、本発明のポリヌクレオチドまたはそれらの相補体が1マイクロアレイ上に複数エレメントの1サブセットを含むような高処理フォーマットでハイブリダイゼーションを発生させる。結果として得られる転写イメージは、遺伝子活性のプロフィールを提供し得る。

10

【0256】

転写イメージは、組織、細胞株、生検またはその生物学的サンプルから単離した転写物を用いて生成し得る。転写イメージはしたがって、組織または生検サンプルの場合には i n v i v o、細胞株の場合には i n v i t r oでの遺伝子発現を反映する。

【0257】

本発明のポリヌクレオチドの発現プロフィールを作製する転写イメージはまた、i n v i t r oモデル系および薬剤の前臨床評価に、あるいは工業的または天然の環境化合物の毒性試験に関連して使用し得る。全ての化合物は、作用および毒性のメカニズムを標示する、しばしば分子フィンガープリントまたは毒性シグネチャと称される、特徴的な遺伝子発現パターンを惹起する（Nuway s i r, E. F. 他（1999）Mol. Ca r c i n o g. 24:153-159、Ste i n e r, S. および N. L. A n d e r s o n（2000）Tox i c o l. L e t t. 112-113:467-471、該文献は特に引用を以って本明細書の一部となす）。試験化合物が、既知毒性を有する化合物のシグネチャと同一のシグネチャを有する場合には、毒性特性を共有している可能性がある。フィンガープリントまたはシグネチャは、多数の遺伝子および遺伝子ファミリーからの発現情報を含んでいる場合には、最も有用且つ正確である。理想的には、発現の、ゲノム全域にわたる測定が、最高品質のシグネチャを提供する。たとえ、発現が任意の試験された化合物によって変容しない遺伝子があったとしても、それらの発現レベルを残りの発現データを標準化するために使用できるため、それらの遺伝子は重要である。標準化手順は、種々の化合物で処理した後の発現データの比較に有用である。或る毒性シグネチャのエレメント群に遺伝子機能を割り当てることは毒性機序の解釈に役立つが、毒性の予測につながる、シグネチャ群の統計的マッチングには、遺伝子機能の知識は必要でない（例えば2000年2月29日に米国国立環境健康科学研究所（National I n s t i t u t e o f E n v i r o n m e n t a l H e a l t h S c i e n c e s）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）。したがって、毒性シグネチャを用いる中毒学的スクリーニングの際に、全ての発現した遺伝子配列を含めることは、重要且つ望ましい。

20

30

40

【0258】

或る実施態様では、核酸を有する生物学的サンプルを試験化合物で処理することにより、この試験化合物の毒性を算定する。処理した生物学的サンプル中で発現した核酸は、本発明のポリヌクレオチドに特異的な1つ以上のプローブでハイブリダイズし、それによって本発明のポリヌクレオチドに対応する転写レベルを定量し得る。処理した生物学的サンプル中の転写レベルを、未処理生物学的サンプル中のレベルと比較する。両サンプルの転写レベルの差は、処理済サンプル中で試験化合物が引き起こす毒性反応を標示する。

【0259】

別の実施態様は、本発明のポリペプチド配列を用いて組織または細胞タイプのプロテオ-

50

ムを分析することに関連する。プロテオームの語は、特定の組織または細胞タイプでのタンパク質発現の包括的パターンを指す。プロテオームの各タンパク質成分は、個々に更なる分析の対象とすることができる。プロテオーム発現パターンすなわちプロフィールは、所与の条件下で所与の時間に発現したタンパク質の数および相対存在量を定量することにより分析し得る。したがって、或る細胞のプロテオームのプロフィールは、特定の組織または細胞タイプのポリペプチドを分離および分析することにより作成し得る。或る実施態様では、1次元等電点電気泳動によりサンプルからタンパク質を分離し、2次元ドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に応じて分離するような2次元ゲル電気泳動により、分離が達成される（前出のSteiner および Anderson）。タンパク質は、通常はクーマシーブルー、あるいは銀染色液または蛍光染色液などの物質を用いてゲルを染色することにより、分散した、独自の位置にある点としてゲル中で可視化される。各タンパク質スポットの光学密度は、通常、サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処理または未処理のいずれかの生物学的サンプルからの、同等に位置したタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を同定する。スポット内のタンパク質は、例えば化学的または酵素的切断とそれに続く質量分析を用いる標準的な方法を用いて部分的にシーケンシングする。或るスポット内のタンパク質の同一性は、その部分配列を、好適には少なくとも5個の連続するアミノ酸残基を、本発明のポリペプチド配列と比較することにより決定し得る。場合によっては、決定的なタンパク質同定のための更なる配列データが得られる。

10

20

【0260】

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

30

【0261】

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に分析するべきである。いくつかの組織のいくつかのタンパク質については、転写物の存在量とタンパク質の存在量との相関が乏しいので（Anderson, N. L. および J. Seilhamer（1997）Electrophoresis 18:533-537）、転写イメージにはそれ程影響しないがプロテオームのプロフィールを改変するような化合物の分析において、プロテオーム毒性シグネチャは有用たり得る。更に、体液中の転写物の分析はmRNAの急速な分解のために困難なので、プロテオームのプロフィール作成はこのような場合により信頼し得、情報価値があり得る。

40

【0262】

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

【0263】

別の実施態様では、タンパク質を含有する生物学的サンプルを試験化合物で処理すること

50

により、試験化合物の毒性を算定する。生物学的サンプルから得たタンパク質は、本発明のポリペプチドに特異的な抗体を用いてインキュベートする。抗体が認識したタンパク質の量を定量する。処理された生物学的サンプル中のタンパク質の量を、未処理生物学的サンプル中のタンパク質の量と比較する。両サンプルのタンパク質の量の差は、処理サンプル中の試験化合物に対する毒性反応を標示する。

【0264】

マイクロアレイは、本技術分野で既知の方法で調製し、使用し、分析する (Brennan, T. M. 他 (1995) 米国特許第 5, 474, 796 号、Schena, M. 他 (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619、Baldeschweiler 他 (1995) PCT 出願第 WO 95/251116 号、Shalon, D. 他 (1995) PCT 出願第 WO 95/35505 号、Heller, R. A. 他 (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155、Heller, M. J. 他 (1997) 米国特許第 5, 605, 662 号などを参照)。様々なタイプのマイクロアレイが公知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, 編集 (1999) Oxford University Press, London に記載されている。該文献は、特に引用を以て本明細書の一部となす。

10

【0265】

本発明の別の実施態様ではまた、TMP をコードする核酸配列を用いて、天然のゲノム配列をマッピングするのに有用なハイブリダイゼーションプローブを作製することが可能である。コード配列または非コード配列のいずれかを用いることができ、いくつかの例では、コード配列より非コード配列が好ましい。例えば、多重遺伝子族のメンバー内でのコード配列の保存により、染色体マッピング中に望ましくないクロスハイブリダイゼーションが生じる可能性がある。配列は、或る特定の染色体に、または或る染色体の或る特定領域に、または人為形成の染色体、例えば、ヒト人工染色体 (HAC)、酵母人工染色体 (YAC)、細菌人工染色体 (BAC)、細菌 P1 構成物、あるいは単一染色体 cDNA ライブラリ群に対してマッピングされる (例えば Harrington, J. J. 他 (1997) Nat. Genet. 15:345-355; Price, C. M. (1993) Blood Rev. 7:127-134; Trask, B. J. (1991) Trends Genet. 7:149-154 を参照)。一度マッピングすると、本発明の核酸配列を用いて例えば病状の遺伝を特定の染色体領域の遺伝または制限酵素断片長多型 (RFLP) と関連させるような遺伝子連鎖地図を開発し得る (例えば、Lander, E. S. および D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357 を参照)。

20

30

【0266】

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

40

【0267】

確定した染色体マーカー類を用いた連鎖分析などの物理的マッピング技術、および染色体標本の原位ハイブリッド形成法を用いて、遺伝地図を拡張することができる。例えばマウスなど別の哺乳動物の染色体上に遺伝子を配置することにより、正確な染色体上の遺伝子座が未知でも、関連するマーカー類をしばしば明らかにし得る。この情報は、位置クローニングその他の遺伝子発見技術を用いて疾患遺伝子を探る研究者にとって価値がある。

50

疾患または症候群に關与する遺伝子（群）が、血管拡張性失調症の11q22-23領域など、特定のゲノム領域への遺伝的連鎖によって、大まかに位置決めがなされると、該領域に対するいかなる配列マッピングも、更なる調査のための関連遺伝子あるいは調節遺伝子を表し得る（Gatti, R. A. 他（1988）Nature 336:577-580などを参照）。転座、反転などに起因する、健常者、保有者、罹病者の三者間における染色体位置の相違を発見するために、本発明のヌクレオチド配列を用いてもよい。

【0268】

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

【0269】

別の薬剤スクリーニング方法は、目的のタンパク質に対して好適な結合親和性を有する化合物を高い処理能力でスクリーニングするために用いられる（Geysen他（1984）PCT出願番号 WO84/03564などを参照）。この方法においては、多数の様々な小さな試験用化合物を固体基板上で合成する。試験用化合物は、TMP、或いはその断片と反応してから洗浄される。精製したTMPはまた、上記した薬剤のスクリーニング技術において用いるプレート上で直接コーティングすることもできる。別法では、非中和抗体を用いてペプチドを捕捉し、ペプチドを固体支持物に固定することもできる。

【0270】

別の実施例では、TMPと特異結合可能な中和抗体がTMPとの結合について試験化合物と競合する、競合的薬剤スクリーニングアッセイを用いることができる。この方法では、抗体を用いて、1つ以上の抗原決定基をTMPと共有するどのペプチドの存在をも検出できる。

【0271】

別の実施例では、TMPをコードするヌクレオチド配列を、将来に開発される分子生物学技術であり、現在知られているヌクレオチド配列の特性（限定はされないが、トリプレット遺伝コード、特異的な塩基対相互作用等を含む）に依存する新技術に用い得る。

【0272】

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

【0273】

前述した、および以下に記載する全ての特許出願、特許、刊行物、米国特許出願第60/244,017号、同第60/252,855号、同第60/251,825号、および同第60/255,085号も含めて、言及することをもって本明細書の一部とする。

【0274】

（実施例）

1 cDNAライブラリの作製

Incyte cDNA群の由来は、LIFESEQ GOLDデータベース（Incyte Genomics, Palo Alto CA）に記載されたcDNAライブラリ群であり、表4の列5に列記した。幾つかの組織はホモジナイズしてグアニジニウムイソチオシアネート溶液に溶解し、また、ホモジナイズしてフェノールまたは変性剤類の好適な混合液に溶解した組織もある。変性剤の混合液は、例えばフェノールとグアニジンイソチオシアネートとの単相溶液であるTRIZOL（Life Technologies）である。結果として得られた溶解物は、塩化セシウムのクッション液の上に重層して遠心分離するか、クロロホルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方、あるいは別の慣例的方法を用いて、溶解物からRNAを沈殿させた。

【0275】

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

【0276】

場合によってはStratagene社へのRNA提供を行い、対応するcDNAライブラリをStratagene社が作製した。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)を用いて本技術分野で既知の推奨方法または類似の方法でcDNAを合成し、cDNAライブラリを作製した(前出のAusubel, 1997, 5.1-6.6ユニットなどを参照)。逆転写は、オリゴd(T)またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖cDNAに連結反応させ、好適な制限酵素または酵素群でcDNAを消化した。殆どのライブラリに対して、cDNAのサイズ選択(300~1000bp)は、SEPHACRYL S1000、SEPHAROSE CL2BまたはSEPHAROSE CL4Bカラムクロマトグラフィー(Amersham Pharmacia Biotech)、あるいは調製用アガロースゲル電気泳動法を用いて行った。cDNAは好適なプラスミドのポリリンカーの適合する制限酵素部位でライゲーションされた。好適なプラスミドは、例えばPBUESCRIPTプラスミド(Stratagene)、PSPORT1プラスミド(Life Technologies)PCDNA2.1プラスミド(Invitrogen, Carlsbad CA)、PBK-CMVプラスミド(Stratagene)、PCR2-TOPOTAプラスミド(Invitrogen)、PCMV-ICISプラスミド(Stratagene)、pIGEN(Incyte Genomics, Palo Alto CA)またはpINCY(Incyte Genomics)、またはこれらの誘導体である。組換えプラスミドは、Stratagene社のXL1-Blue、XL1-BlueMRFまたはSOLR、あるいはLife Technologies社のDH5α、DH10BまたはElectroMAX DH10Bなど適格な大腸菌細胞に形質転換した。

【0277】

2 cDNAクローンの単離

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

【0278】

別法では、高処理フォーマットで直接結合PCR法を用い、宿主細胞溶解物からプラスミドDNAを増幅した(Rao, V. B. (1994) Anal. Biochem. 216:1-14)。宿主細胞の溶解および熱サイクリング過程は、単一反応混合液中で行った。サンプルを加工し、それを384穴プレート内で保管し、増幅したプラスミドDNAの濃度をPICOGREEN色素(Molecular Probes, Eugene OR)およびFLUOROSKAN II 蛍光スキャナ(Labsystems Oy,

Helsinki, Finland) を用いて蛍光分析的に定量した。

【0279】

3 シークエンシングおよび分析

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

【0280】

Incyte cDNA に由来するポリヌクレオチド配列は、ベクター、リンカーおよびポリ (A) 配列を除去し、あいまいな塩基対をマスクすることによって有効性を確認した。その際、BLAST と、動的プログラミングと、隣接ジヌクレオチド頻度分析とに基づく、アルゴリズムとプログラムとを用いた。次に、Incyte cDNA 配列またはそれらの翻訳の問い合わせを、以下のデータベース群より選択した 1 つに対して行った。すなわち、公共のデータベース群、例えば GenBank の霊長類およびげっ歯類、哺乳動物、脊椎動物、真核生物のデータベースと、BLOCKS、PRINTS、DOMO、PRODOM、PROTEOME の、ヒト、ラット、マウス、線虫 (Caenorhabditis elegans)、酵母 (Saccharomyces cerevisiae)、分裂酵母 (Schizosaccharomyces pombe) および Candida albicans (Incyte Genomics, Palo Alto CA) からの配列群を持つデータベース群、および、PFAM 等の、隠れマルコフモデル (HMM) ベースのタンパク質ファミリーデータベース群である (HMM は、遺伝子ファミリーのコンセンサス 1 次構造を分析する確率的アプローチである。例えば、Eddy, S. R. (1996) Curr. Opin. Struct. Biol. 6:361-365 を参照)。問合せは、BLAST、FASTA、BLIMPS、および HMMER に基づくプログラムを用いて行った。Incyte cDNA 配列を構築し、完全長のポリヌクレオチド配列を産出した。あるいは、GenBank cDNA 群、GenBank EST 群、ステッチされた配列群、ストレッチされた配列群、または Genscan 予測コード配列群 (実施例 4 および 5 を参照) を用い、Incyte cDNA の集団を完全長まで伸長させた。構築には Phred と Phrap と Consed とに基づくプログラムを用い、GenMark と BLAST と FASTA とに基づくプログラムを用いて、cDNA の集団を、オープンリーディングフレームについてスクリーニングした。完全長ポリヌクレオチド配列を翻訳し、対応する完全長ポリペプチド配列を誘導した。あるいは、本発明のポリペプチドは、完全長翻訳ポリペプチドの任意のメチオニン残基で開始し得る。完全長ポリペプチド配列群の続いての分析としての問い合わせを、GenBank タンパク質データベース群 (genpept)、SwissProt、PROTEOM E データベース群、BLOCKS、PRINTS、DOMO、PRODOM および Pro

s i t e等のデータベースや、P F A Mなど隠れマルコフモデル（HMM）ベースのタンパク質ファミリーデータベース群に対し行った。完全長ポリヌクレオチド配列はまた、M A C D N A S I S P R Oソフトウェア（日立ソフトウェアエンジニアリング、S o u t h S a n F r a n c i s c o C A）およびL A S E R G E N Eソフトウェア（D N A S T A R）を用いて分析した。ポリヌクレオチドとポリペプチドとの配列アラインメントは、M E G A L I G Nマルチシーケンサアラインメントプログラム（D N A S T A R）に組み込まれたC L U S T A Lアルゴリズムが指定する、デフォルトパラメータを用いて生成する。これは、アラインメントした配列間の一致率も計算する。

【0281】

I n c y t e c D N Aと完全長配列との分析と構築とに利用したツールとプログラムとアルゴリズムとの概略と、適用可能な説明、引用文献、閾値パラメータを表7に示す。用いたツール、プログラムおよびアルゴリズムを表7の列1に、それらの簡単な説明を列2に示す。列3は好適な参考文献であり、全ての文献は引用を以って全文を本明細書の一部となす。適用可能な場合には、列4は、2つの配列が一致する強さを評価するために用いた、スコア、確率値その他のパラメータを示す（スコアが高いほど、または確率値が低いほど、2配列間の同一性が高い）。

10

【0282】

完全長ポリヌクレオチド配列およびポリペプチド配列の構築および分析に用いる上記のプログラムは、S E Q I D N O：18-34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および増幅技術に有用な約20～約4000ヌクレオチドの断片を、表4の列4に示した。

20

【0283】

4 ゲノムDNAからのコード配列の同定および編集

推定上の膜貫通タンパク質類の同定には、先ずG e n s c a n遺伝子同定プログラムを、公共のゲノム配列データベース（例えば、g b p r iやg b h t g）に対して実行した。G e n s c a nは、様々な生物からのゲノムDNA配列を分析する汎用遺伝子同定プログラムである（B u r g e , C . および S . K a r l i n（1997）J . M o l . B i o l . 268：78-94、B u r g e , C . および S . K a r l i n（1998）C u r r . O p i n . S t r u c t . B i o l . 8：346-354参照）。このプログラムは予測されたエキソン群を連結し、メチオニンから終止コドンに及ぶ、構築されたc D N A配列を形成する。G e n s c a nの出力は、ポリヌクレオチドおよびポリペプチド配列のF A S T Aデータベースである。G e n s c a nが一度に分析する配列の最大範囲は、30kbに設定した。これらのG e n s c a n予測c D N A配列の内、どの配列が膜貫通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、P F A Mモデルにおいて膜貫通タンパク質について問合せて分析した。潜在的な膜貫通タンパク質はまた、既に膜貫通タンパク質として注釈が付けられたI n c y t e c D N A配列群に対する相同性により、同定した。これら選択されたG e n s c a n予測配列を、次にB L A S T分析により、公共データベースg e n p e p tおよびg b p r iと比較した。必要であれば、g e n p e p tからのトップB L A S Tヒットと比較することによりG e n s c a n予測配列を編集し、余分なまたは取り除かれたエキソンなど、G e n s c a nが予測した配列におけるエラーを補正した。B L A S T分析はまた、G e n s c a n予測配列の、いかなるI n c y t e c D N Aまたは公共c D N Aカバレッジ（c o v e r a g e）の発見にも用いられ、したがって転写の証拠を提供する。I n c y t e c D N Aカバレッジが利用できる場合には、この情報を用いてG e n s c a n予測配列を補正または確認した。完全長ポリヌクレオチド配列は、実施例3に記載した構築プロセスを用いて、I n c y t e c D N A配列および／または公共c D N A配列でG e n s c a n予測コード配列を構築して得た。あるいは、完全長ポリヌクレオチド配列は、編集後または非編集のG e n s c a n予測コード配列に、完全に由来する。

30

40

【0284】

5 c D N A配列データを使ったゲノム配列データの構築

50

スティッチ配列 (Stitched Sequence)

部分cDNA配列群は、実施例4に記載のGenscan遺伝子同定プログラムが予測したエキソン群を用いて伸長させた。実施例3に記載したように構築した部分cDNA群を、ゲノムDNAにマッピングした。また、関連するcDNA群と、1つ以上のゲノム配列からGenscan予測されたエキソン群とを含むクラスター群に分解した。cDNAとゲノムとの情報を統合すべく、グラフ理論および動的プログラミングに基づく或るアルゴリズムを用いて各クラスターを分析し、潜在的スプライシング変異配列群を生成した。配列群を続いて確認、編集または伸長し、完全長配列を創出した。区間の全長が、2つ以上の配列に在るような配列区間群をクラスター内で同定し、そのように同定された区間群は、推移性により、等しいと考えた。例えば、或る区間が、1つのcDNAと2つのゲノム配列とに在る場合、3つの区間は全て等しいと考えた。このプロセスにより、無関係だが連続したゲノム配列群を、cDNA配列で結び合わせて架橋し得る。このようにして同定した区間を、それらの親配列 (parent sequences) に沿って現われるようにスティッチアルゴリズムで「縫い合わせ」、可能な限り最長の配列と変異配列群とを生成する。1種類の親配列に沿って続く区間間の連鎖 (cDNA-cDNAまたはゲノム配列-ゲノム配列) は、親の種類を変える連鎖 (cDNA-ゲノム配列) より優先した。結果として得たスティッチ配列群を、BLAST分析で翻訳し、公共データベース genpept および gbpr i と比較した。Genscanが予測した不正確なエキソン群は、genpeptからのトップBLASTヒットとの比較により補正した。必要な場合、追加cDNA配列群を用いるかゲノムDNAの検査により、配列群を更に伸長させた。

10

20

【0285】

ストレッチ配列 (Stretched Sequence)

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

30

【0286】

6 TMPをコードするポリヌクレオチドの染色体マッピング

SEQ ID NO: 18-34を構築するために用いた配列を、BLAST他のSmith-Watermanアルゴリズムの実装を用いて、Incyte LIFESEQデータベースおよび公共のドメインデータベース群の配列と比較した。SEQ ID NO: 18-34と一致するこれらのデータベースの配列を、Phrapなどの構築アルゴリズム (表7) を使用して、連続および重複した配列のクラスター群に組み入れた。スタンフォード・ヒトゲノムセンター (SHGC)、ホワイトヘッド・ゲノム研究所 (WIGR)、GeneThonなどの公的な情報源から入手可能な放射線ハイブリッドおよび遺伝地図データを用いて、いずれかのクラスター化された配列が既にマッピングされているかを判定した。マッピングされた配列が或るクラスターに含まれている場合、そのクラスターの全配列が、個々の配列番号と共に、地図上の位置に割り当てられた。

40

【0287】

地図上の位置は、ヒト染色体の範囲または間隔として表される。センチモルガン間隔の地図上の位置は、染色体のpアームの末端に対して測定する (センチモルガン (cM) は、

50

染色体マーカー間の組換え頻度に基づく計測単位である。平均して、1 cMは、ヒトDNAの1メガベース (Mb) にほぼ等しい。尤も、この値は、組換えのホットスポットおよびコールドスポットに起因して広範囲に変化する)。cM距離は、各クラスター内に配列が含まれる放射線ハイブリッドマーカー類に対して境界を提供するGenethonによってマッピングされた遺伝マーカー群に基づく。NCBI「GeneMap'99」(<http://www.ncbi.nlm.nih.gov/genemap/>) など一般個人が入手可能なヒトゲノム地図その他の資源を用いて、既に同定されている疾患遺伝子類が、上記の区間内若しくは近傍にマップされているかを判定できる。

【0288】

このようにして、SEQ ID NO: 22は、11番染色体の59.50から62.50センチモルガンの区間内にマップされ、またSEQ ID NO: 26は、1番染色体の179.2から186.4センチモルガンの区間内にマップされた。

【0289】

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技術であり、特定の細胞種または組織からのRNAが結合している膜への、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に関与している (前出のSambrook, 7章、同Ausbell (1995) 4章および16章などを参照)。

【0290】

BLASTを応用した類似のコンピュータ技術を用いて、GenBankやLIFESEQ (Incyte Genomics) などのcDNAデータベースにおいて、同一または関連分子を検索した。ノーザン分析は、いくつかの膜系ハイブリダイゼーションよりも非常に速い。更に、特定の一致を厳密なあるいは相同的なものとして分類するか否かを決定するため、コンピュータ検索の感度を修正することができる。検索の基準は積スコアであり、次式で定義される。

【0291】

【数1】

(BLAST スコア×配列一致率)

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

積スコアは、2つの配列間の類似度と、配列が一致する長さとの両方を考慮している。積スコアは、0～100の、正規化された値であり、次のようにして求める。BLASTスコアにヌクレオチドの配列一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除する。或る高スコアリングセグメント対 (HSP) 内の一致する各塩基に+5のスコアを割り当て、各不一致塩基に-4を割り当てることにより、BLASTスコアを計算する。2つの配列は、2以上のHSPを共有し得る (ギャップにより隔離される)。2以上のHSPがある場合には、最高BLASTスコアのセグメント対を用いて積スコアを計算する。積スコアは、断片的重畳とBLASTアラインメントの質とのバランスを表す。例えば積スコア100は、比較した2つの配列の短い方の長さ全体にわたって100%一致する場合のみ得られる。積スコア70は、一端が100%一致し、70%重畳しているか、他端が88%一致し、100%重畳しているかのいずれかの場合に得られる。積スコア50は、一端が100%一致し、50%重畳しているか、79%一致し、100%重畳しているかのいずれかの場合に得られる。

【0292】

或いは、TMPをコードするポリヌクレオチド配列を、由来する組織に対して分析する。例えば或る完全長配列は、Incyte cDNA配列 (実施例3を参照) と少なくとも一部は重畳するように構築される。各cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各ヒト組織は、以下の臓器/組織カテゴリーの1つに分類される。すなわち心血管系、結合組織、消化器系、胎芽構造、内分泌系、外分泌腺、女性生殖器

、男性生殖器、生殖細胞、血液および免疫系、肝、筋骨格系、神経系、脾臓、呼吸器系、感覚器、皮膚、顎口腔系、非分類性／混合性または尿路である。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。同様に、各ヒト組織は、以下の疾患／病状カテゴリーすなわち癌、細胞株、発達、炎症、神経性、外傷、心血管、プール、その他の1つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。得られるパーセンテージは、TMPをコードするcDNAの、組織特異的発現および疾患特異的発現を反映する。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。

【0293】

10

8 TMPをコードするポリヌクレオチドの伸長

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

【0294】

20

選択したヒトcDNAライブラリ群を用い、配列を伸長した。2段階以上の伸長が必要または望ましい場合には、付加的プライマーあるいはプライマーのネステッドセットを設計した。

【0295】

高忠実度の増幅が、当業者によく知られている方法を利用したPCR法によって得られた。PCRは、PTC-200 サーマルサイクラー (MJ Research) を用いて96穴プレート内で行った。反応混合液は、鋳型DNAを有し、また、200 nmolの各プライマーを有する。また、 Mg^{2+} と $(NH_4)_2SO_4$ と2-メルカプトエタノールとを含有する反応バッファーと、Taq DNAポリメラーゼ (Amersham Pharmacia Biotech) と、ELONGASE酵素 (Life Technologies) と、Pfu DNAポリメラーゼ (Stratagene) とを含有する。プライマー対であるPCI AとPCI Bとに対し、以下のパラメータで増幅を行った。

30

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

別法では、プライマー対であるT7とSK+に対して以下のパラメータで増幅を行った。

40

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

各ウェルのDNA濃度は、 $1 \times TE$ および $0.5 \mu l$ の希釈していないPCR産物に溶解した $100 \mu l$ のPICOGREEN定量試薬 (0.25 (v/v) PICOGREEN; Molecular Probes, Eugene OR) を不透明な蛍光光度計ブ

50

レート (Corning Costar, Acton MA) の各ウェルに分配してDNAが試薬と結合できるようにして測定する。サンプルの蛍光を計測してDNAの濃度を定量すべく、プレートはFluoroskan II (Lab Systems Oy, Helsinki, Finland) でスキャンした。反応混合物のアリコット5~10 μ lを1%アガロースゲル上で電気泳動法によって解析し、どの反応が配列の伸長に成功したかを判定した。

【0296】

伸長したヌクレオチドは、脱塩および濃縮して384穴プレートに移し、CviJIコラウイルスエンドヌクレアーゼ (Molecular Biology Research, Madison WI) を用いて消化し、音波処理またはせん断し、pUC 18ベクター (Amersham Pharmacia Biotech) への再連結を行った。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で分離し、断片を切除し、寒天をAgar ACE (Promega) で消化した。伸長させたクローンをT4リガーゼ (New England Biolabs, Beverly MA) を用いてpUC 18ベクター (Amersham Pharmacia Biotech) に再連結し、Pfu DNAポリメラーゼ (Stratagene) で処理して制限部位のオーバーハングを満たし、適格な大腸菌細胞に形質移入した。形質移入した細胞群の選択を抗生物質を含む培地で行い、それぞれのコロニーを採取し、LB/2Xカルベニシリン培養液内の384ウェルプレート群に37℃で一晩培養した。

【0297】

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

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

DNAの定量化には、上記したようにPICOGREEN試薬 (Molecular Probes) を用いた。DNAの回収率が低いサンプルは、上記と同一の条件を用いて再増幅した。サンプルは20%ジメチルスルホキシド (1:2, v/v) で希釈し、DYENAMIC エネルギートランスファー シークエンシングプライマー、およびDYENAMIC DIRECT kit (Amersham Pharmacia Biotech) またはABI PRISM BIGDYE ターミネーターサイクル シークエンシングレディ反応キット (Terminator cycle sequencing ready reaction kit) (Applied Biosystems) を用いてシーケンシングした。

【0298】

同様に、上記手順を用いて完全長ポリヌクレオチド配列を検証した。あるいは、完全長ポリヌクレオチドを用い、上記手順で、そのような伸長のために設計したオリゴヌクレオチド類と、或る適切なゲノムライブラリとを用いて5'調節配列を得た。

【0299】

9 個々のハイブリダイゼーションプローブの標識および使用

SEQ ID NO: 18-34由来のハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記載するが、より大きなヌクレオチド断片に対しても本質的に同一の手順を用いる。オリゴヌクレオチドは、OLIGO 4.06ソフトウェア

10

20

30

40

50

(National Biosciences)等の最新ソフトウェアを用いて設計し、各オリゴマー50 pmolと、 $[\gamma\text{-}^{32}\text{P}]$ アデノシン3リン酸 (Amersham Pharmacia Biotech) 250 μCi と、T4ポリヌクレオチドキナーゼ (DuPont NEN, Boston MA)を混合することにより標識する。標識したオリゴヌクレオチドは、SEPHADEX G-25超細繊維分子サイズ排除デキストラン ビーズカラム (Amersham Pharmacia Biotech)を用いて実質的に精製する。Ase I、Bgl II、Eco RI、Pst I、Xba IまたはPvu II (DuPont NEN)のいずれか1つのエンドヌクレアーゼで消化されたヒトゲノムDNAの、典型的な膜ベースのハイブリダイゼーション解析において、毎分107カウントの標識されたプローブを含むアリコットを用いる。

10

【0300】

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

【0301】

10 マイクロアレイ

20

或るマイクロアレイ上でのアレイエレメント群の結合または合成は、フォトリソグラフィ、圧電印刷 (インクジェット印刷、前出のBaldeschweilerなどを参照)、機械的マイクロスポッティング技術およびこれらから派生した手段を用いて達成し得る。上記各技術において基板は、均一な、非多孔性の表面を持つ固体とすべきである (Schena (1999) 前出)。推奨する基板には、シリコン、シリカ、スライドガラス、ガラスチップおよびシリコンウエハがある。あるいは、ドットプロット法またはスロットプロット法に類似した手順を利用し、熱的、紫外線の、化学的または機械的結合手順を用いて基板の表面にエレメントを配置および結合させてもよい。通常のアレイは、利用可能な、当業者に公知の方法と機械とを用いて作製でき、任意の適正数のエレメントを有し得る (例えばSchena, M. 他 (1995) Science 270: 467-470、Shalon, D. 他 (1996) Genome Res. 6: 639-645、Marshall, A. および J. Hodgson (1998) Nat. Biotechnol. 16: 27-31を参照)。

30

【0302】

完全長cDNA、発現配列タグ (EST)、またはそれらの断片またはオリゴマーが、マイクロアレイのエレメントと成り得る。ハイブリダイゼーションに好適な断片またはオリゴマーを、LASERGENEソフトウェア (DNASTAR) など本技術分野で公知のソフトウェアを用いて選択することが可能である。アレイエレメントは、生物学的サンプル中でポリヌクレオチドを用いてハイブリダイズする。生物学的サンプル中のポリヌクレオチドは、検出を容易にするために蛍光標識その他の分子タグに抱合させる。ハイブリダイゼーション後、ハイブリダイズされていないヌクレオチドを生物学的サンプルから除去し、蛍光スキャナを用いて各アレイエレメントでのハイブリダイゼーションを検出する。あるいは、レーザ脱離および質量スペクトロメトリを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上の或るエレメントにハイブリダイズする各ポリヌクレオチドの、相補性の度合と相対存在度とを算定し得る。一実施態様におけるマイクロアレイの調整および使用について、以下に詳述する。

40

【0303】

組織または細胞サンプルの調製

全RNAを組織サンプル群から単離するためグアニジニウム チオシアネート法を用い、ポリ (A)⁺ RNAを精製するためオリゴ (dT)セルロース法を用いる。各ポリ (A)

50

⁺ RNAサンプルを、MMLV逆転写酵素、0.05 pg/μlのオリゴ(dT)プライマー(21mer)、1×第一鎖バッファー、0.03 unit/μlのRNアーゼ阻害因子、500 μMのdATP、500 μMのdGTP、500 μMのdTTP、40 μMのdCTP、40 μMのdCTP-Cy3(BDS)またはdCTP-Cy5(Amersham Pharmacia Biotech)を用いて逆転写する。逆転写反応は、GEMBRIGHTキット(Incyte)を用い、200 ngのポリ(A)⁺ RNAを含む体積25 mlで行う。特異的対照ポリ(A)⁺ RNA群は、非コード酵母ゲノムDNAから*in vitro*転写により合成する。37℃で2時間インキュベートした後、各反応サンプル(1つはCy3、もう1つはCy5標識)は、2.5 mlの0.5 M水酸化ナトリウムで処理し、85℃で20分間インキュベートし、反応を停止させてRNAを分解させる。サンプルを精製するため、2つの連続するCHROMA SPIN 30ゲル濾過スピナラム(CLONTECH Laboratories (CLONTECH), Palo Alto CA)を用いる。混合後、2つの反応サンプルをエタノール沈殿させるため、1 mlのグリコーゲン(1 mg/ml)、60 mlの酢酸ナトリウム、および300 mlの100%エタノールを用いる。サンプルを次にSpeed VAC(Savant Instruments, Holbrook NY)を用いて完全に乾燥させ、14 μlの5×SSC/0.2% SDS中で再懸濁する。

10

【0304】

マイクロアレイの調製

本発明の配列を用いて、アレイエレメント群を作製する。各アレイエレメントは、クローン化したcDNAインサート群により、ベクター含有細菌細胞群から増幅する。PCR増幅は、cDNAインサートの側面に位置するベクター配列群に相補的なプライマー類を用いる。30サイクルのPCRによって、1~2 ngの初期量から5 μgを超える最終量まで、アレイエレメントを増幅する。増幅したアレイエレメントは、SEPHACRYL-400(Amersham Pharmacia Biotech)を用いて精製する。

20

【0305】

精製したアレイエレメントは、ポリマーコートされたスライドガラス上に固定する。顕微鏡スライドガラス(Corning)は、0.1%のSDSおよびアセトン中で超音波洗浄し、処理中および処理後に多量の蒸留水で洗浄する。スライドガラスは、4%フッ化水素酸(VWR Scientific Products Corporation (VWR), West Chester PA)中でエッチングし、多量の蒸留水中で洗浄し、95%エタノール中で0.05%アミノプロピルシラン(Sigma)を用いてコーティングする。コーティングしたスライドガラスは、110℃のオーブンで硬化させる。

30

【0306】

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。この特許は引用することを以って本明細書の一部とする。平均濃度100 ng/μlのアレイエレメントDNA 1 μlを、高速ロボット装置(robotic apparatus)により、開放型キャピラリープリンティングエレメント(open capillary printing element)に充填する。装置はここで、スライド毎に約5 nlのアレイエレメントサンプルを置く。

40

【0307】

マイクロアレイをUV架橋するため、STRATALINKER UV架橋剤(Stratagene)を用いる。マイクロアレイの洗浄を、室温において、0.2% SDSで1回、蒸留水で3回行う。各非特異結合部位をブロックするため、マイクロアレイのインキュベーションをリン酸緩衝生理食塩水(PBS)(Tropix, Bedford MA)中の0.2%カゼイン中において60℃で30分間行った後、前に行ったように0.2% SDSおよび蒸留水で洗浄する。

【0308】

ハイブリダイゼーション

ハイブリダイゼーション反応に用いる9 μlのサンプル混合体には、Cy3またはCy5

50

で標識したcDNA合成産物群の各0.2μgを、5×SSC、0.2%SDSハイブリダイゼーション緩衝液中に含む。サンプル混合体は、65℃まで5分間加熱し、マイクロアレイ表面上で等分して1.8cm²のカバーガラスで覆う。アレイを、顕微鏡用スライドよりわずかに大きい空洞を有する防水チェンバーに移す。チェンバー内を湿度100に保つため、140μlの5×SSCをチェンバーの1コーナーに加える。アレイを含むチェンバーは、60℃で約6.5時間インキュベートする。アレイは、第1洗浄緩衝液中(1×SSC、0.1%SDS)において45℃で10分間洗浄し、第2洗浄緩衝液中(0.1×SSC)において45℃で10分間ずつ3回洗浄して乾燥させる。

【0309】

検出

レポーター標識したハイブリダイゼーション複合体を検出するには、Cy3の励起のために488nm、Cy5の励起のために632nmでスペクトル線を発生し得るInnova70混合ガス10Wレーザ(Coherent, Santa Clara CA)を備えた顕微鏡を用いる。励起レーザ光の焦点をアレイ上に置くため、20×顕微鏡対物レンズ(Nikon, Melville NY)を用いる。アレイを含むスライドを、顕微鏡の、コンピュータ制御のX-Yステージに置き、対物レンズを通してラスタースキャンする。本実施例で用いる1.8cm×1.8cmのアレイは、解像度20μmでスキャンする。

【0310】

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

【0311】

通常、各スキャンの感度を較正するため、cDNA対照種を或る既知濃度でサンプル混合体に添加し、対照種が発生するシグナル強度を用いる。アレイ上の或る特定の位置には或る相補的DNA配列が含まれ、その位置におけるシグナルの強度をハイブリダイゼーション種の重量比1:100, 000で相関させる。異なる源泉(例えば代表的な試験細胞および対照細胞など)からの2つのサンプルを、各々異なる蛍光色素で標識し、異なる発現をする遺伝子群を同定するために単一のアレイにハイブリダイズする場合には、較正するcDNAのサンプルを2種の蛍光色素で標識し、ハイブリダイゼーション混合液に各々等量を加えることによって較正を行う。

【0312】

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

【0313】

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

10

20

30

40

50

【0314】

1.1 相補的ポリヌクレオチド

TMPをコードする配列群或いはその任意の部分に対する相補配列群を用いることにより、天然TMPの発現が検出、低減、または阻害される。約15～30塩基対を含むオリゴヌクレオチドの使用について記すが、これより小さなあるいは大きな配列の断片の場合でも、本質的に同じ手順を用いる。適切なオリゴヌクレオチド群を設計するため、Oligo 4.06ソフトウェア(National Biosciences)および、TMPのコーディング配列を用いる。転写を阻害するためには、最も独特な5'配列から相補的オリゴヌクレオチドを設計して用いることにより、プロモーターがコーディング配列に結合するのを防止する。翻訳を阻害するには、相補的なオリゴヌクレオチドを設計することにより、リボソームが、TMPをコードする転写物に結合するのを防ぐ。

10

【0315】

1.2 TMPの発現

TMPの発現および精製は、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。細菌内でTMPが発現するよう、cDNAを好適なベクターにサブクローニングする。ベクターは、抗生物質耐性遺伝子と、cDNA転写レベルを高める誘導性のプロモーターとを持つものを用いる。このようなプロモーターの例には、限定するものではないが、lacオペレーター調節エレメントに関連するT5またはT7バクテリオファージプロモーターおよびtrp-lac(tac)ハイブリッドプロモーターが含まれる。組換えベクターを、BL21(DE3)などの好適な細菌宿主に形質転換する。抗生物質耐性細菌にTMPを発現させるには、イソプロピルβ-Dチオガラクトピラノシド(IPTG)で誘発する。真核細胞でのTMP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るAutographica californica核多角体病ウイルス(AcMNPV)の組換え型を感染させて行う。バキュロウイルスの非必須ポリヘドリン遺伝子をTMPをコードするcDNAと置換するには、相同組換えを行うか、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を行う。ウイルスの感染力は維持され、強力な多角体プロモーターによって高レベルのcDNA転写が行われる。組換えバキュロウイルスは、多くの場合は夜蛾の1種Spodoptera frugiperda(Sf9)昆虫細胞への感染に用いるが、ヒト肝細胞への感染に用いることもある。後者の感染の場合は、バキュロウイルスへの更なる遺伝的修飾が必要になる(Engelhardt, E. K. 他(1994)Proc. Natl. Acad. Sci. USA 91:3224-3227, Sandig, V. 他(1996)Hum. Gene Ther. 7:1937-1945を参照)。

20

30

【0316】

殆どの発現系では、TMP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を、迅速に1回で行うことができる。GSTは日本住血吸虫からの26kDaの酵素であり、タンパク質の活性および抗原性を維持した状態で、固定化したグルタチオン上での融合タンパク質の精製を可能とする(Amersham Pharmacia Biotech)。精製後、GST要素を、操作した特定の部位においてTMPからタンパク質分解的に切断することが可能である。FLAGは8アミノ酸のペプチドであり、市販されているモノクローナルおよびポリクローナル抗FLAG抗体(Eastman Kodak)を用いた免疫親和性精製を可能にする。6ヒスチジン残基が連続して伸長した6-Hisは、金属キレート樹脂(QIAGEN)上での精製を可能にする。タンパク質の発現および精製の方法は、前出のAusubel(1995)10章、16章に記載されている。これらの方法で精製したTMPを直接用いて、適用可能な場合は以下の実施例16、および17のアッセイを行うことができる。

40

【0317】

1.3 機能的アッセイ

50

TMPの機能は、哺乳動物細胞培養系において生理的に高められたレベルでの、TMPをコードする配列の発現によって算定する。cDNAを高いレベルで発現する強いプロモーターを持つ哺乳動物発現ベクターにcDNAをサブクロニングする。選択されるベクターの例には、pCMV SPORT (Life Technologies) および pCR3.1 (Invitrogen, Carlsbad CA) があり、どちらもサイトメガロウイルスプロモーターを持つ。リポソーム製剤あるいは電気穿孔法を用いて、5～10 μ gの組換えベクターをヒト細胞株、例えば内皮由来または造血由来の細胞株に、一過的に形質移入する。更に、標識タンパク質をコードする配列を含む1～2 μ gのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入細胞と非形質移入細胞とを区別する手段が与えられる。また、標識タンパク質の発現によって、組換えベクターからのcDNA発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質 (GFP; Clontech)、CD64またはCD64-GFP融合タンパク質から選択できる。自動化された、レーザー光学に基づく技術であるフローサイトメトリー (FCM) を用いて、GFPまたはCD64-GFPを発現する形質移入された細胞を同定し、それらの細胞のアポトーシス状態や他の細胞特性を評価する。FCMは、細胞死に先行するかあるいは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、プロピジウムヨウ化物によるDNA染色によって計測される核DNA内容物の変化、前方散乱光と90°側方散乱光によって計測される細胞サイズと顆粒性の変化、プロモデオキシウリジンの取込量の低下によって計測されるDNA合成の下方調節、特異抗体との反応性によって計測される細胞表面および細胞内におけるタンパク質の発現の変容、および蛍光複合アネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変容とがある。フローサイトメトリー法については、Ormerod, M. G. (1994) Flow Cytometry, Oxford, New York, NYに記述がある。

10

20

30

40

50

【0318】

遺伝子発現におけるTMPの影響は、TMPをコードする配列と、CD64またはCD64-GFPのどちらかまたは両方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に対する抗体かのどちらかで被覆された磁気ビーズを用いて効率的に分離することができる (DYNAL, Lake Success NY)。mRNAは、当業者に周知の方法で細胞から精製することができる。TMPと、目的とする他の遺伝子とをコードする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。

【0319】

1.4 TMPに特異的な抗体の作製

実質的に精製されたTMPを、ポリアクリルアミドゲル電気泳動法 (PAGE; 例えば、Harrington, M. G. (1990) Methods Enzymol. 182: 488-495を参照) または他の精製技術で作製し、これ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。

【0320】

別法では、TMP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)。

【0321】

通常は、長さ約15残基のオリゴペプチドを、FMOCケミストリを用いるABI 431A ペプチドシンセサイザ (Applied Biosystems) を用いて合成し、N-マレイミドベンゾイル-N-ヒドロキシスクシンイミドエステル (MBS) を用いた

反応によってKLH (Sigma-Aldrich, St. Louis MO) に結合させて、免疫原性を高める (例えば前出のAusubel, 1995を参照)。完全フロイントアジュバントにおいて、オリゴペプチド-KLH複合体を用いてウサギを免疫化する。得られた抗血清の抗ペプチド活性および抗TMP活性を検査するには例えば、ペプチドまたはTMP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、放射性ヨウ素標識したヤギ抗ウサギIgGと反応させる。

【0322】

15 特異的抗体を用いる天然TMPの精製

天然TMPあるいは組換えTMPを実質的に精製するため、TMPに特異的な抗体類を用いるイムノアフィニティー (免疫親和性) クロマトグラフィーを行う。イムノアフィニティーカラムは、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TMP抗体とを共有結合させることにより形成する。結合後に、製造者の使用説明書に従ってこのレジンをブロックし、洗浄する。

【0323】

TMPを含む培養液をイムノアフィニティーカラムに通し、TMP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TMP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イオンのようなカオトロップで) 溶出させ、TMPを回収する。

【0324】

16 TMPと相互作用する分子の同定

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

【0325】

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

【0326】

TMP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PATHC ALLING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)。

【0327】

17 TMP活性の実証

TMPのギャップ結合活性

TMPのギャップ結合活性を実証するには、細胞間チャネルの形成を誘発する能力を、1対のアフリカツメガエル卵母細胞にTMPのcRNAを注入して調べる (Hennemann, 前出)。TMP cRNAを実験注入する1週間前に、卵母細胞に、TMPに対するアンチセンスオリゴヌクレオチドを注入してバックグラウンドを低減する。TMP cRNA注入卵母細胞のインキュベートを1晩行い、卵黄膜を剥ぎ、対にして、接合部電流の記

10

20

30

40

50

録を、2細胞電圧固定で行う。測定したコンダクタンスは、TMPのギャップ結合活性に比例する。

【0328】

別法では、TMP活性のアッセイのための、TMPのイオンチャネル活性の測定に、イオンコンダクタンスについての電気生理学的アッセイを用いる。哺乳動物細胞株、(例えばCOS7、HeLa、若しくはCHOなど)を、TMPをコードする真核細胞発現ベクターで形質転換させて、TMPを発現し得る。真核細胞発現ベクターは市販されており、それらを細胞内に導入する技術は当業者には周知である。 β -ガラクトシダーゼなど多数のマーカー遺伝子のいずれか1つを発現させる第二のプラスミドをこれら細胞へと同時形質転換することにより、外来DNAを取り込んで発現させた細胞群の迅速な同定が可能となる。形質転換の後、この細胞株がTMPおよび β -ガラクトシダーゼを発現し蓄積するのに適した条件下でこれらの細胞を48~72時間、インキュベーションする。

10

【0329】

β -ガラクトシダーゼを発現する形質転換細胞は、好適な比色用基質が本技術分野で公知の条件下で培地へ添加されると、青く染色される。染色した細胞は、本技術分野で公知の電気生理技術を用いて膜コンダクタンスの違いを試験する。形質転換していない細胞、および/または、形質転換に、ベクター配列群のみ、または β -ガラクトシダーゼ配列群のみを用いた細胞を対照として用い、平行して試験する。TMPを発現する細胞は、対照細胞に比べてより高いアニオンコンダクタンスまたはカチオンコンダクタンスを持つこととなる。コンダクタンスに対するTMPの寄与は、TMPに特異的な抗体類を用いて細胞をインキュベートすることで確認し得る。抗体は、TMPの細胞外側に結合することにより、イオンチャネル内のポアと、それに伴うコンダクタンスとをブロックする。

20

【0330】

TMPの膜貫通タンパク質活性

TMP活性に対する或るアッセイは、細胞表面におけるTMPの発現を測定する。TMPをコードするcDNAを、好適な哺乳動物細胞株に形質移入する。細胞表面タンパク質を、記載されているようにビオチンで標識する(de la Fuente, M. A. 他(1997) Blood 90:2398-2405)。TMP特異抗体を用いて免疫沈降を行い、SDS-PAGEおよび免疫ブロット技術を用いて免疫沈降サンプルを分析する。標識された免疫沈降素と非標識免疫沈降素との比は、細胞表面に発現したTMPの量に比例する。

30

【0331】

TMP活性の或る別のアッセイは、リガンド/受容体が媒介する、細胞増殖の変調についての原型のアッセイに基づいている。このアッセイは、TMPを発現するスィスマウス3T3細胞において新規に合成されたDNAの量を測定する。TMPをコードするcDNAを持つ好適な哺乳動物発現ベクターを、当分野で周知のトランスフェクション法を用いて静止状態の3T3培養細胞に加える。形質転換細胞を、 $[^3\text{H}]$ チミジンと可変量のTMPリガンドとの存在下でインキュベートする。酸沈殿性DNA内へ取り込まれた $[^3\text{H}]$ チミジンを、トリチウムラジオアイソトープカウンタを用いて適当な時間間隔で測定する。取り込まれた量は、新規に合成されたDNAの量に正比例する。少なくとも100倍のTMPリガンド濃度範囲に対する用量反応曲線がリニアであることは、受容体活性を示唆するものである。ミリリットル当りの1単位の活性の定義は、50%の反応レベルを産出するTMPの濃度であり、100%であれば、 $[^3\text{H}]$ チミジンを酸沈殿性DNAに最大限、取り込むことを表す(McKay, I. および Leigh, I., eds. (1993) Growth Factors: A Practical Approach, Oxford University Press, New York NY, 73ページ)。

40

【0332】

当業者は、本発明の範囲および精神から逸脱することなく、本発明の記載した方法およびシステムの種々の修正を行い得る。本発明について説明するにあたり幾つかの実施態様に

50

関連して説明を行ったが、本発明の請求の範囲が、そのような特定の実施態様に不当に制限されるべきではないことを理解されたい。分子生物学または関連分野の専門家には明らかな、本明細書に記載する本発明の実施方法の様々な修正は、明確に特許請求の範囲内にあるものとする。

【0333】

(表の簡単な説明)

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

【0334】

表2は、GenBank識別番号と、本発明のポリペプチドに最も近いGenBank相同体の注釈とを示す。また、各ポリペプチドとその相同体(1つ以上)が一致する確率スコアも併せて示す。 10

【0335】

表3は、予測されたモチーフおよびドメインを持つ本発明のポリヌクレオチド配列の構造的特徴を、ポリペプチドの分析に用いる方法、アルゴリズムおよび検索可能なデータベースと共に示す。

【0336】

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

【0337】

表5は、本発明のポリヌクレオチドの代表的なcDNAライブラリを示す。 20

【0338】

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

【0339】

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

【表1】

表 1

IncYTE プロジェクト ID	ポリペプチド SEQ ID NO:	IncYTE ポリペプチド ID	ポリスクレオチド SEQ ID NO:	IncYTE ポリスクレオチド ID
6431478	1	6431478CD1	18	6431478CB1
3584654	2	3584654CD1	19	3584654CB1
3737084	3	3737084CD1	20	3737084CB1
71426238	4	71426238CD1	21	71426238CB1
7475123	5	7475123CD1	22	7475123CB1
7481952	6	7481952CD1	23	7481952CB1
382654	7	382654CD1	24	382654CB1
1867351	8	1867351CD1	25	1867351CB1
3323104	9	3323104CD1	26	3323104CB1
4769306	10	4769306CD1	27	4769306CB1
2720058	11	2720058CD1	28	2720058CB1
7481255	12	7481255CD1	29	7481255CB1
1510242	13	1510242CD1	30	1510242CB1
162131	14	162131CD1	31	162131CB1
1837725	15	1837725CD1	32	1837725CB1
3643847	16	3643847CD1	33	3643847CB1
6889872	17	6889872CD1	34	6889872CB1

【表 2】

表 2-1

ポリペプチド SEQ ID NO:	Incyte ポリペプ チド ID	GenBank ID NO:	標準スコア	GenBank 相同体
1	6431478CD1	G310100	4.0E-216	[ラット]発達に応じて調節されるタンパク質
2	3584654CD1	G1054884	0.0	[ラット]プロスタグランジン F2a 受容体関連タンパク質 Orlicky, D.J. (1996) Negative regulatory activity of a prostaglandin F2a receptor associated protein (FPRP). Prostaglandins Leukot. Essent. Fatty Acids 54, 247-259
3	3737084CD1	G3513451	1.7E-233	[ラット] カリウムチャネル調節因子 1 Hoshi, N., et al (1998) KCRI, a membrane protein that facilitates functional expression of non-inactivating K+ currents associates with rat EAG voltage-dependent K+ channels. J. Biol. Chem. 273, 23080-23085
4	71426238CD1	G15990851	1.0E-130	[ヒト] MEGF7 (多量 EGF 様 7 タンパク質)
5	7475123CD1	G3449306	0.0	Nakayama, M. 他, (1998) Genomics 51:27-34
6	7481952CD1	G8926243	1.5E-40	[マウス] 低密度リポタンパク質受容体関連タンパク質 LRP1B/LRP-DIT

10

20

30

【表 3】

表 2-2

ポリペプチド SEQ ID NO:	Incyte ポリペプ チド ID	GenBank ID NO:	確率スコア	GenBank 相合体
9	3323104CD1	g11414879	0.0	[f][ヒト] マンノシルトランスフェラーゼ Maeda, Y. 他 (2001) FIG-M transfers the first mannose to glycosylphosphatidylinositol on the luminal side of the ER. EMBO J. 20:250-261
11	2720058CD1	g6013381	1.5E-12	TM6P1 (膜内在性タンパク質) [ラット] (Zhang, J. 他 (2000) Biochim. Biophys. Acta 1492:280-284)
12	7481255CD1	g4150839	4.2E-27	[マウス] GSG1
13	1510242CD1	g4263743	5.3E-163	UNC-93 に類似; U09424 に類似 (PID:g3642687)
14	162131CD1	g13374351	0.0	[f][ヒト] アポリポタンパク質 L4
15	1837725CD1	g6850311	6.0E-52	[シロイヌナズナ]からの液胞選別受容体相合体 との類似性を持つ gb U79959
16	3643847CD1	g9944332	2.8E-95	[マウス] Ttyhl Campbell, H.D. 他 (2000) Genomics 58:89-92
17	6869872CD1	g3191975	1.3E-95	dJ63G5.3 (推定上のロイシンリッチタンパク 質) [ヒト]

10

20

30

【表 4】

【表 5】

SEQ ID NO.	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコ ル化部位	シグネチャ配列、 ドメインおよびモチーフ	分析方法及びデ ータベース
1	6431478CD1	461	S117 S190 S225 S280 S336 S34 T109 T221 T333 T386	N113 N193 N294	膜貫通ドメイン： N38-T56, V94-T112, Y156-V174, Y195- M216, I231-I248, Y388-M406, タンパク質 胎盤性 DIFFF33 発生的に調節される R11HG.2 PD011773: I76-P283, K192-S460 タンパク質 胎盤性 DIFFF33 R11HG.2 発生的に調節さ れる PD018175: M1-E70	HMMER BLAST_PRODOM BLAST_PRODOM
2	3584654CD1	879	S150 S158 S244 S288 S349 S46 S574 S580 S598 S623 S66 S72 S747 S758 S772 T303 T402 T472 T505 T536 T570 T757 T820 T866 T97 Y640	N286 N300 N383 N413 N44 N525 N600 N618 N691	signal_cleave: M1-G21 signal_peptide: M1-R25 膜貫通ドメイン: I833-0851 免疫グロブリンドメイン： G36-T121, G162-V249, G292-V375, D422-V517, G564-V657, G704-V795 タンパク質 プロスタグランジン F2α 受容体調節前駆体関 連シグナル 免疫グロブリンフォールド PD025644: L9-S814 プロスタグランジン F2α 受容体調節タンパク質前駆体関 連シグナル 免疫グロブリンフォールド 膜貫通 糖タンパ ク質 PD185779: K815-D879 免疫グロブリン DM000001 I39207 25-139:V26-Y132 DM000001 I39207 141-259:V140-E255 RGD モチーフ: R703-D705	SPSCAN HMMER HMMER HMMER_FAM BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM
3	3737094CD1	473	S243 S280 S383 S387 S56 T368 T434 Y417 Y444	N246 N457	膜貫通ドメイン： M95-Y112, P141-M166, Y257-G277, N396-I419	HMMER MOTIFS

表 3 - 2

SEQ ID NO.	Incyte ID/	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
4	71426238CD1	223	T132 T174		タンパク質 チャネル DIE2 カリウム 調節因子 イオン性 C56F8.06C 染色体 I 前駆体 PD025171: E33-S387, W395-N428, S435-W473	BLAST_PRODOM
					カリウムチャネル調節因子 1 イオンチャネル PD184319:M1-R32	BLAST_PRODOM
					コネキシン:M1-L201	HMMER_Pfam
					コネキシン DM00590 Q02739 1-268:M1-P218 P28231 1-269:M1-K208 Q02738 1-262:M1-L206 P08034 1-276:M1-K204	BLAST_DOMO
5	7475123CD1	1553	S1007 S1011	N1063 N1115	ギャップ結合 コネキシン タンパク質 膜貫通 ALPHA1 CX43 ALPHA8 ALPHA5 BETA1 PD001135:W3-K108, M16-L201	BLAST_PRODOM
					ギャップ結合タンパク質 コネキシン 膜貫通 BETA2 CX26 BETA4 CX31.1 疾患 PD001118:G119-L201	BLAST_PRODOM
					コネキシン タンパク質 BL00407:S2-V38, A39-F69, P70-Y97, G118-D147, C157-L201	BLIMPS_BLOCKS
					コネキシン シグネチャ PR00206:I20-W44, F51-Q73, L76-A96, L120-Y146, C157-T177, I178-L201	BLIMPS_PRINTS
					膜貫通ドメイン:I23-A39, L120-Y143, T177-F200	HMMER
					コネキシン シグネチャ 2:C157-P173	MOTIFS
5	7475123CD1	1553	S1007 S1011	N1063 N1115	コネキシンシグネチャ connexins 1:L33-T86	PROFILESCAN
					コネキシンシグネチャ connexins 2:I136-L193	PROFILESCAN
5	7475123CD1	1553	S1007 S1011	N1063 N1115	signal_cleavage:M1-A20	SPSCAN

【表 6】

表 3-3

SEQ ID NO:	Incyte ID/	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
		S1016 S1065 S1103 S1108 S1117 S1152 S1234 S1313 S1337 S1346 S1416 S1420 S1465 S1479 S149 S1504 S1535 S1548 S177 S21 S287 S366 S390 S404 S430 S496 S529 S540 S562 S583 S61 S72 S761 S804 S816 S843 S848 T1095 T1174 T1211 T1331 T1336 T1508 T384 T389 T453 T461 T475 T485 T623 T629 T727 T793 T880 T907	N1409 N1449 N549 N725	LDL 受容体リガンド結合リピート DM00045 P01130 37-111:C44-D113, G81-S149, G120-P183 LDL 受容体リガンド結合リピート DM00045 P98160 295-336:G81-C122 LDL 受容体リガンド結合リピート DM00045 I48623 82-120:C83-C122 糖タンパク質 タンパク質受容体 EGF 様ドメイン リポタ ンパク質前駆体シグナル 膜貫通 受容体関連 PD149641:R431-D559, D1044-D1171, Q739-D867 低密度リポタンパク質受容体関連タンパク質前駆体 LRP 受容体膜貫通リピート エンドサイトーシス 糖タンパク 質シグナル カルシウム結合 EGF 様ドメイン 被覆ピッチ PD126644:V440-I707, I753-Y1040, I1057-C1287 仮説上 294.4 KD タンパク質 仮説上タンパク質 PD126659:C661-I1015 LDL 受容体クラス A (LDLRA) ドメイン タンパク質 BL01209:C90-E102 低密度リポタンパク質 PR00261:G81-E102, C42- E63, G158-E179, G201-E222 シグナルペプチド: M1-G18 膜貫通ドメイン: T1372-A1390	BLAST_DOMO BLAST_DOMO BLAST_DOMO BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLIMPS_BLOCKS BLIMPS_PRINTS HMMER HMMER	

10

20

30

【表 7】

表 3-4

SEQ ID NO.	Incyte ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
5					低密度リポタンパク質受容体リポドメイン: D433-V474 N476-W517 G519-Q561 Q563-E604 R605-E645 G741-I782 R784-E825 S827-A869 S871-S911 Y912-A953 G1045-V1086 R1088-R1129 G1131-D1173 R1175-T1220	HMMER_Pfam
5					低密度リポタンパク質受容体ドメイン: P23-L68, P69-P107, Q237-W275, P189-E227, R146-S184, R108-M145.	HMMER_Pfam
6	748195201	1718	S1027 S1031 S1095 S1111 S1165 S1273 S1310 S1463 S1482 S1510 S1517 S1544 S1583 S1686 S200 S214 S240 S274 S28 S366 S553 S565 S580 S812 S866 S877 S938 T1178 T1238 T133 T1359 T1371 T1391 T468 T504 T597 T613 T676 T763 T827	N1235 N126 N1299 N1345 N1545 N1634 N1684 N183 N260 N275 N496 N551 N611 N703 N761 N910 N969 N976	膜貫通ドメイン: I1642-A1660 低密度リポタンパク質受容体ドメイン: E383-P423, H610-M649, G1043-S1081, S1463-L1502, P1507-N1545, K1546-E1586, R824-T864, E1244-H1283 MAM ドメイン C38-T199, C216-L378, C427-L586, C652-P818, C869-S1027, C1083-T1238, C1291-T1454 LDL 受容体クラス A BL01209: C1485-E1497 MAM ドメインタンパク質: BL00740A: C434-W446, BL00740B: R1011-S1031 低密度リポタンパク質受容体 PRO0261: L1476-E1497, R397-E418, K836-E857 MAM ドメインタンパク質: DMO1344 P98072 352-509: E418-D578 DMO1344 P28824 595-796: E424-D577 DMO1344 A55620 961-1128: C421-L576	HMMER HMMER_Pfam HMMER_Pfam BLIMPS-BLOCKS BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-DMO

【表 8】

表 3-5

SEQ ID NO:	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシ 化部位	シグネチャ配列、 ドメインおよびモチーフ	分析方法及びデータベース
7	382654CD1	224	S139 S165 S48 S74 T179 T202		細胞接着配列 (RGD): R1011-D1013, R1138-D1140 膜貫通ドメイン: I51-W71	MOTIFS HMMER
8	1867351CD1	570	S317 S336 S356 S538 S94 T367	N82	signal peptide: M1-C63 膜貫通ドメイン: L38-I58, A302-N322, V338-C357, I414-M436, I438-I457, L542-R564 特散付けられていない膜タンパク質ファミリー UPF0013: G43-L204, I264-A426 膜内タンパク質 PD004336: V344-W462 膜内タンパク質 PD149928: I55-P129	SPScan HMMER HMMER-PFAM BLAST-PRODOM BLAST-PRODOM
9	3323104CD1	423	S141 S203 S3 T263 T417 T48 T73 Y52	N400	ロイシンジッパーモチーフ: L537-L558 膜貫通ドメイン: L228-Y247, F386-I405 ロイシンジッパーモチーフ: L215-L236 遺伝子閉鎖域膜貫通タンパク質: PD040574: W131-I408	MOTIFS HMMER MOTIFS BLAST-PRODOM
10	4769306CD1	388	S141 S188 S232 S279 S294 S303 S357 S361 S377 S4 S96 S97 T187 T252 T385 T50	N161 N195 N301 N336 N384	膜貫通ドメイン: V200-I220 CUB ドメイン: C27-F139 低密度リポタンパク質受容体ドメイン: F145-E183 LDL 受容体クラス A BL01209: C166-E178	HMMER HMMER-PFAM HMMER-PFAM BLMP5-BLOCKS

10

20

30

【表 9】

表 3-6

SEQ ID NO:	Incyte ID/	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
11	2720068CD1	231	S137 T177 T220 T42 Y190	N54 N75	signal_peptide:MI-A14 膜貫通ドメイン: F6-S26, V95-H115, I196-D215	SPScan HAMMER
12	7481255CD1	293	S109 S148 S251 S95 T236 T35	N243 N277 N59 N67	signal_peptide:MI-S28 signal_peptide:MI-S22 膜貫通ドメイン: I126-G144, W156-I184 PMP-22 / BMP / MP20 family BL01221: F68-N101, Y203-T229	HAMMER SPScan HAMMER
13	1510242CD1	526	S279 S479 T22 Y125	N180 N201 N378	膜貫通ドメイン: L146-L172, V217-L235, P333-W351	HAMMER
14	162131CD1	348	S45 S187 S344 T33 T54 T207 T246 T271 Y39 Y77		膜貫通ドメイン: I149-L175 アポリポタンパク質 L 前駆体 APOL 血漿 脂質輸送糖タンパク質シグナル DJ6802.1 PD042084:T9-A347	HAMMER BLAST_PRODOM
15	1837725CD1	520	S8 S54 S141 T37 T151 T157 T202 T331 T505	N35 N58 N66 N74 N116 N126 N149 N155	signal_cleavage:MI-A25 シグナルペプチド:MI-A28 膜貫通ドメイン: M172-G191, F218-L237, W243-Y261, V287-F306, L346-E371	SPSCAN HAMMER HAMMER
16	3643847CD1	534	S35 S40 S124 S133 S338 S512 S519 T143 T170 T339 T418 Y207	N31 N129 N283 N352	signal_cleavage:MI-G57 膜貫通ドメイン: L46-F62, L214-G233, L390-C408 TWEETY F42E11.2 タンパク質 PD043235:V20-D429 細胞接着配列 R164-D166	SPSCAN HAMMER BLAST_PRODOM MOTIFS
17	6889872CD1	820	S92 S160 S164	N54 N60 N85	signal_cleavage:MI-A22	SPSCAN

10

20

30

【表 10】

表 3 - 7

SEQ ID NO:	Incyte ポリペプチド ID/	アミノ酸 残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					シグナルペプチド：MI-A22	HMMER
					膜貫通ドメイン：T394-Y416	HMMER
					ロイシンリッチリピート：T56-G79, N80-S103, S104-S127, R128-P151, S152-A175	HMMER_PFAM

【表 1 1】

10

20

30

表 4 - 1

ポリヌクレオチド SEQ ID NO:	Incyte ID	ポリヌクレオチド 配列長	選択された断片	配列断片	5'位置	3'位置
18	6431478CB1	2653	1-606, 2109-2306	679882JJ1 (COLENO03) 6882379J1 (BRAHTR03) 7162277F8 (PLACNO01) 7757207J1 (SPINTUE01) 6109577F6 (MCLDTXT03) 7196064F8 (LUNGER04)	2076 1679 1427 418 870 1	2653 2342 2173 1133 1464 735
19	3584654CB1	3531	1-1782	71760758V1 71760788V1 71761160V1 7658349J1 (UTREDME06) 70815172V1 71764416V1 71754337V1 7183593H1 (BONREFC01) 7709623J1 (PANCNOE02) 6807068F6 (SKINOR01) 70409519D1	2094 2694 2577 1341 1481 2978 1965 1 895 293 1372	2675 3350 3331 1989 2064 3531 2648 603 1480 1156 1947
20	3737084CB1	2280	947-997, 1840-2280	70408370D1 70400676D1 2310625T6 (NGANNOT01) 71685057V1 71683716V1	1720 859 1282 609 1	2280 1371 1832 1268 748
21	71426238CB1	1104	1-115, 550-624, 198-513	1947421R6 (PITUNOT01)	1	508
				71433875V1	435	1104

10

20

30

表 4-2

ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID	配列長	選択された断片	配列断片	5'位置	3'位置
22	7475123CB1	4966	1-957, 1328-3163	70879893V1	4038	4682
				8124425H1 (HEAONOC01)	4309	4966
				8253556H1 (BRAHDIT10)	3948	4663
				6891485J1 (BRAHTR03)	2762	3423
				6885135H1 (BRAHTR03)	1	468
				593337R6 (BRAVUNT02)	3439	3968
				70673072V1	3303	3965
22				GBI.G7705145_17_12_25 27.regenscan ()	337	3773
23	7481952CB1	5401	4237-4630, 2834- 3834, 1-2283, 4757-4851	55153856J1	1947	2415
				GNN.G10120165_000001_1 006.edit	1	342
				6933885R6 (SINTMR02)	2518	3182
				6933885H1 (SINTMR02)	2339	3011
				72434542D1	4614	5401
				55153330J1	3010	3752
				899395H1 (BRSTTUT03)	1589	1891
				55153261H1	311	899
				71876387V1	4346	5105
				71932270V1	3716	4305
				55153293H1	364	1151
				55153285J1	886	1816
				72432008D1	4234	5099
				GNN.G9843576_000009_0 06	1678	2105

10

20

30

【表 1 3】

表 4-3

ポリヌクレオチド SEQ ID NO:	Incyte ID	ポリヌクレオチド 配列長	選択された断片	配列断片	5'位置	3'位置
24	382654CB1	1949	1-247, 1882-1949	7629905H1 (BRAFTUE03)	1	517
				6618345J2 (BRAUTDR03)	1282	1949
				6884626H1 (BRAHTDR03)	541	1060
				6952509H1 (BRAITDR02)	1035	1600
				6891983H1 (BRAITDR03)	468	1056
				6891983J1 (BRAITDR03)	1121	1671
25	1867351CB1	2133	987-1049, 1927-2133, 779-832	70571014V1	1520	2133
				70568556V1	1259	1924
				7261108H1 (UTREWC01)	646	1135
				8004078H1 (MUSCTDC01)	1	704
				70570876V1	1151	1828
				55026730J1 (ADMEDNV22)	692	1215
26	3323104CB1	2090	1-34, 415-1111	702759T6 (SYNORAT03)	1128	1767
				71220067V1	585	1119
				70833651V1	553	1114
				70769661V1	1061	1692
				8133288H1 (SCOMDIC01)	1	578
				70771538V1	1625	2090
27	4769306CB1	1618	1-208, 561-723	70954880V1	1126	1618
				3269667H1 (BRAINOT20)	1	244
				70955603V1	648	1272
				8103247H1 (MIXDDIE02)	229	797
				7076427H1 (BRAUTDR04)	49	677
				70953586V1	795	1368

10

20

30

【表 1 4】

表 4 - 4

ポリヌクレオチド SEQ ID NO:	Incyte ID	ポリヌクレオチド 配列長	選択された断片	配列断片	5'位置	3'位置
28	2720058CB1	3269	2536-2564, 1- 220, 2785-3269	2557501T6 (LUNGTT09)	2502	3056
				71520348V1	1839	2547
				71524606V1	1139	1824
				2657501F6 (LUNGTT09)	594	1126
				7225843H1 (LUNGTC01)	392	983
				71670642V1	1721	2424
				1305113F6 (PLACNOT02)	1	571
				1955363F6 (CONNOT01)	2570	3058
				1375665F6 (LUNGNOT10)	2677	3269
				71521430V1	1054	1714
29	7481255CB1	1227	1-1227	3269676H1 (BRAINT02)	5	96
				GBI_g7622436_000021_0	1	1227
				00024.edit		
				FL7481255_98469082_00	99	304
				0035_g4150939_1_1		
				4538535F6 (THYRINT01)	654	1227
30	1510242CB1	2618	1-515, 2475-2618	6823120J1 (SINTNOR01)	1889	2437
				7036758F7 (UTRSTNR02)	1	705
30				7355340H1 (HEARNON03)	2005	2618
				7930760H1 (COLNDIS02)	1224	1900
				6837255H1 (BRSTNON02)	673	1368
				70366002D1	1430	1904
				3413605H1 (PTHYNOT04)	2456	2618
				71808592V1	806	1433
31	162131CB1	2188	1-25, 1173-1633	6823120H1 (SINTNOR01)	1820	2436
				g1506355	1667	2188

10

20

30

表 4-5

ポリスクレオチド SEQ ID NO:	Incyte ID	ポリスクレオチド	配列長	選択された断片	配列断片	5'位置	3'位置
					7243158H2 (PROSTM01)	1351	1878
					1998635R6 (BRSTTUT03)	1632	2172
					70559145V1	454	1196
					70558921V1	1069	1787
					2618229F6 (BRSTNOT14)	1	540
					70558931V1	549	1305
32	1837725CB1		1969	1-400, 1927-1969	70377507D1	1400	1955
					91378655	1517	1969
					70378490D1	639	1268
					6799889J1 (COLENOR03)	1	673
					8227475H1 (BRAUTDR02)	778	1463
					2352377H1 (COLSUCT01)	1763	1959
33	3643847CB1		3006	796-1137, 2849-3006	1478479H1 (CORENOT02)	2745	3006
					6993355H1 (BRAQIDR02)	1129	1755
					5964485H1 (BRAINOT05)	1772	2489
					1419930H1 (KIDNNOT09)	2650	2892
					6789167H1 (BRACNOR01)	2091	2801
					6118619H1 (BRAHNON05)	1992	2546
					7436931H1 (ADRETUE02)	594	1202
					72018042V1	1	579
					72018222V1	413	1028
					1992275F6 (CORENOT02)	1268	1809

10

20

30

【表 16】

表 4－6

ポリヌクレオチド Seq ID NO:	Incyte ポリヌクレオチド ID	配列長	選択された断片	配列断片	5'位置	3'位置
34	6889872CB1	2884	1-319, 2090- 2884, 803-1585, 379-418, 1991- 2017	8230901H1 (BRAUTDR02)	873	1610
				8230892H1 (BRAUTDR02)	1	725
				8230892J1 (BRAUTDR02)	676	1356
				7634710H1 (SINTDIE01)	1577	2162
				GNN:g3191973_012	422	2884

【表 1 7】

10

20

30

表 5

ポリヌクレオチド SEQ ID NO:	Incyte プロジェクト ID	代表的ライブラリ
18	6431478CB1	BRAHTR03
19	3584654CB1	PROSNOT11
20	3737084CB1	PROSNON01
21	71426238CB1	PITUNOT01
22	7475123CB1	CONUTUT01
23	7481952CB1	SINTTMR02
24	382654CB1	BRAHTR03
25	1867351CB1	BRAITUT21
26	3323104CB1	LUNGNOT27
27	4769306CB1	BRAMNOT01
28	2720058CB1	LUNGUT10
29	7481255CB1	THYRTMT01
30	3510242CB1	SINTFER02
31	162131CB1	PANCNOT15
32	1837725CB1	SINDNOT01
33	3643847CB1	CORPNOT02
34	6889872CB1	SINTDIE01

【表 1 8】

表 6-1

ライブラリ	ベクター	ライブラリの説明
BRAHTR03	PCDNA2.1	ライブラリはライブラリは、胆管癌で死亡した 55 才白人女性から切除した旧皮質、前庭馬組織から単離した RNA を用いて作製された。病理は、弓状部にわたる軽度の腫瘍線維形成があり、帯状皮質白質および視床に軸索球状体が散在し、内視野および中脳水道灰白質領域に、少数の神経原線維終末が散在した。関連腫瘍組織の病理は、残存腫瘍または再発性腫瘍を持つ高分化型の肝臓胆管癌を示した。患者の病歴には、胆管癌、術後パッド・キヤーリ症候群、胆管腹水症、水腫症、脱水症、栄養不良、乏尿、および急性腎不全がある。過去の手術には、胆管切除術および、肝臓の 85% の摘除がある。
BRAITUT21	pINCY	ライブラリは、61 才の白人女性の脳腫瘍病変の切除時、正中線前頭葉から採取した脳腫瘍組織から単離した RNA を用いて作製した。病理は非定型を伴わない前頭下脳膜皮髄腫 (subfrontal meningotheial meningioma) を示した。1 つの脳骨と結膜組織サンプルは髄膜腫を示した。家族歴には、脳血管疾患、老人性痴呆症、高脂血症、良性の高血圧、アテローム硬化型冠動脈疾患、鬱血性心不全および乳がんが含まれる。
BRANNO01	pINCY	ライブラリは、心不全で死亡した 35 才白人男性の脳から採取した腫瘍組織から単離された RNA を用いて作製された。病理検査では、軽度の軟膜線維症と大脳新皮質の複雑微小梗塞が見られた。顕微鏡によって、人脳半球は、焦点石灰化がある軽度の軟膜線維症と判明した。大脳半球全体に縮んだわずかに好酸性の垂体ニューロンの形跡があった。さらに、大脳皮質全体に散在して、周辺神経節性を伴う細胞の小さな空泡発生があった。当患者の病歴には、拡張型心筋症、鬱血性心不全、心肥大、及び肥大型の脾臓と肝臓が含まれる。
CONITUT01	pINCY	ライブラリは 61 才の女性の膵式子宮全摘出術および局部的リンパ節切除を伴った両側卵巣摘出時に得た S 字結腸腸間膜腫瘍組織から単離した RNA を用いて作製した。病理は S 状結腸腸間膜の 2 つの部位に転移性のグレード 4 の悪性ニューラ管混合腫瘍の存在が示された。
CORPNOT02	pINCY	ライブラリはアルツハイマー病で死亡した 74 才の白人男性の脳から除去した罹患脳梁組織から単離した RNA を用いて作製した。
LUNGNOT27	pINCY	ライブラリは、17 才のヒスパニック系女子から採取した肺組織から単離した RNA を用いて作製した。
LUNGUT10	pINCY	ライブラリは 65 才の白人女性の肺の部分切除中に左上肺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。病理学検査では、転移性グレード 2 の粘液性脂肪肉腫と転移性グレード 4 の脂肪肉腫を示した。患者の病歴では、軟部組織癌、乳癌、二次性肺癌を示した。
PANCNOT15	pINCY	ライブラリは、15 才白人男子の診断的開腹を伴う遠位膵切除および脾臓全摘時に除去された罹患脾臓組織から単離された RNA を用いて作製された。病理検査では膵臓細胞過形成があり、家族歴には、前立腺癌と心血管疾患とがある。
PITUNOT01	PBLUESCRIPT	ライブラリは、Clontech から得た RNA (CLON 6594-2, lot 35278) を用いて作製した。この RNA の単離は、16 才から 70 才の、外傷で死亡した 18 人の白人男女提供者のプールの下から切除した下体から行った。
PROSNON01	PSPOK11	標準化した前立腺ライブラリは、前立腺ライブラリの 440 万個の独立性クローンから作製した。開始 RNA は白ら紙

10

20

30

【表 19】

表 6－2

ライブラリ	ベクター	ライブラリの説明
		ライブラリは、頭部外傷のため死亡した 16 才白人男子の小腸から除去した十二指腸から単離された RNA を用いて作製された。患者の病歴には腎感染があった。
PROSNOT11	pINCY	ライブラリは、自ら招いた発砲による損傷で死亡した 28 才の白人男性の前立腺組織から単離した RNA を用いて作製した。
SINDNOT01	pINCY	ライブラリは、頭部外傷のため死亡した 16 才白人男子の小腸から除去した十二指腸から単離された RNA を用いて作製された。患者の病歴には腎感染があった。
SINTDIE01	PCDNA2.1	この、5'に偏向してランダムプライムしたライブラリは、49 才の白人女性の、胃腸吻合、診断的開腹、および迷走神経切除時に摘出された小腸組織から単離した RNA を用いて作製した。患者は、閉塞を伴う急性胃潰瘍、悪心および嘔吐、並びに異常な体重減少を示した。患者の病歴としては、腰痛、穿孔を伴う急性胃潰瘍、および正常な出産があった。過去の手術には、扁桃腺アデノイド切除および膣式子宮全摘出がある。患者の使用薬利にはプレマリ
SINTFER02	pINCY	ン (Premarin) がある。家族歴には、良性高血圧、II 型糖尿病、および鬱血性心不全が父親にある。このランダムプライムしたライブラリは、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 を用いて作製された。死因は胎児性死亡である。
SINTTMR02	PCDNA2.1	この、ランダムプライムしたライブラリは、59 才男性から摘出した小腸組織から単離した RNA を用いて作製した。一致する腫瘍組織の病理学検査では、複発 (9) のカルチノイド腫瘍 (グレード 1) が小腸にあった。最大の腫瘍には、1 つの大きな腸間膜塊を伴った。腸の複数の回状セグメントがこの腫瘍に付着していた。1 つ (13 の内) の領域リンパ節は、悪性腫瘍に関して陽性であった。腹膜生検によると腸管性脂肪癌死を示した。
THYRTMT01	pINCY	ライブラリは、56 才の白人女性の片側甲状腺切除および頸部甲狀腺生検時に採取した、左甲状腺組織から単離した RNA を用いて作製された。関連する腫瘍組織の病理は、髄様癌の侵襲が上面の骨格筋に及ぶことを示した。転移性の髄様癌の関与は、1 つの頸動脈鞘リンパ節 (9 の内)、1 つの左頸リンパ節 (結節外伸展を伴う)、および 1 つの中央区画節に及ぶ、顕微鏡的病巣のグレード 1 乳頭状癌が、甲状腺の右葉内で同定された。左甲状腺腫瘍生検では、腫瘍に陰性であった。患者の病歴には、高脂血症、頭痛、およびアテローム性冠動脈疾患がある。家族歴には、脳血管疾患、心血管疾患、および骨癌がある。

【表 2 0】

10

20

30

【表 2 1】

表 7 - 1

プログラム	説明	引用文献	パラメータ関値
ABI FACTURA	ベクター配列を除き、核酸配列のまぎらわしい塩基をマスクするプログラム。	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	アミノ酸または核酸配列の比較、注釈に有用なファーストデータファインダ。	Applied Biosystems, Foster City, CA. Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler	核酸配列の構築を行うプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	アミノ酸および核酸配列の配列類似性検索に有用なベシーングローカルアライメント検索ツール (Basic Local Alignment Search Tool)。BLAST には5つの機能がある: blastp、blastn、blastx、tblastnおよびblastz。	Altschul S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul S.F. 他 (1997) Nucleic Acids Res. 25:3389-3402.	EF: 確率値=1.0E-8 以下 完全長配列: 確率値= 1.0E-10以下
FASTA	一群の同じタイプの配列と問合せ配列との間の類似性について検索するPearson およびLipman アルゴリズム。FASTAは最低5つの機能からなる: fasta、tfasta、fastx、tfastaおよびsssearch。	Pearson, W. R. および D. J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W. R. (1990) Methods Enzymol. 183:63-98, Smith, T.F. および M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	EF: fasta E 値=1.0E-6 構築された EST:fasta 同一性=95%以上 一致した長さ=200塩基以上 fasta E 値=1.0E-8以下 完全長配列: fastx スコア=100以上
BLIMPS	配列をBLOCKS、PRINTS、DOMO、PRODOM およびPFAM データベースの配列と対応させて遺伝子ファミリー、配列相同性および構造的フィンガープリント領域を検索するBLOCKS IMPROVED Searcher。	Henikoff, S. および J. G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J. G. および S. Henikoff (1996) Methods Enzymol. 266:88-105, Attwood, T.K. 他. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	確率値= 1.0E-3以下
HMMER	問合せ配列をPFAMのようなタンパク質ファミリーコンセンサス配列の隠れマルコフモデル (HMM)に基づいたデータベースに対して検索するアルゴリズム。	Krogh, A. 他 (1994) J. Mol. Biol. 235:1501-1531, Sonnhammer, E.L.L. 他 (1998) Nucleic Acids Res. 26:320-322; Durbin, R. 他 (1998) Our World View, in a Nutshell, Cambridge Univ. Press, 1-350 ページ。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	説明	引用文献	パラメータ閾値
ProfileScan	Prositeで定義された配列パターンと一致するタンパク質配列の構造的モチーフおよび配列モチーフを検索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4:61-66; Gribskov, M. 他 (1989) Methods Enzymol. 183:146-159; Baroch, A. 他 (1997) Nucleic Acids Res. 25:217-221.	標準化された質スコアと特定の Prosite モチーフに対する GCG 指定「HIGH」値 通常、スコア=1.4~2.1.
Phred	高度感受性および確立性のある自動シーケンサートレースを調べる塩基コーリングアルゴリズム。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. および P. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的実行に基づいたプログラムである、SWATとCrossMatchを含むPhls改訂構築プログラムで、配列相同性検索やDNA配列の構築に有用である。	Smith, T. F. および M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T. F. および M. S. Waterman (1981) J. Mol. Biol. 147:195-197; Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=56以上
Consed	Phrap 構築の表示、編集用グラフィカルツール。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキャンし、分泌シグナルペプチドの存在を調べる重量マトリックス分析プログラム。	Nielson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J. M. および S. Audic (1997) CABIOS 12:431-439.	スコア=3.6以上
TMAP	重量マトリックスを使ってタンパク質配列上の膜貫通部分を描写し、方向を決定するプログラム。	Persson, B. および P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. および P. Argos (1996) Protein Sci. 5:363-371.	
TWHLHER	隠れマルコフモデル(HMM)を使ってタンパク質配列上の膜貫通部分を描写し、方向を決定するプログラム。	Sonnhammer, E. L. 他 (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol., Biol., Glasgow 他, 編集, The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, 175-182ページ.	175-182.
Motifs	Prositeで定義されたものと一致するパターンのアミノ酸配列を検索するプログラム。	Baroch, A. 他 (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, バージョン9, M51-59ページ, Genetics Computer Group, Madison, WI	

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
2 May 2002 (02.05.2002)

PCT

(10) International Publication Number
WO 02/34783 A2

(51) International Patent Classification: C07K 14/705

(21) International Application Number: PCT/US01/49670

(22) International Filing Date: 26 October 2001 (26.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/244,017 27 October 2000 (27.10.2000) US
60/252,855 22 November 2000 (22.11.2000) US
60/251,825 7 December 2000 (07.12.2000) US
60/255,085 12 December 2000 (12.12.2000) 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): WARREN, Bridget,
A. [US/US]; 10130 Parkwood Drive #2, Capertino, CA
95014 (US). XU, Yanning [US/US]; 1739 Walnut Drive,
Mountain View, CA 94040 (US). YUE, Henry [US/US];
826 Lois Avenue, Sunnyvale, CA 94087 (US). BATRA,
Sajeev [US/US]; 555 El Camino Real, #709, San Leand-
ro, CA 94087 (US). BURFORD, Neil [GB/US]; 105
Wildwood Circle, Durham, CT 06422 (US). GANDHI,
Ameena, R. [US/US]; 705 5th Avenue, San Francisco, CA
94118 (US). WALIA, Nariader, K. [US/US]; 890 David
Street, #205, San Leandro, CA 94577 (US). ARVIZU,
Chaudra [US/US]; 490 Sherwood Way #1, Menlo Park,
CA 94025 (US). TANG, Y., Tom [US/US]; 4230 Ranwic
Court, San Jose, CA 95118 (US). LU, Dyung, Alina, M.
[US/US]; 233 Goy Drive, San Jose, CA 95123 (US). DUG-
GAN, Brendan, M. [AU/US]; 243 Buena Vista Avenue
#306, Sunnyvale, CA 94086 (US). BAUGHN, Mariah, R.
[US/US]; 14244 Santiago Road, San Leandro, CA 94577
(US). LEE, Ernestine, A. [US/US]; 624 Kains Street, Al-
bany, CA 94706 (US). KHAN, Farrah, A. [IN/US]; 9445Harrison Street, Des Plaines, IL 60016 (US). NGUYEN,
Daniel, B. [US/US]; 1403 Ridgwood Drive, San Jose,
CA 95118 (US). AZIMZAI, Yalda [US/US]; 5518 Boul-
der Canyon Drive, Castro Valley, CA 94552 (US). YAO,
Monique, G. [US/US]; 1189 Woodgate Drive, Carmel,
IN 46033 (US). LAL, Preeti, G. [IN/US]; P.O. Box 5142,
Santa Clara, CA 95056 (US). THANGAVELU, Kavitha
[IN/US]; 1950 Montecito Avenue #23, Mountain View, CA
94043 (US). RAMKUMAR, Jayalaxmi [IN/US]; 34359
Maybird Circle, Fremont, CA 94555 (US). TRAN, Bao
[US/US]; 750 Salberg Avenue, Santa Clara, CA 95051
(US). DING, Li [CN/US]; 3353 Alma Street #146, Palo
Alto, CA 94306 (US). AU-YOUNG, Janice [US/US]; 233
Golden Eagle Lane, Brisbane, CA 94005 (US).(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG).

Published:

— without international search report and to be republished
upon receipt of that reportFor two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: TRANSMEMBRANE PROTEINS

(57) Abstract: The invention provides human transmembrane proteins (TMP) and polynucleotides which identify and encode TMP.
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 TMP.

WO 02/34783 A2

WO 02/34783

PCT/US01/49670

TRANSMEMBRANE PROTEINS

TECHNICAL FIELD

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

10

BACKGROUND OF THE INVENTION

Eukaryotic organisms are distinct from prokaryotes in possessing many intracellular membrane-bound compartments such as organelles and vesicles. Many of the metabolic reactions which distinguish eukaryotic biochemistry from prokaryotic biochemistry take place within these compartments. In particular, many cellular functions require very stringent reaction conditions, and the organelles and vesicles enable compartmentalization and isolation of reactions which might otherwise disrupt cytosolic metabolic processes. The organelles include mitochondria, smooth and rough endoplasmic reticulum, sarcoplasmic reticulum, and the Golgi body. The vesicles include phagosomes, lysosomes, endosomes, peroxisomes, and secretory vesicles. Organelles and vesicles are bounded by single or double membranes.

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

Plasma Membrane Proteins

Transmembrane proteins (TMP) are characterized by extracellular, transmembrane, and intracellular domains. TMP domains are typically comprised of 15 to 25 hydrophobic amino acids which are predicted to adopt an α -helical conformation. TMP are classified as bitopic (Types I and II) proteins, which span the membrane once, and polytopic (Types III and IV) (Singer, S.J. (1990) Annu. Rev. Cell Biol. 6:247-96) proteins, which contain multiple membrane-spanning segments. TMP that act as cell-surface receptor proteins involved in signal transduction include growth and differentiation factor receptors, and receptor-interacting proteins such as *Drosophila* pecanex and frizzled proteins, LIV-1 protein, NF2 protein, and GNS1/SUR4 eukaryotic integral membrane

35

WO 02/34783

PCT/US01/49670

proteins. TMP also act as transporters of ions or metabolites, such as gap junction channels (connexins) and ion channels, and as cell anchoring proteins, such as lectins, integrins, and fibronectins. TMP function as vesicle and organelle-forming molecules, such as caveolins; or cell recognition molecules, such as cluster of differentiation (CD) antigens, glycoproteins, and mucins.

5 The transport of hydrophilic molecules across membranes is facilitated by the presence of channel proteins which form aqueous pores which can perforate a lipid bilayer. Many channels consist of protein complexes formed by the assembly of multiple subunits, at least one of which is an integral membrane protein that contributes to formation of the pore. In some cases, the pore is constructed to allow selective passage of only one or a few molecular species. Distinct types of
10 membrane channels that differ greatly in their distribution and selectivity include: (1) aquaporins, which transport water; (2) protein-conducting channels, which transport proteins across the endoplasmic reticulum membrane; (3) gap junctions, which facilitate diffusion of ions and small organic molecules between neighboring cells; and (4) ion channels, which regulate ion flux through various membranes.

15 Gap junctions (also called connexons) are specialized regions of the plasma membrane comprising transmembrane channels that function chemically and electrically to couple the cytoplasm of neighboring cells in many tissues. Gap junctions function as electrical synapses for intercellular propagation of action potentials in excitable tissues. In nonexcitable tissues, gap junctions have roles in tissue homeostasis, coordinated physiological response, metabolic cooperation, growth control, and the
20 regulation of development and differentiation.

Each connexon, which spans the lipid bilayer of the plasma membrane, is composed of six identical subunits called connexins. At least fourteen distinct connexin proteins exist, with each having similar structures but differing tissue distributions. Structurally, the connexins consist of a short cytoplasmic N-terminal domain connected to four transmembrane spanning regions (M1, M2, M3 and
25 M4) which separate two extracellular and one cytoplasmic loop followed by a C-terminal, cytoplasmic domain of variable length (20 residues in Cx26 to 260 residues in Cx56). The M2-M3 loop and the N- and C-termini are oriented towards the cell cytoplasm. Conserved regions include the membrane spanning regions and the two extracellular loops. Within the extracellular loops are three conserved cysteines which are involved in disulfide bond formation. Signature patterns for these two loops are
30 either: C-[DN]-T-x-Q-P-G-C-x-(2)-V-C-Y-D or C-x(3,4)-P-C-x(3)-[LIVM]-[DEN]-C-[FY]-[LIVM]-[SA]-[KR]-P (PDOC00341, Profilescan and S. Rahman and W.H. Evans, (1991) J. Cell Sci. 100:567-578). The variable regions, which include the cytoplasmic loop and the C-terminal region, may be responsible for the regulation of different connexins. (See Hennemann, H. et al. (1992) J. Biol. Chem. 267:17225-17233; PRINTS PR00206 connexin signature; Yeager, M. et al., (1998) Curr. O. Struct.
35 Biol. 8:517-524.)

WO 02/34783

PCT/US01/49670

Gap junctions help to synchronize heart and smooth muscle contraction, speed neural transmission, and propagate extracellular signals. Gap junctions can open and close in response to particular stimuli (e.g., pH, Ca^{+2} , and cAMP). The effective pore size of a gap junction is approximately 1.5 nm, which enables small molecules (e.g., those under 1000 daltons) to diffuse freely through the pore. Transported molecules include ions, small metabolites, and second messengers (e.g., Ca^{+2} and cAMP).

Connexins have many disease associations. Female mice lacking connexin 37 (Cx37) are infertile due to the absence of the oocyte-granulosa cell signaling pathway. Mice lacking Cx43 die shortly after birth and show cardiac defects reminiscent of some forms of stenosis of the pulmonary artery in humans. Mutations in Cx32 are associated with the X-linked form of Charcot-Marie-Tooth disease, a motor and sensory neuropathy of the peripheral nervous system. Cx26 is expressed in the placenta, and Cx26-deficient mice show decreased transplacental transport of a glucose analog from the maternal to the fetal circulation. In humans, Cx26 has been identified as the first susceptibility gene for non-syndromic sensorineural autosomal deafness. Mutations in Cx31 have been linked with an autosomal-dominant hearing impairment (a nonsense or missense mutation in the second extracellular loop) and in a dominantly transmitted skin disorder, erythrokratoderma variabilis (missense mutations in either the N-terminal domain or the M2 domain.) (See A. M. Simon, (1999) Trends Cell Biol. 9:169-170). Cx46 is expressed in lens fiber cells, and Cx46-deficient mice develop early-onset cataracts that resemble human nuclear cataracts. (See Nicholson, S.M. and R. Bruzzone (1997) Curr. Biol. 7:R340-R344.)

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

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

Chemical modification of amino acid residue side chains alters the manner in which MPs interact with other molecules, such as membrane phospholipids. Examples of such chemical modifications include the formation of covalent bonds with glycosaminoglycans, oligosaccharides,

WO 02/34783

PCT/US01/49670

phospholipids, acetyl and palmitoyl moieties, ADP-ribose, phosphate, and sulphate groups.

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

5 Transmembrane proteins of the plasma membrane also include cell surface receptors. These receptors recognize hormones such as catecholamines, e.g., epinephrine, norepinephrine, and histamine; peptide hormones, e.g., glucagon, insulin, gastrin, secretin, cholecystokinin, adrenocorticotrophic hormone, follicle stimulating hormone, luteinizing hormone, thyroid stimulating hormone, parathyroid hormone, and vasopressin; growth and differentiation factors, e.g., epidermal growth factor, fibroblast growth factor, transforming growth factor, insulin-like growth factor, platelet-derived growth factor, nerve growth factor, colony-stimulating factors, and erythropoietin; cytokines, e.g., chemokines, interleukins, interferons, and tumor necrosis factor; small peptide factors such as bombesin, oxytocin, endothelin, angiotensin II, vasoactive intestinal peptide, and bradykinin; neurotransmitters such as neuropeptide Y, neurotensin, neuropeptide N, melanocortins, opioids, e.g., 15 enkephalins, endorphins and dynorphins; galanin, somatostatin, and tachykinins; and circulatory system-borne signaling molecules, e.g., angiotensin, complement, calcitonin, endothelins, and formyl-methionyl peptides. Cell surface receptors on immune system cells recognize antigens, antibodies, and major histocompatibility complex (MHC)-bound peptide. Other cell surface receptors bind ligands to be internalized by the cell. This receptor-mediated endocytosis functions in the uptake 20 of low density lipoproteins (LDL), transferrin, glucose- or mannose-terminal glycoproteins, galactose-terminal glycoproteins, immunoglobulins, phosphovitellogenins, fibrin, proteinase-inhibitor complexes, plasminogen activators, and thrombospondin. (Lodish, H. et al. (1995) *Molecular Cell Biology*, Scientific American Books, New York, NY, p. 723; and Mikhailenko, I. et al. (1997) *J. Biol. Chem.* 272:6784-6791.)

25 Many cell surface receptors have seven transmembrane regions, with an extracellular N-terminus that binds ligand and a cytoplasmic C-terminus that interacts with G proteins. (Strosberg, A.D. (1991) *Eur. J. Biochem.* 196:1-10.) Cysteine-rich domains are found in two families of cell surface receptors, the LDL receptor family and the tumor necrosis factor receptor/nerve growth factor (TNFR/NGFR) receptor family. Seven successive cysteine-rich repeats of about forty amino acids in 30 the N-terminal extracellular region of the LDL receptor form the binding site for LDL and calcium; similar repeats have been found in vertebrate very low density lipoprotein receptor, vertebrate low-density lipoprotein receptor-related protein 1 (LRP1) (also known as α_2 -macroglobulin receptor), and vertebrate low-density lipoprotein receptor-related protein 2 (also known as gp330 or megalin) (ExPASy PROSITE document PDOC00929; and Bairoch, A. et al. (1997) *Nucl. Acids. Res.* 25:217-35 221.) The structure of the repeat is a β -hairpin followed by a series of β -turns; there are six disulfide-

WO 02/34783

PCT/US01/49670

bonded cysteines within each repeat.

The LDL receptor is an integral membrane protein which functions in lipid uptake by removing cholesterol from the blood. Most cells outside the liver and intestine take up cholesterol from the blood rather than synthesize it themselves. Cell surface LDL receptors bind LDL particles which are then
 5 internalized by endocytosis (Meyers, R.A. (1995) Molecular Biology and Biotechnology, VCH Publishers, New York NY, pp. 494-501). Absence of the LDL receptor, the cause of the disease familial hypercholesterolemia, leads to increased plasma cholesterol levels and ultimately to atherosclerosis (Stryer, L. (1995) Biochemistry, W.H. Freeman, New York NY, pp. 691-702).
G-Protein Coupled Receptors

10 G-protein coupled receptors (GPCR) comprise a superfamily of integral membrane proteins which transduce extracellular signals. GPCRs include receptors for biogenic amines, lipid mediators of inflammation, peptide hormones, and sensory signal mediators.

The structure of these highly-conserved receptors consists of seven hydrophobic transmembrane (serpentine) regions, cysteine disulfide bridges between the second and third
 15 extracellular loops, an extracellular N-terminus, and a cytoplasmic C-terminus. Three extracellular loops alternate with three intracellular loops to link the seven transmembrane regions. The most conserved parts of these proteins are the transmembrane regions and the first two cytoplasmic loops. Cysteine disulfide bridges connect the second and third extracellular loops. A conserved, acidic-Arg-aromatic residue triplet present in the second cytoplasmic loop may interact with G
 20 proteins. A GPCR consensus pattern is characteristic of most proteins belonging to this superfamily (ExPASy PROSITE document PS00237; and Watson, S. and S. Arkinstall (1994) The G-protein Linked Receptor Facts Book, Academic Press, San Diego, CA, pp 2-6). Mutations and changes in transcriptional activation of GPCR-encoding genes have been associated with neurological disorders such as schizophrenia, Parkinson's disease, Alzheimer's disease, drug addiction, and feeding
 25 disorders. The juvenile development and fertility-2 (jdf-2) locus, also called runty-jerky-sterile (rjs), is associated with deletions and point mutations in HERC2, a gene encoding a guanine nucleotide exchange factor protein involved in vesicular trafficking (Walkowicz, M. et al. (1999) Mamm. Genome 10:870-878).

A GPCR known as FP is the receptor for prostaglandin $F_{2\alpha}$ (PGF_{2 α}). The prostaglandins
 30 belong to a large family of naturally occurring paracrine/autocrine mediators of physiologic and inflammatory responses. PGF_{2 α} plays a role in responses of certain tissues such as reproductive tract, lung, bone, and heart, including the stimulation of myometrial contraction, corpus luteum breakdown, and bronchoconstriction. An FP-associated molecule (FPRP) is copurified with FP and is expressed only in those tissues where a physiological role for PGF_{2 α} has been described. FPRP is predicted to
 35 be a transmembrane protein with glycosylated extracellular immunoglobulin loops and a short, highly

WO 02/34783

PCT/US01/49670

charged intracellular domain. FPRP appears to be a negative regulator of $\text{PGF}_{2\alpha}$ binding to FP. As such, FPRP may be associated with $\text{PGF}_{2\alpha}$ related diseases, which may include dysmenorrhea, infertility, asthma, or cardiomyopathy (Orlicky, D. J. et al. (1996) Hum. Genet. 97:655-658).

Scavenger Receptors

5 Macrophage scavenger receptors with broad ligand specificity may participate in the binding of low density lipoproteins (LDL) and foreign antigens. Scavenger receptors types I and II are trimeric membrane proteins with each subunit containing a small N-terminal intracellular domain, a transmembrane domain, a large extracellular domain, and a C-terminal cysteine-rich domain. The extracellular domain contains a short spacer domain, an α -helical coiled-coil domain, and a triple
10 helical collagenous domain. These receptors have been shown to bind a spectrum of ligands, including chemically modified lipoproteins and albumin, polyribonucleotides, polysaccharides, phospholipids, and asbestos (Matsumoto, A. et al. (1990) Proc. Natl. Acad. Sci. 87:9133-9137; and Elomaa, O. et al. (1995) Cell 80:603-609). The scavenger receptors are thought to play a key role in atherogenesis by mediating uptake of modified LDL in arterial walls, and in host defense by binding
15 bacterial endotoxins, bacteria, and protozoa.

Tetraspan family proteins

The transmembrane 4 superfamily (TM4SF), or tetraspan family, is a multigene family encoding type III integral membrane proteins (Wright, M.D. and Tomlinson, M.G. (1994) Immunol. Today 15:588-594). TM4SF is comprised of membrane proteins which traverse the cell membrane
20 four times. Members of the TM4SF include platelet and endothelial cell membrane proteins, melanoma-associated antigens, leukocyte surface glycoproteins, colonic carcinoma antigens, tumor-associated antigens, and surface proteins of the schistosome parasites (Jankowski, S.A. (1994) Oncogene 9:1205-1211). Members of the TM4SF share about 25-30% amino acid sequence identity with one another.

25 A number of TM4SF members have been implicated in signal transduction, control of cell adhesion, regulation of cell growth and proliferation, including development and oncogenesis, and cell motility, including tumor cell metastasis. Expression of TM4SF proteins is associated with a variety of tumors, and the level of expression may be altered when cells are growing or activated.

Tumor Antigens

30 Tumor antigens are surface molecules that are differentially expressed in tumor cells relative to normal cells. Tumor antigens distinguish tumor cells immunologically from normal cells and provide diagnostic and therapeutic targets for human cancers (Takagi, S. et al. (1995) Int. J. Cancer 61: 706-715; Liu, E. et al. (1992) Oncogene 7: 1027-1032).

Ion channels

35 Ion channels are found in the plasma membranes of virtually every cell in the body. For

WO 02/34783

PCT/US01/49670

example, chloride channels mediate a variety of cellular functions including regulation of membrane potential and absorption and secretion of ions across epithelial membranes. When present in intracellular membranes of the Golgi apparatus and endocytic vesicles, chloride channels also regulate organelle pH (see, e.g., Greger, R. (1988) *Annu. Rev. Physiol.* 50:111-122). Electrophysiological and pharmacological properties of chloride channels, including ion conductance, current-voltage relationships, and sensitivity to modulators, suggest that different chloride channels exist in muscles, neurons, fibroblasts, epithelial cells, and lymphocytes.

Many channels have sites for phosphorylation by one or more protein kinases including protein kinase A, protein kinase C, casein kinase II, and tyrosine kinases, all of which regulate ion channel activity in cells. Inappropriate phosphorylation of membrane proteins has been correlated with pathological changes in cell cycle progression and cell differentiation. Changes in the cell cycle have been linked to induction of apoptosis or cancer. Changes in cell differentiation have been linked to diseases and disorders of the reproductive system, immune system, and skeletal muscle.

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

Proton pumps

Proton ATPases are a large class of membrane proteins that use the energy of ATP hydrolysis to generate an electrochemical proton gradient across a membrane. The resultant gradient may be used to transport other ions across the membrane (Na^+ , K^+ , or Cl^-) or to maintain organelle pH. Proton ATPases are further subdivided into the mitochondrial F-ATPases, the plasma membrane ATPases, and the vacuolar ATPases. The vacuolar ATPases establish and maintain an acidic pH within various vesicles involved in the processes of endocytosis and exocytosis (Mellman, I. et al. (1986) *Ann. Rev. Biochem.* 55:663-700).

Proton-coupled, 12 membrane-spanning domain transporters such as PEPT 1 and PEPT 2 are responsible for gastrointestinal absorption and for renal reabsorption of peptides using an electrochemical H^+ gradient as the driving force. Another type of peptide transporter, the TAP transporter, is a heterodimer consisting of TAP 1 and TAP 2 and is associated with antigen

WO 02/34783

PCT/US01/49670

processing. Peptide antigens are transported across the membrane of the endoplasmic reticulum by TAP so they can be expressed on the cell surface in association with MHC molecules. Each TAP protein consists of multiple hydrophobic membrane spanning segments and a highly conserved ATP-binding cassette (Boll, M. et al. (1996) Proc. Natl. Acad. Sci. 93:284-289). Pathogenic
 5 microorganisms, such as herpes simplex virus, may encode inhibitors of TAP-mediated peptide transport in order to evade immune surveillance (Marusina, K. and Manaco, J.J. (1996) Curr. Opin. Hematol. 3:19-26).

ABC Transporters

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

Membrane Proteins Associated with Intercellular Communication

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

Synaptobrevins are synaptic vesicle-associated membrane proteins (VAMPs) which were first
 30 discovered in rat brain. These proteins were initially thought to be limited to neuronal cells and to function in the movement of vesicles from the plasmalemma of one cell, across the synapse, to the plasmalemma of another cell. Synaptobrevins are now known to occur and function in constitutive vesicle trafficking pathways involving receptor-mediated endocytotic and exocytotic pathways of many non-neuronal cell types. This regulated vesicle trafficking pathway may be blocked by the highly
 35 specific action of clostridial neurotoxins which cleave the synaptobrevin molecule.

WO 02/34783

PCT/US01/49670

In vitro studies of various cellular membranes (Galli et al. (1994) J. Cell. Biol. 125:1015-24; Link et al. (1993) J. Biol. Chem. 268:18423-6) have shown that VAMPs are widely distributed. These important membrane trafficking proteins appear to participate in axon extension via exocytosis during development, in the release of neurotransmitters and modulatory peptides, and in endocytosis.

5 Endocytotic vesicular transport includes such intracellular events as the fusions and fissions of the nuclear membrane, endoplasmic reticulum, Golgi apparatus, and various inclusion bodies such as peroxisomes or lysosomes. Endocytotic processes appear to be universal in eukaryotic cells as diverse as yeast, *Caenorhabditis elegans*, *Drosophila*, and mammals.

VAMP-1B is involved in subcellular targeting and is an isoform of VAMP-1A (Isenmann, S. et al. (1998) Mol. Biol. Cell 9:1649-1660). Four additional splice variants (VAMP-1C to F) have recently been identified. Each variant has variable sequences only at the extreme C-terminus, suggesting that the C-terminus is important in vesicle targeting (Berglund, L. et al. (1999) Biochem. Biophys. Res. Commun. 264:777-780).

Lysosomes are the site of degradation of intracellular material during autophagy, and of 15 extracellular molecules following endocytosis. Lysosomal enzymes are packaged into vesicles which bud from the *trans*-Golgi network. These vesicles fuse with endosomes to form the mature lysosome in which hydrolytic digestion of endocytosed material occurs. Lysosomes can fuse with autophagosomes to form a unique compartment in which the degradation of organelles and other intracellular components occurs.

20 Protein sorting by transport vesicles, such as the endosome, has important consequences for a variety of physiological processes including cell surface growth, the biogenesis of distinct intracellular organelles, endocytosis, and the controlled secretion of hormones and neurotransmitters (Rothman, J.E. and Wieland, F.T. (1996) Science 272:227-234). In particular, neurodegenerative disorders and other neuronal pathologies are associated with biochemical flaws during endosomal protein sorting or 25 endosomal biogenesis (Mayer R.J. et al. (1996) Adv. Exp. Med. Biol. 389:261-269).

Peroxisomes are organelles independent from the secretory pathway. They are the site of many peroxide-generating oxidative reactions in the cell. Peroxisomes are unique among eukaryotic organelles in that their size, number, and enzyme content vary depending upon organism, cell type, and metabolic needs (Waterham, H.R. and Cregg, J.M. (1997) BioEssays 19:57-66). Genetic defects in 30 peroxisome proteins which result in peroxisomal deficiencies have been linked to a number of human pathologies, including Zellweger syndrome, rhizomelic chondrodysplasia punctata, X-linked adrenoleukodystrophy, acyl-CoA oxidase deficiency, bifunctional enzyme deficiency, classical Refsum's disease, DHAP alkyl transferase deficiency, and acatalasemia (Moser, H.W. and Moser, A.B. (1996) Ann. NY Acad. Sci. 804:427-441). In addition, Gartner, J. et al. (1991; Pediatr. Res. 29:141- 35 146) found a 22 kDa integral membrane protein associated with lower density peroxisome-like

WO 02/34783

PCT/US01/49670

subcellular fractions in patients with Zellweger syndrome.

Normal embryonic development and control of germ cell maturation is modulated by a number of secretory proteins which interact with their respective membrane-bound receptors. Cell fate during embryonic development is determined by members of the activin/TGF- β superfamily, cadherins, IGF-2, and other morphogens. In addition, proliferation, maturation, and redifferentiation of germ cell and reproductive tissues are regulated, for example, by IGF-2, inhibins, activins, and follistatins (Petraglia, F. (1997) Placenta 18:3-8; Mather, J.P. et al. (1997) Proc. Soc. Exp. Biol. Med. 215:209-222).

Endoplasmic Reticulum Membrane Proteins

The normal functioning of the eukaryotic cell requires that all newly synthesized proteins be correctly folded, modified, and delivered to specific intra- and extracellular sites. Newly synthesized membrane and secretory proteins enter a cellular sorting and distribution network during or immediately after synthesis and are routed to specific locations inside and outside of the cell. The initial compartment in this process is the endoplasmic reticulum (ER) where proteins undergo modifications such as glycosylation, disulfide bond formation, and oligomerization. The modified proteins are then transported through a series of membrane-bound compartments which include the various cisternae of the Golgi complex, where further carbohydrate modifications occur. Transport between compartments occurs by means of vesicle budding and fusion. Once within the secretory pathway, proteins do not have to cross a membrane to reach the cell surface.

Although the majority of proteins processed through the ER are transported out of the organelle, some are retained. The signal for retention in the ER in mammalian cells consists of the tetrapeptide sequence, KDEL, located at the carboxyl terminus of resident ER membrane proteins (Munro, S. (1986) Cell 46:291-300). Proteins containing this sequence leave the ER but are quickly retrieved from the early Golgi cisternae and returned to the ER, while proteins lacking this signal continue through the secretory pathway.

Disruptions in the cellular secretory pathway have been implicated in several human diseases. In familial hypercholesterolemia the low density lipoprotein receptors remain in the ER, rather than moving to the cell surface (Pathak, R.K. (1988) J. Cell Biol. 106:1831-1841). Altered transport and processing of the β -amyloid precursor protein (β APP) involves the putative vesicle transport protein presenilin and may play a role in early-onset Alzheimer's disease (Levy-Lahad, E. et al. (1995) Science 269:973-977). Changes in ER-derived calcium homeostasis have been associated with diseases such as cardiomyopathy, cardiac hypertrophy, myotonic dystrophy, Brody disease, Smith-McCort dysplasia, and diabetes mellitus.

Mitochondrial Membrane Proteins

The mitochondrial electron transport (or respiratory) chain is a series of three enzyme

WO 02/34783

PCT/US01/49670

complexes in the mitochondrial membrane that is responsible for the transport of electrons from NADH to oxygen and the coupling of this oxidation to the synthesis of ATP (oxidative phosphorylation). ATP then provides the primary source of energy for driving the many energy-requiring reactions of a cell.

5 Most of the protein components of the mitochondrial respiratory chain are the products of nuclear encoded genes that are imported into the mitochondria, and the remainder are products of mitochondrial genes. Defects and altered expression of enzymes in the respiratory chain are associated with a variety of disease conditions in man, including, for example, neurodegenerative diseases, myopathies, and cancer.

10 Lymphocyte and Leukocyte Membrane Proteins

The B-cell response to antigens is an essential component of the normal immune system. Mature B cells recognize foreign antigens through B cell receptors (BCR) which are membrane-bound, specific antibodies that bind foreign antigens. The antigen/receptor complex is internalized, and the antigen is proteolytically processed. To generate an efficient response to complex antigens, 15 the BCR, BCR-associated proteins, and T cell response are all required. Proteolytic fragments of the antigen are complexed with major histocompatibility complex-II (MHCII) molecules on the surface of the B cells where the complex can be recognized by T cells. In contrast, macrophages and other lymphoid cells present antigens in association with MHCI molecules to T cells. T cells recognize and are activated by the MHCI-antigen complex through interactions with the T cell receptor/CD3 20 complex, a T cell-surface multimeric protein located in the plasma membrane. T cells activated by antigen presentation secrete a variety of lymphokines that induce B cell maturation and T cell proliferation, and activate macrophages, which kill target cells.

Leukocytes have a fundamental role in the inflammatory and immune response, and include monocytes/macrophages, mast cells, polymorphonucleoleukocytes, natural killer cells, neutrophils, 25 eosinophils, basophils, and myeloid precursors. Leukocyte membrane proteins include members of the CD antigens, N-CAM, I-CAM, human leukocyte antigen (HLA) class I and HLA class II gene products, immunoglobulins, immunoglobulin receptors, complement, complement receptors, interferons, interferon receptors, interleukin receptors, and chemokine receptors.

Abnormal lymphocyte and leukocyte activity has been associated with acute disorders such as 30 AIDS, immune hypersensitivity, leukemias, leukopenia, systemic lupus, granulomatous disease, and eosinophilia.

Apoptosis-Associated Membrane Proteins

A variety of ligands, receptors, enzymes, tumor suppressors, viral gene products, pharmacological agents, and inorganic ions have important positive or negative roles in regulating 35 and implementing the apoptotic destruction of a cell. Although some specific components of the

WO 02/34783

PCT/US01/49670

apoptotic pathway have been identified and characterized, many interactions between the proteins involved are undefined, leaving major aspects of the pathway unknown.

A requirement for calcium in apoptosis was previously suggested by studies showing the involvement of calcium levels in DNA cleavage and Fas-mediated cell death (Hewish, D.R. and L.A. Burgoyne (1973) *Biochem. Biophys. Res. Comm.* 52:504-510; Vignaux, F. et al. (1995) *J. Exp. Med.* 181:781-786; Oshimi, Y. and S. Miyazaki (1995) *J. Immunol.* 154:599-609). Other studies show that intracellular calcium concentrations increase when apoptosis is triggered in thymocytes by either T cell receptor cross-linking or by glucocorticoids, and cell death can be prevented by blocking this increase (McConkey, D.J. et al. (1989) *J. Immunol.* 143:1801-1806; McConkey, D.J. et al. (1989) *Arch. Biochem. Biophys.* 269:365-370). Therefore, membrane proteins such as calcium channels and the Fas receptor are important for the apoptotic response.

The discovery of new transmembrane 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 reproductive, developmental, cardiovascular, neurological, gastrointestinal, lipid metabolism, cell proliferative, and autoimmune/inflammatory disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of transmembrane proteins.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, transmembrane proteins, referred to collectively as "TMP" and individually as "TMP-1," "TMP-2," "TMP-3," "TMP-4," "TMP-5," "TMP-6," "TMP-7," "TMP-8," "TMP-9," "TMP-10," "TMP-11," "TMP-12," "TMP-13," "TMP-14," "TMP-15," "TMP-16," and "TMP-17." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-17.

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:1-17, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the

WO 02/34783

PCT/US01/49670

group consisting of SEQ ID NO:1-17, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-17. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:18-34.

5 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:1-17, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

15 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:1-17, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

25 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:1-17, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17.

30 The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34, c) a

WO 02/34783

PCT/US01/49670

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

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-17. The invention additionally provides a

WO 02/34783

PCT/US01/49670

method of treating a disease or condition associated with decreased expression of functional TMP, 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:1-17, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17. 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 TMP, 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:1-17, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17. 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 TMP, 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:1-17, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID

WO 02/34783

PCT/US01/49670

NO:1-17. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-17. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34, iii) a

WO 02/34783

PCT/US01/49670

polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

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

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

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

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

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

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

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

DESCRIPTION OF THE INVENTION

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

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a

WO 02/34783

PCT/US01/49670

reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

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

DEFINITIONS

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

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

An "allelic variant" is an alternative form of the gene encoding TMP. Allelic variants may
 20 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
 25 a given sequence.

"Altered" nucleic acid sequences encoding TMP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as TMP or a polypeptide with at least one functional characteristic of TMP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of
 30 the polynucleotide encoding TMP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding TMP. 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 TMP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility,
 35 hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological

WO 02/34783

PCT/US01/49670

or immunological activity of TMP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, 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 TMP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of TMP either by directly interacting with TMP or by acting on components of the biological pathway in which TMP 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 TMP 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 "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an *in vitro* evolutionary process (e.g., SELEX

WO 02/34783

PCT/US01/49670

(Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH₂), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker. (See, e.g., Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13.)

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

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

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the 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 TMP, 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'.

WO 02/34783

PCT/US01/49670

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 TMP or fragments of TMP 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.

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

WO 02/34783

PCT/US01/49670

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.

5 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
10 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.

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

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

A "fragment" is a unique portion of TMP or the polynucleotide encoding TMP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to
25 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
30 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.

35 A fragment of SEQ ID NO:18-34 comprises a region of unique polynucleotide sequence that

WO 02/34783

PCT/US01/49670

specifically identifies SEQ ID NO:18-34, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:18-34 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:18-34 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:18-34 and the region of SEQ ID NO:18-34 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:1-17 is encoded by a fragment of SEQ ID NO:18-34. A fragment of SEQ ID NO:1-17 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-17. For example, a fragment of SEQ ID NO:1-17 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-17. The precise length of a fragment of SEQ ID NO:1-17 and the region of SEQ ID NO:1-17 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 parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several

WO 02/34783

PCT/US01/49670

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

Matrix: BLOSUM62
Reward for match: 1
Penalty for mismatch: -2
Open Gap: 5 and Extension Gap: 2 penalties
Gap x drop-off: 50
Expect: 10
Word Size: 11
Filter: on

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

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

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, 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.

WO 02/34783

PCT/US01/49670

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
 5 "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
 10 (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

15 *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,
 20 a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

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

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

30 "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
 35 of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e.,

WO 02/34783

PCT/US01/49670

binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

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

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

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

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

WO 02/34783

PCT/US01/49670

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 TMP 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 TMP 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.

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

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

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

"Probe" refers to nucleic acid sequences encoding TMP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short

WO 02/34783

PCT/US01/49670

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).

5 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
10 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
15 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
20 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
25 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
30 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.
35 Hence, this program is useful for identification of both unique and conserved oligonucleotides and

WO 02/34783

PCT/US01/49670

polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

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

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

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

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

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

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

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure

WO 02/34783

PCT/US01/49670

of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

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

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

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

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

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

20 A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with 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,

WO 02/34783

PCT/US01/49670

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 alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

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

THE INVENTION

The invention is based on the discovery of new human transmembrane proteins (TMP), the polynucleotides encoding TMP, and the use of these compositions for the diagnosis, treatment, or prevention of reproductive, developmental, cardiovascular, neurological, gastrointestinal, lipid metabolism, cell proliferative, and autoimmune/inflammatory 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

WO 02/34783

PCT/US01/49670

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

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are transmembrane proteins. For example, SEQ ID NO:2 is 89% identical to rat prostaglandin F2a receptor regulatory protein (GenBank ID g1054884) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:2 also contains six immunoglobulin domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) In addition, SEQ ID NO:2 contains a signal peptide, a transmembrane domain, and an RGD motif, providing further corroborative evidence that SEQ ID NO:2 is a human transmembrane protein.

In the alternative, SEQ ID NO:4 is 56% identical to human connexin 31.1 (GenBank ID g4336903) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 5.8e-68, which indicates the probability of obtaining the observed

WO 02/34783

PCT/US01/49670

polypeptide sequence alignment by chance. SEQ ID NO:4 also contains a connexin domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:4 is a connexin.

5 Note that six identical connexins compose a connexon (gap junction), a transmembrane channel in the plasma membrane which functions chemically and electrically to couple the cytoplasm of neighboring cells in many tissues. SEQ ID NO:5 is 1554 amino acids in length and is 99% identical over 1157 amino acids to human MEGF7 (GenBank ID g3449306) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the

10 probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:5 also contains low-density lipoprotein receptor repeats and low-density lipoprotein receptor domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS analyses provide further corroborative evidence that SEQ ID NO:5 is a member of the LDL receptor family of

15 proteins.

In another alternative, SEQ ID NO:6 is 36% identical to mouse low density lipoprotein receptor related protein LRP1B/LRP-DIT (GenBank ID g8926243) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.5e-40, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:6 also

20 contains low-density lipoprotein receptor domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS analyses provide further corroborative evidence that SEQ ID NO:6 is a low-density lipoprotein receptor-related molecule. Further, SEQ ID NO:14 is 59% identical to human TNF-inducible protein CG12-1 (GenBank ID g3978246) as

25 determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.2e-94, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. Data from HMMER analysis provides further corroborative evidence that SEQ ID NO:14 contains a transmembrane domain. SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:7-13, and SEQ ID NO:15-17 were analyzed and annotated in a similar manner. The algorithms and

30 parameters for the analysis of SEQ ID NO:1-17 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

35 consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention.

WO 02/34783

PCT/US01/49670

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:18-34 or that distinguish between SEQ ID NO:18-34 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences in column 5 relative to their respective full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 679882711 is the identification number of an Incyte cDNA sequence, and COLENOR03 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., 71760758V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g1506353) which contributed to the assembly of the full length polynucleotide sequences. In addition, the identification numbers in column 5 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the identification numbers in column 5 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, FL_XXXXXX_N₁_N₂_YYYYY_N₃_N₄ represents a "stitched" sequence in which XXXXXX is the identification number of the cluster of sequences to which the algorithm was applied, and YYYYY is the number of the prediction generated by the algorithm, and N_{1,2,3,...}, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the identification numbers in column 5 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, FL_XXXXXX_gAAAAA_gBBBBB_1_N is the identification number of a "stretched" sequence, with XXXXXX being the Incyte project identification number, gAAAAA being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, gBBBBB being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and N referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, gBBBBB).

WO 02/34783

PCT/US01/49670

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

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

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

15 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.

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

The invention also encompasses polynucleotides which encode TMP. In a particular 25 embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:18-34, which encodes TMP. The polynucleotide sequences of SEQ ID NO:18-34, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

30 The invention also encompasses a variant of a polynucleotide sequence encoding TMP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least

WO 02/34783

PCT/US01/49670

about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding TMP. 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:18-34 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:18-34. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of TMP.

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide sequence encoding TMP. A splice variant may have portions which have significant sequence identity to the polynucleotide sequence encoding TMP, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing of exons during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to the polynucleotide sequence encoding TMP over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide sequence encoding TMP. Any one of the splice variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of TMP.

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 TMP, 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 TMP, and all such variations are to be considered as being specifically disclosed.

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

WO 02/34783

PCT/US01/49670

The invention also encompasses production of DNA sequences which encode TMP and TMP 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 TMP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:18-34 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (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 TMP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al.

WO 02/34783

PCT/US01/49670

(1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060).

- 5 Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a
- 10 GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

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

- Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the
- 20 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
- 25 which may be present in limited amounts in a particular sample.

- In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode TMP may be cloned in recombinant DNA molecules that direct expression of TMP, 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
- 30 amino acid sequence may be produced and used to express TMP.

- The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter TMP-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
- 35 oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-

WO 02/34783

PCT/US01/49670

mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen, Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of TMP, 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 TMP 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, TMP 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 TMP, 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 TMP, the nucleotide sequences encoding TMP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains

WO 02/34783

PCT/US01/49670

- 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 TMP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding TMP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding TMP 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.)
- Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding TMP 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.)
- A variety of expression vector/host systems may be utilized to contain and express sequences encoding TMP. 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., T3 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.*

WO 02/34783

PCT/US01/49670

USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding TMP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding TMP can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPO1 plasmid (Life Technologies). Ligation of sequences encoding TMP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of TMP are needed, e.g. for the production of antibodies, vectors which direct high level expression of TMP 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 TMP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

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

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding TMP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses TMP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc.

WO 02/34783

PCT/US01/49670

Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

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

For long term production of recombinant proteins in mammalian systems, stable expression of TMP in cell lines is preferred. For example, sequences encoding TMP 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 TMP is inserted within a marker gene sequence, transformed cells containing sequences encoding TMP can be identified by the absence of marker gene function. Alternatively, a

WO 02/34783

PCT/US01/49670

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

In general, host cells that contain the nucleic acid sequence encoding TMP and that express

- 5 TMP 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 TMP using either
10 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 TMP 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.
15 (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

- A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization
20 or PCR probes for detecting sequences related to polynucleotides encoding TMP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding TMP, 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
25 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 radiolabeled nucleotides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

- 30 Host cells transformed with nucleotide sequences encoding TMP 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 TMP may be designed to contain signal sequences which direct secretion
35 of TMP through a prokaryotic or eukaryotic cell membrane.

WO 02/34783

PCT/US01/49670

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
 5 which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding TMP 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 TMP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of TMP activity. Heterologous protein and peptide moieties
 15 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,
 20 respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the TMP encoding sequence and the heterologous protein sequence, so that TMP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein
 25 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 TMP 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
 30 SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

TMP of the present invention or fragments thereof may be used to screen for compounds that specifically bind to TMP. At least one and up to a plurality of test compounds may be screened for specific binding to TMP. Examples of test compounds include antibodies, oligonucleotides, proteins
 35 (e.g., receptors), or small molecules.

WO 02/34783

PCT/US01/49670

In one embodiment, the compound thus identified is closely related to the natural ligand of TMP, 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 TMP
 5 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 TMP, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing TMP or cell membrane fractions which contain TMP are then contacted with a
 10 test compound and binding, stimulation, or inhibition of activity of either TMP 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 TMP, either in
 15 solution or affixed to a solid support, and detecting the binding of TMP 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.

TMP of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of TMP. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for TMP activity, wherein TMP is combined with at least one test compound, and the activity of TMP in the presence of a test compound is compared with the activity of TMP in the absence of the test compound.
 20 A change in the activity of TMP in the presence of the test compound is indicative of a compound that modulates the activity of TMP. Alternatively, a test compound is combined with an in vitro or cell-free system comprising TMP under conditions suitable for TMP activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of TMP 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
 25 compounds may be screened.

In another embodiment, polynucleotides encoding TMP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example,
 30 mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and

WO 02/34783

PCT/US01/49670

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

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

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

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of TMP and transmembrane proteins. In addition, the expression of TMP is closely associated with brain, prostate, smooth muscle, cardiovascular, pituitary, gastrointestinal, lung, pancreatic, and small intestine tissues. Therefore, TMP appears to play a role in reproductive, developmental, cardiovascular, neurological, gastrointestinal, lipid metabolism, cell proliferative, and autoimmune/inflammatory disorders. In the treatment of disorders associated with increased TMP expression or activity, it is desirable to decrease the expression or activity of TMP. In the treatment of disorders associated with decreased TMP expression or activity, it is desirable to increase the expression or activity of TMP.

Therefore, in one embodiment, TMP 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 TMP.

WO 02/34783

PCT/US01/49670

Examples of such disorders include, but are not limited to, a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, ectopic pregnancy, teratogenesis, cancer of the breast, fibrocystic breast disease, galactorrhea, a

5 disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, gynecomastia, hypergonadotropic and hypogonadotropic hypogonadism, pseudohermaphroditism, azoospermia, premature ovarian failure, acrosin deficiency, delayed puberty, retrograde ejaculation and

10 anejaculation, haemangioblastomas, cystsphaechromocytomas, paraganglioma, cystadenomas of the epididymis, and endolymphatic sac tumours; 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,

15 hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; a cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis, Raynaud's disease,

20 aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebothrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral valve prolapse,

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

30 obstructive pulmonary disease, restrictive pulmonary disease, chronic obstructive pulmonary disease, emphysema, chronic bronchitis, bronchial asthma, bronchiectasis, bacterial pneumonia, viral and mycoplasmal pneumonia, lung abscess, pulmonary tuberculosis, diffuse interstitial diseases, pneumoconioses, sarcoidosis, idiopathic pulmonary fibrosis, desquamative interstitial pneumonitis, hypersensitivity pneumonitis, pulmonary eosinophilia bronchiolitis obliterans-organizing pneumonia,

35 diffuse pulmonary hemorrhage syndromes, Goodpasture's syndromes, idiopathic pulmonary

WO 02/34783

PCT/US01/49670

hemosiderosis, pulmonary involvement in collagen-vascular disorders, pulmonary alveolar proteinosis, lung tumors, inflammatory and noninflammatory pleural effusions, pneumothorax, pleural tumors, drug-induced lung disease, radiation-induced lung disease, and complications of lung transplantation; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha₁-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; a lipid metabolism disorder such as fatty liver, cholestasis, primary biliary cirrhosis,

WO 02/34783

PCT/US01/49670

- carnitine deficiency, carnitine palmitoyltransferase deficiency, myoadenylate deaminase deficiency, hypertriglyceridemia, lipid storage disorders such as Fabry's disease, Gaucher's disease, Niemann-Pick's disease, metachromatic leukodystrophy, adrenoleukodystrophy, GM₂ gangliosidosis, and ceroid lipofuscinosis, abetalipoproteinemia, Tangier disease, hyperlipoproteinemia, diabetes mellitus,
- 5 lipodystrophy, lipomatosis, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipoid adrenal hyperplasia, minimal change disease, lipomas, atherosclerosis, hypercholesterolemia, hypercholesterolemia with hypertriglyceridemia, primary hypoalphalipoproteinemia, hypothyroidism, renal disease, liver disease, lecithin:cholesterol acyltransferase deficiency, cerebrotendinous xanthomatosis, sitosterolemia, hypocholesterolemia, Tay-Sachs disease, Sandhoff's disease,
- 10 hyperlipidemia, hyperlipemia, lipid myopathies, and obesity; 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,
- 15 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 an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia,
- 20 autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel
- 25 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,
- 30 parasitic, protozoal, and helminthic infections, and trauma.

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

- In a further embodiment, a composition comprising a substantially purified TMP in conjunction
- 35 with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder

WO 02/34783

PCT/US01/49670

associated with decreased expression or activity of TMP including, but not limited to, those provided above.

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

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

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding TMP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TMP including, but not limited to, those described above.

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

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

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with TMP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG

WO 02/34783

PCT/US01/49670

(*bacilli Calmette-Guerin*) and *Cornebacterium parvum* are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to TMP 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 TMP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

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

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6851-6855; Neuberger, M.S. et al. (1984) *Nature* 312:604-608; and Takeda, S. et al. (1985) *Nature* 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce TMP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) *Proc. Natl. Acad. Sci. USA* 88:10134-10137.)

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

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

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

WO 02/34783

PCT/US01/49670

immunoassays typically involve the measurement of complex formation between TMP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering TMP epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

5 Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for TMP. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of TMP-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 TMP
10 epitopes, represents the average affinity, or avidity, of the antibodies for TMP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular TMP epitope, 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 TMP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7
15 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of TMP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

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

In another embodiment of the invention, the polynucleotides encoding TMP, 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 TMP.
30 Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding TMP. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press, 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
35 intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence

WO 02/34783

PCT/US01/49670

- 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* 5 76:271; Ausubel, *supra*; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, I.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.)
- 10 In another embodiment of the invention, polynucleotides encoding TMP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined
- 15 immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaise, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii)
- 20 express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA.* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides*
- 25 *brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in TMP expression or regulation causes disease, the expression of TMP 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 TMP
- 30 are treated by constructing mammalian expression vectors encoding TMP and introducing these vectors by mechanical means into TMP-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).

WO 02/34783

PCT/US01/49670

Expression vectors that may be effective for the expression of TMP include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). TMP may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding TMP from a normal individual.

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

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to TMP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding TMP under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and A.D. Miller (1988) *J. Virol.* 62:3802-3806; Dull, T. et al. (1998) *J. Virol.* 72:8463-8471; Zufferey, R. et al. (1998) *J. Virol.* 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference.

WO 02/34783

PCT/US01/49670

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 TMP to cells which have one or more genetic abnormalities with respect to the expression of TMP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) *Transplantation* 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) *Annu. Rev. Nutr.* 19:511-544 and Verma, I.M. and N. Sornia (1997) *Nature* 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding TMP to target cells which have one or more genetic abnormalities with respect to the expression of TMP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing TMP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding TMP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV

WO 02/34783

PCT/US01/49670

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 TMP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of TMP-coding RNAs and the synthesis of high levels of TMP 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 (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of TMP 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 TMP.

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

WO 02/34783

PCT/US01/49670

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.

5 Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding TMP. 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
10 modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine,
15 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 TMP. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides,
20 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 TMP expression or activity, a compound which specifically inhibits expression of the polynucleotide
25 encoding TMP may be therapeutically useful, and in the treatment of disorders associated with decreased TMP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding TMP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method
30 commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a
35 polynucleotide encoding TMP is exposed to at least one test compound thus obtained. The sample

WO 02/34783

PCT/US01/49670

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

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

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

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

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

WO 02/34783

PCT/US01/49670

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the 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.

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 TMP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, TMP 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 TMP or fragments thereof, antibodies of TMP, and agonists, antagonists or inhibitors of TMP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

WO 02/34783

PCT/US01/49670

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.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg , up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. 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 TMP may be used for the diagnosis of disorders characterized by expression of TMP, or in assays to monitor patients being treated with TMP or agonists, antagonists, or inhibitors of TMP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for TMP include methods which utilize the antibody and a label to detect TMP 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 TMP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of TMP expression. Normal or standard values for TMP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to TMP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of TMP 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 TMP 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 TMP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of

WO 02/34783

PCT/US01/49670

TMP, and to monitor regulation of TMP levels during therapeutic intervention.

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

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the TMP 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:18-34 or from genomic sequences including promoters, enhancers, and introns of the TMP gene.

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

Polynucleotide sequences encoding TMP may be used for the diagnosis of disorders associated with expression of TMP. Examples of such disorders include, but are not limited to, a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, ectopic pregnancy, teratogenesis, cancer of the breast, fibrocystic breast disease, galactorrhea, a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, gynecomastia, hypergonadotropic and hypogonadotropic hypogonadism, pseudohermaphroditism, azoospermia, premature ovarian failure, acrosin deficiency, delayed puberty, retrograde ejaculation and anejaculation, haemangioblastomas, cystsphaeochromocytomas, paraganglioma, cystadenomas of the epididymis, and endolymphatic sac tumours; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary

WO 02/34783

PCT/US01/49670

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; a cardiovascular disorder such as arteriovenous fistula, atherosclerosis, hypertension, vasculitis,

5 Raynaud's disease, aneurysms, arterial dissections, varicose veins, thrombophlebitis and phlebotrombosis, vascular tumors, and complications of thrombolysis, balloon angioplasty, vascular replacement, coronary artery bypass graft surgery, congestive heart failure, ischemic heart disease, angina pectoris, myocardial infarction, hypertensive heart disease, degenerative valvular heart disease, calcific aortic valve stenosis, congenitally bicuspid aortic valve, mitral annular calcification, mitral

10 valve prolapse, rheumatic fever and rheumatic heart disease, infective endocarditis, nonbacterial thrombotic endocarditis, endocarditis of systemic lupus erythematosus, carcinoid heart disease, cardiomyopathy, myocarditis, pericarditis, neoplastic heart disease, congenital heart disease, complications of cardiac transplantation, congenital lung anomalies, atelectasis, pulmonary congestion and edema, pulmonary embolism, pulmonary hemorrhage, pulmonary infarction,

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

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

25 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,

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

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

WO 02/34783

PCT/US01/49670

- nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD),
- 5 akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis,
- 10 gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation,
- 15 gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha₁-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic
- 20 cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; a lipid metabolism disorder such as fatty liver, cholestasis, primary biliary cirrhosis, carnitine deficiency, carnitine palmitoyltransferase deficiency, myoadenylate deaminase deficiency, hypertriglyceridemia, lipid storage disorders such as Fabry's disease, Gaucher's disease, Niemann-Pick's disease, metachromatic leukodystrophy, adrenoleukodystrophy, GM₂ gangliosidosis, and
- 25 ceroid lipofuscinosis, abetalipoproteinemia, Tangier disease, hyperlipoproteinemia, diabetes mellitus, lipodystrophy, lipomatosis, acute panniculitis, disseminated fat necrosis, adiposis dolorosa, lipoid adrenal hyperplasia, minimal change disease, lipomas, atherosclerosis, hypercholesterolemia, hypercholesterolemia with hypertriglyceridemia, primary hypoalphalipoproteinemia, hypothyroidism, renal disease, liver disease, lecithin:cholesterol acyltransferase deficiency, cerebrotendinous
- 30 xanthomatosis, sitosterolemia, hypocholesterolemia, Tay-Sachs disease, Sandhoff's disease, hyperlipidemia, hyperlipemia, lipid myopathies, and obesity; 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,
- 35 myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone,

WO 02/34783

PCT/US01/49670

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 an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma. The polynucleotide sequences encoding TMP 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 TMP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding TMP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding TMP 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 TMP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of TMP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding TMP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values

WO 02/34783

PCT/US01/49670

from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

5 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.

10 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
15 progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding TMP 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 TMP, or a fragment of a polynucleotide complementary to the polynucleotide encoding TMP, and will be
20 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 TMP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions,
25 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 TMP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily
30 fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of
35 identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which

WO 02/34783

PCT/US01/49670

assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system
 5 (Sequenom, San Diego CA).

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

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

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

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by
 30 quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the
 35 hybridization takes place in high-throughput format, wherein the polynucleotides of the present

WO 02/34783

PCT/US01/49670

invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

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

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

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

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome

WO 02/34783

PCT/US01/49670

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

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

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

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

WO 02/34783

PCT/US01/49670

in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

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

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

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

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical and genetic 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 TMP on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA

WO 02/34783

PCT/US01/49670

associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps.

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

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

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with TMP, or fragments thereof, and washed. Bound TMP is then detected by methods well known in the art. Purified TMP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding TMP specifically compete with a test compound for binding TMP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with TMP.

In additional embodiments, the nucleotide sequences which encode TMP 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

WO 02/34783

PCT/US01/49670

in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/244,017, U.S. Ser. No. 60/252,835, U.S. Ser. No. 60/251,825, and U.S. Ser. No. 60/255,085, are hereby expressly incorporated by reference.

5

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 SUPERSCRIPIT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX

WO 02/34783

PCT/US01/49670

DH10B from Life Technologies.

II. Isolation of cDNA Clones

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

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

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing 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,

WO 02/34783

PCT/US01/49670

DOMO, PRODOM; PROTEOME databases with sequences from *Homo sapiens*, *Rattus norvegicus*, *Mus musculus*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Candida albicans* (Incyte Genomics, Palo Alto CA); and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV 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, the PROTEOME databases, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San 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 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).

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:18-34. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

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

WO 02/34783

PCT/US01/49670

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

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the

WO 02/34783

PCT/US01/49670

longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpi public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

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

VI. Chromosomal Mapping of TMP Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:18-34 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:18-34 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM

WO 02/34783

PCT/US01/49670

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:22 was mapped to chromosome 11 within the interval from 59.50 to 62.50 centiMorgans and SEQ ID NO:26 was mapped to chromosome 1 within the interval from 179.2 to 186.4 centiMorgans.

VII. Analysis of Polynucleotide Expression

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

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

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

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

WO 02/34783

PCT/US01/49670

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 TMP. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

15 VIII. Extension of TMP 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.

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

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research). 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.

WO 02/34783

PCT/US01/49670

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 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:18-34 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

WO 02/34783

PCT/US01/49670

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. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, *supra*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Skena (1999), *supra*). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Skena, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)

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

WO 02/34783

PCT/US01/49670

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 μ l volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories (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, Holbrook NY) and resuspended in 14 μ l 5X SSC/0.2% SDS.

Microarray Preparation

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

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

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

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.

WO 02/34783

PCT/US01/49670

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

Hybridization

5 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 10 140 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 15 Innova 70 mixed gas 10 W laser (Coherent, 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, 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 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 25 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 30 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 35 fluorophores and adding identical amounts of each to the hybridization mixture.

WO 02/34783

PCT/US01/49670

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

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

XI. Complementary Polynucleotides

Sequences complementary to the TMP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring TMP. 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 TMP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the TMP-encoding transcript.

XII. Expression of TMP

Expression and purification of TMP is achieved using bacterial or virus-based expression systems. For expression of TMP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*lac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express TMP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of TMP 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 TMP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional

WO 02/34783

PCT/US01/49670

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

15 XIII. Functional Assays

20 TMP function is assessed by expressing the sequences encoding TMP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G.

WO 02/34783

PCT/US01/49670

(1994) Flow Cytometry, Oxford, New York NY.

The influence of TMP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding TMP 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

5 immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding TMP and other genes of interest can be analyzed by northern analysis or microarray techniques.

10 XIV. Production of TMP Specific Antibodies

TMP 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 TMP amino acid sequence is analyzed using LASERGENE software
15 (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
20 peptide synthesizer (Applied Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-TMP activity by, for example, binding the peptide or TMP to a substrate, blocking with 1%
25 BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring TMP Using Specific Antibodies

Naturally occurring or recombinant TMP is substantially purified by immunoaffinity chromatography using antibodies specific for TMP. An immunoaffinity column is constructed by covalently coupling anti-TMP antibody to an activated chromatographic resin, such as CNBr-activated
30 SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing TMP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of TMP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/TMP binding (e.g.,
35 a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and

WO 02/34783

PCT/US01/49670

TMP is collected.

XVI. Identification of Molecules Which Interact with TMP

TMP, or biologically active fragments thereof, are labeled with ^{125}I 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 TMP, washed, and
 5 any wells with labeled TMP complex are assayed. Data obtained using different concentrations of TMP are used to calculate values for the number, affinity, and association of TMP with the candidate molecules.

Alternatively, molecules interacting with TMP 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).

TMP 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
 15 No. 6,057,101).

XVII. Demonstration of TMP Activity

Gap Junction Activity of TMP

Gap junction activity of TMP is demonstrated as the ability to induce the formation of intercellular channels between paired *Xenopus laevis* oocytes injected with TMP cRNA (Hennemann,
 20 *supra*). One week prior to the experimental injection with TMP cRNA, oocytes are injected with antisense oligonucleotide to TMP to reduce background. TMP cRNA-injected oocytes are incubated overnight, stripped of vitelline membranes, and paired for recording of junctional currents by dual cell voltage clamp. The measured conductances are proportional to gap junction activity of TMP.

Alternatively, an assay for TMP activity measures the ion channel activity of TMP using an electrophysiological assay for ion conductance. TMP can be expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector encoding TMP.
 25 Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. A second plasmid which expresses any one of a number of marker genes, such as β -galactosidase, is co-transformed into the cells to allow rapid
 30 identification of those cells which have taken up and expressed the foreign DNA. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression and accumulation of TMP and β -galactosidase.

Transformed cells expressing β -galactosidase are stained blue when a suitable colorimetric substrate is added to the culture media under conditions that are well known in the art. Stained cells
 35 are tested for differences in membrane conductance by electrophysiological techniques that are well

WO 02/34783

PCT/US01/49670

known in the art. Untransformed cells, and/or cells transformed with either vector sequences alone or β -galactosidase sequences alone, are used as controls and tested in parallel. Cells expressing TMP will have higher anion or cation conductance relative to control cells. The contribution of TMP to conductance can be confirmed by incubating the cells using antibodies specific for TMP. The antibodies will bind to the extracellular side of TMP, thereby blocking the pore in the ion channel, and the associated conductance.

Transmembrane Protein Activity of TMP

An assay for TMP activity measures the expression of TMP on the cell surface. cDNA encoding TMP is transfected into an appropriate mammalian cell line. Cell surface proteins are labeled with biotin as described (de la Fuente, M. A. et al. (1997) *Blood* 90:2398-2405). Immunoprecipitations are performed using TMP-specific antibodies, and immunoprecipitated samples are analyzed using SDS-PAGE and immunoblotting techniques. The ratio of labeled immunoprecipitant to unlabeled immunoprecipitant is proportional to the amount of TMP expressed on the cell surface.

An alternative assay for TMP activity is based on a prototypical assay for ligand/receptor-mediated modulation of cell proliferation. This assay measures the amount of newly synthesized DNA in Swiss mouse 3T3 cells expressing TMP. An appropriate mammalian expression vector containing cDNA encoding TMP is added to quiescent 3T3 cultured cells using transfection methods well known in the art. The transfected cells are incubated in the presence of [^{3} H]thymidine and varying amounts of TMP ligand. Incorporation of [^{3} H]thymidine into acid-precipitable DNA is measured over an appropriate time interval using a tritium radioisotope counter, and the amount incorporated is directly proportional to the amount of newly synthesized DNA. A linear dose-response curve over at least a hundred-fold TMP ligand concentration range is indicative of receptor activity. One unit of activity per milliliter is defined as the concentration of TMP producing a 50% response level, where 100% represents maximal incorporation of [^{3} H]thymidine into acid-precipitable DNA (McKay, I. and Leigh, L., eds. (1993) *Growth Factors: A Practical Approach*, Oxford University Press, New York, NY, p. 73).

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments.

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

WO 02/34783

PCT/US01/49670

Table 1

Project ID	IncYTE	Polypeptide SEQ ID NO:	Polypeptide ID	IncYTE	Polynucleotide SEQ ID NO:	Polynucleotide ID
6431478		1	6431478CD1		18	6431478CB1
3584654		2	3584654CD1		19	3584654CB1
3737084		3	3737084CD1		20	3737084CB1
71426238		4	71426238CD1		21	71426238CB1
7475123		5	7475123CD1		22	7475123CB1
7481232		6	7481232CD1		23	7481232CB1
366231		7	366231CD1		24	366231CB1
3323104		8	3323104CD1		25	3323104CB1
4769306		9	4769306CD1		26	4769306CB1
2720058		10	2720058CD1		27	2720058CB1
7481255		11	7481255CD1		28	7481255CB1
1510242		12	1510242CD1		29	1510242CB1
162131		13	162131CD1		30	162131CB1
1837725		14	1837725CD1		31	1837725CB1
3643847		15	3643847CD1		32	3643847CB1
6889872		16	6889872CD1		33	6889872CB1
		17			34	

Table 2

Polypeptide SBO ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
1	6431478CDL	g310100	4.0E-216	[Rattus norvegicus] developmentally expressed protein
2	3584654CDL	g1054884	0.0	[Rattus norvegicus] prostaglandin F2a receptor regulatory protein Orlicky D.J. (1996) Negative regulatory activity of a prostaglandin F2a receptor associated protein (FPRP). Prostaglandins Leukot. Essent. Fatty Acids 58, 247-259
3	3737084CDL	g3513451	1.7E-233	[Rattus norvegicus] potassium channel related protein Hoehli N. et al. (1998) KCRI, a membrane protein that facilitates functional expression of non-inactivating K ⁺ currents associates with rat EAG voltage-dependent K ⁺ channels. F. Biol. Chem. 273, 23080-23085
4	714426238CDL	g15906851	1.0E-130	[Homo sapiens] (AJ414563) comexin238
5	7475123CDL	g3449306	0.0	[Homo sapiens] MS3F7 (multiple KCP-1-like protein) Nakayama M. et al. (1998) Genomics 51:27-34.
6	7481952CDL	g6926243	1.5E-40	[Mus musculus] low density lipoprotein receptor related protein LRP1B/LRP-DTT
9	3323104CDL	g11414879	0.0	[Homo sapiens] mannosyltransferase Maeda, Y. et al. (2001) PIG-M (mannosyltransferase) is a novel glycosylphosphatidylinositol on the luminal side of the ER. EMBO J. 20:250-261

WO 02/34783

PCT/US01/49670

Table 2 (cont.)

Polypeptide Seq. ID NO.	Incyte Polypeptide ID	GenBank ID NO.	Probability score	GenBank Homolog
11	2720058CD1	96013381	1.5E-12	TM6P1 (integral membrane protein) [Rattus norvegicus] [Zhang, J. et al. (2000) Biochim. Biophys. Acta 1492:280-284]
12	7481255CD1	54150939	4.2E-27	[Mus musculus] GSG1
13	1310242CD1	94263743	5.3E-163	[Rattus norvegicus] similar to U89424 (P166268)
14	162131CD1	613374351	0.0	[Homo sapiens] apolipoprotein L4 (P166268)
15	1837725CD1	96850311	6.0E-52	Contains similarity to a vacuolar sorting receptor homolog from [Arabidopsis thaliana] gb U79959
16	3643847CD1	95944332	2.8E-95	[Mus musculus] Ttyal Campbell, H.D. et al. (2000)
17	6889872CD1	93191975	1.3E-95	Genomics 68:89-92 G6305.3 (putative leucine rich protein) [Homo sapiens]

Table 3

SEQ ID NO.	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Assays
1	6434478CD1	461	S147, S180, S225, S240, S335, S34, T109, T221, T333, T386	M113, M123, M294	Transmembrane domains: N358-T216, Y94-T112, Y156-V174, Y395-R516, Y381-T248, Y388-M406. REGULATED PLACENTAL DIFF33 DEVELOPMENTALLY PROTEIN R11H6.2 PD011773:176-P283, K192-S460 PROTEIN PLACENTAL DIFF33 R11H6.2 DEVELOPMENTALLY REGULATED PD018175:M1-E70	BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM
2	3584654CD1	879	S150, S158, S244, S248, S249, S46, S47, S56, S72, S623, S66, S72, S747, S758, S772, T303, T402, T472, T505, T506, T570, T757, T820, T866, T97, Y640	M286, N300, N351, M413, M444, M451, M452, M618, M691	Signal cleavage: M1-Q21 Signal peptide: M1-R25 Transmembrane domain: L833-Q851 Immunoglobulin fold: C162-Y249, C292-Y375, D422-V517, G564-V657, G704-V795 PROTEIN PROSTAGLANDIN F2ALPHA RECEPTOR REGULATORY PRECURSOR ASSOCIATED SIGNAL IMMUNOGLOBULIN FOLD PD025644:19-S814 PROSTAGLANDIN F2ALPHA RECEPTOR REGULATORY PROTEIN PRECURSOR ASSOCIATED SIGNAL IMMUNOGLOBULIN FOLD TRANSFORMER GLYCOPROTEIN PD185779:K815-S879 IMMUNOGLOBULIN PD040001:139207 25-139, Y26-V132, DM00001, I39207 141-259, V140-E255 RGD motif: S703-D705	SPSCAN HMMER HMMER_PP4M BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_DOMO BLAST_DOMO MOTIFS

Table 3 (cont.)

SEQ NO:	Inocyte clayp- totype	Amino acid Residues	Potential Phosphoryl- ation sites	Potential Glycosyl- ation sites	Signature Sequences, Domains and Motifs	Analytical Methods and References
3	3737084CD1	473	S243 S280 S383 S387 S56 T368 Y434 Y417 Y444	N246 N457	Transmembrane domains: M95-Y112, P141-M166, Y257-G277, N396-T419 PROTEIN CHANNEL DIE2 POTASSIUM REGULATOR IONIC C56F8.06C CHROMOSOME I PRECURSOR PD025171: E33-S387, W395-W428, S435-W473 POTASSIUM CHANNEL REGULATOR 1 IONIC CHANNEL PD184319.M1-F32 Connexin: M1-L201 CONEXINS DMO0590 Q02739.1-268: M1-E218 Q02738.1-268: M1-E208 Q02738.1-262: M1-K206 P08034.1-276: M1-K204 GAP JUNCTION CONNEXIN PROTEIN TRANSMEMBRANE ALPHA1 CX43 ALPHA8 ALPHA5 BETA1 PD0011135: W3-K108, M16-L201 GAP JUNCTION PROTEIN CONNEXIN TRANSMEMBRANE BETA2 CX26 BETA4 CX31.1 DISEASE PD001118: G119-L201 Connexins proteins: BL00407: S2-V38, A39-F69, F70-Y97, G118-D147, C157-L201 CONNEXIN SIGNATURE PR00206: L20-W44, F51- Q73, L76-A96, L420-Y146, C157-Y177, IL78- L201 Transmembrane domain: T23-A39, L120-Y143, Y177-P200 Connexins signature 2: C157-P173 Connexins signatures connexins 1: L43-T86 Connexins signatures connexins 2: H36-Y193 HMMER	HMMER
4	71426238CD1	223	T132 T174		BLAST_PRODOM HMMER PFAM BLAST_DOMO BLAST_PRODOM	BLAST_PRODOM

92

[illegible]

WO 02/34783

PCT/US01/49670

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
6	7481952CD1	L718	S1037 S1031 S1085 S1111 S1165 S1171 S1310 S1423 S1310 S1463 S1482 S1510 S1517 S1544 S1583 S1686 S200 S214 S240 S274 S28 S366 S312 S366 S380 S312 S366 S377 S338 T1178 T1238 T133 T1359 T1371 T1391 T468 T504 T597 T613 T676 T763 T827	N1235 N126 N1299 N1345 N1545 N1634 N1684 N183 N260 N375 N496 N551 N611 N703 N761 N910 N969 N976	Low-density lipoprotein receptor domain: P21-L68, P69-P107, Q237-W275, P189-E227, R146-S184, R103-H145. Transmembrane domain: H142-L160 Low-density lipoprotein receptor domain: S183-P123, H610-W649, C106-S1091, S1463-T1502, P1507-N1545, K1546-E1586, R834-T864, E1244-H1283 MM domain: C38-I199, C216-L378, C427-L586, C652-P818, C869-S1027, C1083-T1238, C1291-T1454 LDL-receptor class A H101209: MM domain: MM domain: E100740A: P1031 E100740B: R1011-S1031 Low density lipoprotein receptor PR00261: L1476-E1497, R397-E418, K836-E857 MM domain proteins: DML1344: P98072 352-509:E418-D578 DML1344: P28824 595-796:E424-D577 DML1344: A55620 961-1128:C821-L576 C821 attachment sequences (RGD): K1008-E1038, R397-E418, R397-E418 Transmembrane domain: I51-W71	HMMER-PPAM HMMER HMMER-PPAM HMMER-PPAM HMMER-PPAM BLIMPS-BLOCKS BLIMPS-BLOCKS BLIMPS-PRINTS ELAST-DMO MOTIFS HMMER
7	382654CD1	224	S139 S165 S48 S74 T179 T202			

WO 02/34783

PCT/US01/49670

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
8	1867351CD1	570	S317 S336 S356 S38 594 T367	N82	Signal peptide: M1-C63 Transmembrane domains: L38-I58, A302-N322, V338-C357, I414-M436, L438-I457, L542-R564 Uncharacterized membrane protein family (Pfam: PF00043)	SPScan HMMER
9	3323104CD1	423	S141 S203 S3 T263 T417 T48 T73 Y52	N400	Integral membrane protein PD004336: V344-W462 Integral membrane protein PD149928: I55-P129 Leucine zipper pattern: I537-L558 Transmembrane domains: L228-Y247, F386-I405 Leucine zipper pattern: I215-L236 Interferon region transmembrane protein: I215-L236 Transmembrane domain: V200-L220 Cytochrome domain: C272-F139 Low-density lipoprotein receptor domain: P145-E183 LDL-receptor class A BL01209: C166-E178 Signal peptide: M1-A14 Transmembrane domains: F6-S26, V95-E115, T196-D215 Signal peptide: M1-S28 Transmembrane domains: L122-G144, M152-T184 FMP-22 / NRP / M20 family BL01221: F88-N101, Y203-T229 Transmembrane domains: L146-L172, V217-L235, P333-W351	BLAST-PRODOM BLAST-PRODOM MOTIFS HMMER MOTIFS BLAST-PRODOM HMMER HMMER-PFAM HMMER-PFAM BLIMPS-BLOCKS SPScan HMMER HMMER SPScan HMMER BLIMPS-BLOCKS HMMER
10	4769306CD1	388	S141 S188 S232 S279 S294 S303 S357 S361 S377 S4 596 597 T187 T252 T385 T50	N161 N195 N301 N336 N384	Signal peptide: M1-S28 Transmembrane domains: L122-G144, M152-T184 FMP-22 / NRP / M20 family BL01221: F88-N101, Y203-T229 Transmembrane domains: L146-L172, V217-L235, P333-W351	BLIMPS-BLOCKS SPScan HMMER HMMER-PFAM HMMER-PFAM BLIMPS-BLOCKS SPScan HMMER
11	2720058CD1	231	S137 T177 T220 T42 Y190	N54 N75	Signal peptide: M1-S28 Transmembrane domains: L122-G144, M152-T184 FMP-22 / NRP / M20 family BL01221: F88-N101, Y203-T229 Transmembrane domains: L146-L172, V217-L235, P333-W351	BLIMPS-BLOCKS SPScan HMMER HMMER-PFAM HMMER-PFAM BLIMPS-BLOCKS SPScan HMMER
12	7481255CD1	293	S109 S148 S251 S95 T236 T35	N243 N277 N39 N67	Signal peptide: M1-S28 Transmembrane domains: L122-G144, M152-T184 FMP-22 / NRP / M20 family BL01221: F88-N101, Y203-T229 Transmembrane domains: L146-L172, V217-L235, P333-W351	BLIMPS-BLOCKS SPScan HMMER HMMER-PFAM HMMER-PFAM BLIMPS-BLOCKS SPScan HMMER
13	1510242CD1	526	S279 S479 T22 Y125	N180 N201 N378	Signal peptide: M1-S28 Transmembrane domains: L122-G144, M152-T184 FMP-22 / NRP / M20 family BL01221: F88-N101, Y203-T229 Transmembrane domains: L146-L172, V217-L235, P333-W351	BLIMPS-BLOCKS SPScan HMMER HMMER-PFAM HMMER-PFAM BLIMPS-BLOCKS SPScan HMMER

WO 02/34783

PCT/US01/49670

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
14	16213CD1	348	S45 S187 S344 T33 T54 T207 T246 T271 Y39 T27 T36 S54 S141 T37 T35 T157 T202 T331 T505	N35 N69 N66 N74 N116 N126 N149 N155	Transmembrane domain: I149-L175 APOLIPROTEIN L PRECURSOR ABOL PLASMA LIPID TRANSPORT GLYCOPROTEIN SIGNAL D96802.1 PD042084: T9-A347 Signal cleavage: M1-A25 Transmembrane domain: M172-G191, P218-L237, W243-Y261, V287-F306, L346-E371 Signal cleavage: M1-G57 Transmembrane domain: L46-F62, L214-G233, L390-C408 Cell attachment sequence R164-D166 signal cleavage: M1-A22 Transmembrane domain: T394-Y416 Transmembrane domain: T352-F371, R80-S103, S104-S127, R128-F151, S152-A175	HMMER BLAST_PRODOM SPSCAN HMMER SPSCAN HMMER BLAST_PRODOM MOTIFS SPSCAN HMMER HMMER_Pfam
15	1837725CD1	520	S27 S35 S40 S124 S133 S338 S512 S519 T143 T170 T339 T418 Y207	N31 N129 N283 N352		
16	3643847CD1	534	S22 S160 S164 S217 S289 S363 S217 S289 S363 S236 S639 S872 S725 S737 S741 S747 S757 T56 T67 T338 T378 T437 T483 T493 T625 T766 Y443 Y743 Y785	N54 N80 N85 N117 N205 N247 N349 N371		
17	6889872CD1	820				

WO 02/34783

PCT/US01/49670

Table 4

Polynucleotide Seq ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
18	643147CB1	2653	1-606, 2109-2306	6798827J1 (COLMOR03)	2076	2653
				6882379J1 (BRAHTR03)	1679	2342
				7162277F8 (PLACOR01)	1427	2173
				775207J1 (SPANT001)	418	1133
				6109577F6 (NCIDT03)	870	1464
				719064F8 (HUNGER04)	1	235
19	358465CB1	3531	1-1782	71160758V1	2094	2675
				71160758V1	2094	2675
				71161160V1	2597	3331
				7658349J1 (UPREME06)	1341	1389
				70815172V1	1481	2064
				71764416V1	2978	3531
				71754337V1	1965	2668
				7183593H1 (BONREFC01)	1	603
				7709623J1 (PANCNOE02)	895	1480
				6807068F6 (SKINOR01)	293	1156
20	3737084CB1	2280	947-997, 1840-2280	70409519D1	1372	1947
				70408370D1	1720	2280
				70400678D1	859	1371
				71862878V1 (NSANNOT01)	4282	1268
				7168507V1	609	748
				71683716V1	1	508
21	71426238CB1	1104	1-115, 550-624, 198-513	1947421R6 (PTUNOT01)	1	508
				71433875V1	435	1104
22	7475123CB1	4966	1-957, 1328-3163	70879833V1	4038	4682
				8124425H1 (HEANOC01)	4309	4966
				8253556H1 (BRAHDT10)	3948	4663
				6891485J1 (BRAHTR03)	2762	3423
				6885135H1	1	468
				593337R6 (BRAVINT02)	3439	3968
				70873072V1	3303	3965

WO 02/34783

PCT/US01/49670

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
22				GBL.97705145_17_12_25	337	3773
23	7481952CB1	5401	4237-4630, 4834- 3834, 1-2283, 4757-4851	27.7segscan () 55153856J1	1947	2415
				GNM.010120165_000001_1	1	342
				69338856 (SINTNR02)	2518	3102
				69338856 (SINTNR02)	2339	3011
				69338856 (SINTNR02)	4614	5401
				72434542D1	3010	3752
				55153330J1	1589	1891
				899395H1 (BRSTU03)	311	899
				55153261H1	4346	5105
				71876387V1	3716	4305
				71932270V1	364	1151
				55153293H1	886	1816
				55153285V1	4234	5099
				72432080D1	1678	2105
				GNM.99843576_000009_0		
24	382654CB1	1949	1-247, 1882-1949	723206H1 (BRAITDR03)	1	517
				66183492 (BRAITDR03)	1282	1949
				688426H1 (BRAITDR03)	541	1060
				6952509H1 (BRAITDR02)	1035	1600
				6891983H1 (BRAITDR03)	468	1056
				6891983J1 (BRAITDR03)	1121	1671
25	1867351CB1	2133	987-1049, 1927- 2133, 779-832	70571014V1	1520	2133
				70568556V1	1259	1924
				7261108H1 (UPREMO01)	646	1135
				8004078H1 (MUSCTDC01)	1	704
				70570876V1	1151	1828

WO 02/34783

PCT/US01/49670

Table 4 (cont.)

Polynucleotide SBO ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
25				5502673001 (ADMDNV22)	692	1215
26	3323104CB1	2090	1-34, 415-1111	7075976 (SYNORAT03)	1128	1767
				71220067V1	585	1119
				70833651V1	553	1114
				70769661V1	1061	1692
				8133288H1 (SCOMD1C01)	1	578
27	4769306CB1	1618	1-208, 561-723	70771538V1	1625	2050
				70954860V1	1126	1618
				2259667H1 (BRAINT02)	248	244
				8133288H1	585	1119
				8133247H1 (MIXDDE02)	229	792
				7076427H1 (BRAUTDR04)	49	677
28	2720056CB1	3269	2536-2564, 1- 220, 2785-3269	70953588V1	795	1368
				265750176 (LUNGUT09)	2502	3056
				71520348V1	1839	2547
				71524606V1	1139	1824
				265750176 (LUNGUT09)	594	1126
				7225843H1 (LUNGTRC01)	392	983
				71670642V1	1721	2424
				1305113F6 (PLACNT02)	1	571
				1955263F6 (CORNT01)	2570	3058
				3376653F6 (LUNGUT10)	267	3293
				71524606V1	2674	3114
29	7481255CB1	1227	1-1227	3269675H1 (BRAINT02)	5	86
				0806752436_00021_0	1	1227
				00024 edit	1	1227
				FL7481255_g8469082_00	99	304
				0035_g4150939_1_1	1	1227
30	1510242CB1	2618	1-515, 2475-2618	4538535F6 (TRYRMT01)	654	1227
				682312071 (SINTNOR01)	1889	2437
				703675877 (UTRSTMR02)	1	705

WO 02/34783

PCT/US01/49670

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
30				7355340HL (HEARNOR03)	2005	2618
				7930760HL (COLNLS02)	1224	1900
				6837255HL (BRSTNOR02)	673	1368
				70366002DL	1430	1904
				3413609HL (PHTNOR04)	2456	2618
				7180852VL	806	2432
31	162131CB1	2188	1-25, 1173-1633	6823120HL (SINTNOR01)	1820	2436
				9130335	1667	2188
				1818222 (BRSTNOR01)	1632	2188
				1988635R6 (BRSTNOR03)	454	2132
				70559145VL	1069	1787
				70558921VL	1069	1787
				2818222F6 (BRSTNOR14)	1	540
32	1837725CB1	1969	1-400, 1927-1969	70558931VL	549	1305
				70377507DL	1400	1955
				61378655	1517	1959
				70378490DL	639	1268
				67998857L (COLENOR03)	1	673
				8227475HL (BRAUTDR02)	778	1463
				2352377HL (COLSUC501)	1763	1959
33	3643847CB1	3006	796-1137, 2849-3006	1478479HL (CORPNOR02)	2745	3006
				6903355HL (BRAUTDR02)	1129	1755
				5824185HL (BRAUTDR02)	1772	2438
				3419330HL (KIDNOR09)	2650	2892
				6799167HL (BRACNOR01)	2091	2801
				6118619HL (BRACNOR01)	1932	2546
				7436931HL (ADREUOR02)	594	1202
				72018042VL	1	579
				72018222VL	413	1028
				1992275F6 (CORPNOR02)	1268	1809

WO 02/34783

PCT/US01/49670

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
34	6889872C81	2884	1-319, 2090- 2884, 803-1585, 379-418, 1991- 2017	8230901H1 (BRAUTDR02)	873	1610
				8230892H1 (BRAUTDR02)	1	725
				8230892H1 (BRAUTDR02)	676	1356
				7634710H1 (SMTDFE01)	1377	2482
				GM1:8231973-012	422	2884

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
18	6431478CB1	BRANDR03
19	3584654CB1	PROSNOT11
20	3737084CB1	PROSNON01
21	71426238CB1	PITUNOT01
22	7475123CB1	CONUT01
23	7481952CB1	SINTPR02
24	382654CB1	BRANDR03
25	1867351CB1	BRANDR03
26	3323104CB1	LINGNOT27
27	4769306CB1	BRANDR03
28	2720050CB1	LINGNOT10
29	7481255CB1	BRANDR03
30	1520442CB1	SINTPR02
31	162131CB1	PROSNOT15
32	1837725CB1	SINDNOT01
33	3643847CB1	CORENOT02
34	6889877CB1	SINTPR01

WO 02/34783

PCT/US01/49670

Table 6

Library	Vector	Library Description
BRAUTDR03	PCDNA2.1	This random primed library was constructed using RNA isolated from archaocortex, anterior hippocampus tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominately over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with associated metastatic disease. The patient had a long history of alcohol abuse, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.
BRAUTUT21	P1NCY	Library was constructed using RNA isolated from brain tumor tissue removed from the midline frontal lobe of a 61-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated subfrontal meningotheial meningioma with no atypia. One ethmoid and mucosal tissue sample indicated meningioma. Family history included cerebrovascular disease, senile dementia, hyperlipidemia, benign hypertension, atherosclerotic coronary artery disease, congestive heart failure, and breast cancer.
BRANUT01	P1NCY	Library was constructed using RNA isolated from medulla tissue removed from the brain of a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. Microscopically, the cerebral hemisphere revealed moderate fibrosis of the leptomeninges with focal calcifications. There was evidence of shrunken and slightly eosinophilic pyramidal neurons throughout the cerebral hemispheres. In addition, scattered throughout the cerebral cortex, there were small, dark, eosinophilic areas consistent with neurofibrillary tangles. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly and an enlarged spleen and liver.
COMUTUT01	P1NCY	Library was constructed using RNA isolated from sigmoid mesentery tumor tissue obtained from a 61-year-old female during a total abdominal hysterectomy and bilateral salpingo-oophorectomy with regional lymph node excision. Pathology indicated a metastatic grade 4 malignant mixed müllerian tumor present in the sigmoid mesentery at two sites.
CORNUOT02	P1NCY	Library was constructed using RNA isolated from diseased corpus callosum tissue removed from the brain of a 74-year-old Caucasian male who died from Alzheimer's disease.

WO 02/34783

PCT/US01/49670

Table 6 (cont.)

Library	Vector	Library Description
LUNGNOT27	PINCY	Library was constructed using RNA isolated from lung tissue removed from a 17-year-old Hispanic female.
LUNGNOT10	PINCY	Library was constructed using RNA isolated from lung tumor tissue removed from the left upper lobe of a 65-year-old Caucasian female during a segmental lung resection. Pathology indicated a metastatic grade 2 myxoid liposarcoma and a metastatic grade 4 liposarcoma. Patient history included soft tissue cancer, breast cancer, and secondary lung cancer.
PANCRNOT15	PINCY	Library was constructed using RNA isolated from diseased pancreatic tissue removed from a 16-year-old Caucasian male during an exploratory laparotomy with distal pancreatectomy and total splenectomy. Pathology indicated islet cell hyperplasia.
PITNOT01	PEMBESCRIP T	Family history included prostate cancer and cardiovascular disease. Library was constructed using RNA obtained from Clontech (CLON 6584-2, lot 35278). The RNA was isolated from the pituitary glands removed from a pool of 18 male and female Caucasian donors, 16 to 70 years old, who died from trauma.
PROSNOM01	PSPORT1	This normalized prostate library was constructed from 4.4 M independent clones from a prostate library. Starting RNA was made from prostate tissue removed from a 28-year-old Caucasian male who died from a self-inflicted gunshot wound. The normalized library was constructed using a self-inflicted gunshot wound. The normalized library was constructed using a self-inflicted gunshot wound. The normalized library was constructed using a self-inflicted gunshot wound. Acad. Sci. USA 91:9228-9232, using a longer (19 hour) reannealing hybridization period.
PROSNOT11	PINCY	Library was constructed using RNA isolated from the prostate tissue of a 28-year-old Caucasian male, who died from a self-inflicted gunshot wound.
SINDNOT01	PINCY	Library was constructed using RNA isolated from duodenum tissue removed from the small intestine of a 16-year-old Caucasian male who died from head trauma. Patient history included a kidney infection.
SINDTIE01	PCMR2.1	This 5' biased random primed library was constructed using RNA isolated from small intestine tissue removed from a 16-year-old Caucasian male who died from head trauma. Patient history included a kidney infection. The patient presented with acute stomach ulcer with obstruction, nausea and vomiting, and abnormal weight loss. Patient history included backache, acute stomach ulcer with perforation, and normal delivery. Previous surgeries included adenotomylectomy and total abdominal hysterectomy. Patient medications included Premarin. Family history included benign hypertension, type II diabetes and congestive heart failure in the father.
SINTFER02	PINCY	This random primed library was constructed using RNA isolated from small intestine tissue removed from a Caucasian male fetus who died from fetal demise.

WO 02/34783

PCT/US01/49670

Table 6 (cont.)

Library	Vector	Library Description
SIN7TR02	PCDNA2.1	This random primed library was constructed using RNA isolated from small intestine tissue removed from a 59-year-old male. Pathology for the matched tumor tissue indicated multiple (9) carcinoid tumors, grade 1, in the small bowel. The largest tumor was associated with a large mesenteric mass. Multiple convoluted segments of bowel were adhered to the tumor. A single (1 of 13) regional lymph node was positive for malignancy. The peritoneal biopsy indicated focal fat necrosis.
THYR01	pINCY	Library was constructed using RNA isolated from left thyroid tissue removed from a 56-year-old Caucasian male during a total thyroidectomy. The matched tumor tissue indicated multiple (9) carcinoid tumors, grade 1, in the thyroid gland. The largest tumor was associated with a large mesenteric mass. Multiple convoluted segments of bowel were adhered to the tumor. A single (1 of 13) regional lymph node was positive for malignancy. The peritoneal biopsy indicated focal fat necrosis. The matched tumor tissue indicated metastatic medullary carcinoma involving the overlying skeletal muscle. Metastatic medullary carcinoma involved one carotid sheath lymph node (of 9), one left neck lymph node with extra nodular extension, and a central compartment node. A microscopic focus of grade 1 papillary carcinoma was identified within the right lobe of the thyroid lobe. The left thyroid vein biopsy was negative for tumor. Patient history included hyperlipidemia, headache, and atherosclerotic coronary artery disease. Family history included cerebrovascular disease, cardiovascular disease and bone cancer.

WO 02/34783

PCT/US01/49670

Table 7

Program	Description	Reference	Parameter Threshold
ABI/FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, ffasta, fastx, fasty, and search.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.0E-6 Assembled ESTs: fasta.Identity= 95% or greater and Match length=200 bases or greater; fasta E value=1.0E-8 or less Full Length sequences: fasta score=100 or greater
BLIMPS	A BLOCKS IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Atwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1513; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nuclei, Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater

WO 02/34783

PCT/US01/49670

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfilesScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Batroch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score; GCG-specified "HIGRP" value for that particular Prosite motif Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequence 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 Phil's Revised Assembly Program including SWAT and CrossMatch programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence libraries and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match lengths= 56 or greater
Coarsed	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.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sommadossi, R. et al. (1998) Proc. Sixth Int. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds, The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 02/34783

PCT/US01/49670

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence selected from the group consisting of
SEQ ID NO:1-17,
 - b) a polypeptide comprising a naturally occurring amino acid sequence at least 90%
identical to an amino acid sequence selected from the group consisting of SEQ ID
NO:1-17,
 - c) a biologically active fragment of a polypeptide having an amino acid sequence selected
from the group consisting of SEQ ID NO:1-17, and
 - d) an immunogenic fragment of a polypeptide having an amino acid sequence selected
from the group consisting of SEQ ID NO:1-17.
2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the
group consisting of SEQ ID NO:1-17.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from
the group consisting of SEQ ID NO:18-34.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a
polynucleotide of claim 3.
7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method of producing a polypeptide of claim 1, the method comprising:
 - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein
said cell is transformed with a recombinant polynucleotide, and said recombinant
polynucleotide comprises a promoter sequence operably linked to a polynucleotide
encoding the polypeptide of claim 1, and

WO 02/34783

PCT/US01/49670

- b) recovering the polypeptide so expressed.

10. A method of claim 9, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-17.

5

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

12. An isolated polynucleotide selected from the group consisting of:

- 10 a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34,
 b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:18-34,
 c) a polynucleotide complementary to a polynucleotide of a),
 15 d) a polynucleotide complementary to a polynucleotide of b), and
 e) an RNA equivalent of a)-d).

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

20

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

- 25 a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
 b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

30

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

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

- 35 a) amplifying said target polynucleotide or fragment thereof using polymerase chain

WO 02/34783

PCT/US01/49670

- reaction amplification, and
- b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.
- 5 17. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.
18. A composition of claim 17, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-17.
- 10 19. A method for treating a disease or condition associated with decreased expression of functional TMP, comprising administering to a patient in need of such treatment the composition of claim 17.
- 15 20. A method of screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:
- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting agonist activity in the sample.
- 20 21. A composition comprising an agonist compound identified by a method of claim 20 and a pharmaceutically acceptable excipient.
22. A method for treating a disease or condition associated with decreased expression of functional TMP, comprising administering to a patient in need of such treatment a composition of claim
- 25 21.
23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:
- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- 30 b) detecting antagonist activity in the sample.
24. A composition comprising an antagonist compound identified by a method of claim 23 and a pharmaceutically acceptable excipient.
- 35 25. A method for treating a disease or condition associated with overexpression of functional

WO 02/34783

PCT/US01/49670

TMP, comprising administering to a patient in need of such treatment a composition of claim 24.

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

- 5 a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

10 27. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, the method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test
15 compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity
20 of the polypeptide of claim 1.

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

- 25 a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying
30 amounts of the compound and in the absence of the compound.

29. A method of assessing toxicity of a test compound, the method comprising:

- a) treating a biological sample containing nucleic acids with the test compound,
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising
35 at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target

WO 02/34783

PCT/US01/49670

- polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,
- c) quantifying the amount of hybridization complex, and
 - d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.
30. A diagnostic test for a condition or disease associated with the expression of TMP in a biological sample, the method comprising:
- a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and
 - b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.
31. The antibody of claim 11, wherein the antibody is:
- a) a chimeric antibody,
 - b) a single chain antibody,
 - c) a Fab fragment,
 - d) a F(ab')₂ fragment, or
 - e) a humanized antibody.
32. A composition comprising an antibody of claim 11 and an acceptable excipient.
33. A method of diagnosing a condition or disease associated with the expression of TMP in a subject, comprising administering to said subject an effective amount of the composition of claim 32.
34. A composition of claim 32, wherein the antibody is labeled.
35. A method of diagnosing a condition or disease associated with the expression of TMP in a subject, comprising administering to said subject an effective amount of the composition of claim 34.
36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

WO 02/34783

PCT/US01/49670

- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibodies from said animal, and
- 5 c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17.
37. A polyclonal antibody produced by a method of claim 36.
- 10 38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.
39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:
- 15 a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-17, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibody producing cells from the animal,
- c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
- 20 d) culturing the hybridoma cells, and
- e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17.
- 25 40. A monoclonal antibody produced by a method of claim 39.
41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.
- 30 42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.
43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.
- 35

WO 02/34783

PCT/US01/49670

44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 in a sample, the method comprising:

- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- 5 b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 in the sample.

45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-17 from a sample, the method comprising:

- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID
- 15 NO:1-17.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

- a) labeling the polynucleotides of the sample,
- b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex,
- 25 and
- c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or

30 polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

35

WO 02/34783

PCT/US01/49670

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is
5 completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a
10 nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

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

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

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

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

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

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

30 60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

35 63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

WO 02/34783

PCT/US01/49670

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.
65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.
- 5 66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.
67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
- 10 69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.
- 15 71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.
72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.
73. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:18.
- 20 74. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:19.
75. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:20.
- 25 76. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:21.
77. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:22.
78. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:23.
- 30 79. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:24.
80. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:25.
- 35 81. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:26.

WO 02/34783

PCT/US01/49670

82. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:27.
83. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:28.
- 5 84. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:29.
85. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:30.
86. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:31.
- 10 87. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:32.
88. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:33.
- 15 89. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:34.

WO 02/34783

PCT/US01/49670

<110> INCYTE GENOMICS, INC.

WARREN, Bridget A.

XU, Yuning

YUE, Henry

BATHA, Sajeew

BURFORD, Neil

GANDHI, Ameena R.

WALLA, Marinder K.

ARVIZU, Chandra

TANG, Y. Tom

LU, Dyung Aina M.

DUGGAN, Brendan M.

BAUGHN, Mariah R.

LEE, Ernestine A.

KHAN, Farrah A.

NGUYEN, Dannie B.

AZIMZAI, Yalda

YAO, Monique G.

LAL, Preeti G.

THANGAVELU, Kavitha

RAMKUMAR, Jayalaxmi

TRAN, Bao

DING, Li

AU-YOUNG, Janice

<120> TRANSMEMBRANE PROTEINS

<130> PF-0836 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/244,017; 60/252,855; 60/251,825; 60/255,085

<151> 2000-10-27; 2000-11-22; 2000-12-07; 2000-12-12

<160> 34

<170> PERL Program

<210> 1

<211> 461

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6431478CD1

<400> 1

Met Ser Ala Gln Cys Cys Ala Gly Gln Leu Ala Cys Cys Cys Gly

1 5 10 15

Ser Ala Gly Cys Ser Leu Cys Cys Asp Cys Cys Pro Arg Ile Arg

20 25 30

Gln Ser Leu Ser Thr Arg Phe Met Tyr Ala Leu Tyr Phe Ile Leu

35 40 45

Val Val Val Leu Cys Cys Ile Met Met Ser Thr Thr Val Ala His

50 55 60

WO 02/34783

PCT/US01/49670

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

<210> 2

WO 02/34783

PCT/US01/49670

<211> 879

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3584654CD1

<400> 2

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

WO 02/34783

PCT/US01/49670

350	355	360
Arg Asp Val Ser Lys Glu Asn Ser Gly Tyr Tyr Tyr Cys His Val		
365	370	375
Ser Leu Trp Ala Pro Gly His Asn Arg Ser Trp His Lys Val Ala		
380	385	390
Glu Ala Val Ser Ser Pro Ala Gly Val Gly Val Thr Trp Leu Glu		
395	400	405
Pro Asp Tyr Gln Val Tyr Leu Asn Ala Ser Lys Val Pro Gly Phe		
410	415	420
Ala Asp Asp Pro Thr Glu Leu Ala Cys Arg Val Val Asp Thr Lys		
425	430	435
Ser Gly Glu Ala Asn Val Arg Phe Thr Val Ser Trp Tyr Tyr Arg		
440	445	450
Met Asn Arg Arg Ser Asp Asn Val Val Thr Ser Glu Leu Leu Ala		
455	460	465
Val Met Asp Gly Asp Trp Thr Leu Lys Tyr Gly Glu Arg Ser Lys		
470	475	480
Gln Arg Ala Gln Asp Gly Asp Phe Ile Phe Ser Lys Glu His Thr		
485	490	495
Asp Thr Phe Asn Phe Arg Ile Gln Arg Thr Thr Glu Glu Asp Arg		
500	505	510
Gly Asn Tyr Tyr Cys Val Val Ser Ala Trp Thr Lys Gln Arg Asn		
515	520	525
Asn Ser Trp Val Lys Ser Lys Asp Val Phe Ser Lys Pro Val Asn		
530	535	540
Ile Phe Trp Ala Leu Glu Asp Ser Val Leu Val Val Lys Ala Arg		
545	550	555
Gln Pro Lys Pro Phe Phe Ala Ala Gly Asn Thr Phe Glu Met Thr		
560	565	570
Cys Lys Val Ser Ser Lys Asn Ile Lys Ser Pro Arg Tyr Ser Val		
575	580	585
Leu Ile Met Ala Glu Lys Pro Val Gly Asp Leu Ser Ser Pro Asn		
590	595	600
Glu Thr Lys Tyr Ile Ile Ser Leu Asp Gln Asp Ser Val Val Lys		
605	610	615
Leu Glu Asn Trp Thr Asp Ala Ser Arg Val Asp Gly Val Val Leu		
620	625	630
Glu Lys Val Gln Glu Asp Glu Phe Arg Tyr Arg Met Tyr Gln Thr		
635	640	645
Gln Val Ser Asp Ala Gly Leu Tyr Arg Cys Met Val Thr Ala Trp		
650	655	660
Ser Pro Val Arg Gly Ser Leu Trp Arg Glu Ala Ala Thr Ser Leu		
665	670	675
Ser Asn Pro Ile Glu Ile Asp Phe Gln Thr Ser Gly Pro Ile Phe		
680	685	690
Asn Ala Ser Val His Ser Asp Thr Pro Ser Val Ile Arg Gly Asp		
695	700	705
Leu Ile Lys Leu Phe Cys Ile Ile Thr Val Glu Gly Ala Ala Leu		
710	715	720
Asp Pro Asp Asp Met Ala Phe Asp Val Ser Trp Phe Ala Val His		
725	730	735
Ser Phe Gly Leu Asp Lys Ala Pro Val Leu Leu Ser Ser Leu Asp		
740	745	750
Arg Lys Gly Ile Val Thr Thr Ser Arg Arg Asp Trp Lys Ser Asp		
755	760	765
Leu Ser Leu Glu Arg Val Ser Val Leu Glu Phe Leu Leu Gln Val		

WO 02/34783

PCT/US01/49670

```

770      775      780
His Gly Ser Glu Asp Gln Asp Phe Gly Asn Tyr Tyr Cys Ser Val
785      790      795
Thr Pro Trp Val Lys Ser Pro Thr Gly Ser Trp Gln Lys Glu Ala
800      805      810
Glu Ile His Ser Lys Pro Val Phe Ile Thr Val Lys Met Asp Val
815      820      825
Leu Asn Ala Phe Lys Tyr Pro Leu Leu Ile Gly Val Gly Leu Ser
830      835      840
Thr Val Ile Gly Leu Leu Ser Cys Leu Ile Gly Tyr Cys Ser Ser
845      850      855
His Trp Cys Cys Lys Lys Glu Val Gln Glu Thr Arg Arg Glu Arg
860      865      870
Arg Arg Leu Met Ser Met Glu Met Asp
875

```

```

<210> 3
<211> 473
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3737084CD1

```

```

<400> 3
Met Ala Gln Leu Glu Gly Tyr Tyr Phe Ser Ala Ala Leu Ser Cys
1      5      10      15
Thr Phe Leu Val Ser Cys Leu Leu Phe Ser Ala Phe Ser Arg Ala
20      25      30
Leu Arg Glu Pro Tyr Met Asp Glu Ile Phe His Leu Pro Gln Ala
35      40      45
Gln Arg Tyr Cys Glu Gly His Phe Ser Leu Ser Gln Trp Asp Pro
50      55      60
Met Ile Thr Thr Leu Pro Gly Leu Tyr Leu Val Ser Ile Gly Val
65      70      75
Ile Lys Pro Ala Ile Trp Ile Phe Gly Trp Ser Glu His Val Val
80      85      90
Cys Ser Ile Gly Met Leu Arg Phe Val Asn Leu Leu Phe Ser Val
95      100     105
Gly Asn Phe Tyr Leu Leu Tyr Leu Leu Phe Cys Lys Val Gln Pro
110     115     120
Arg Asn Lys Ala Ala Ser Ser Ile Gln Arg Val Leu Ser Thr Leu
125     130     135
Thr Leu Ala Val Phe Pro Thr Leu Tyr Phe Phe Asn Phe Leu Tyr
140     145     150
Tyr Thr Glu Ala Gly Ser Met Phe Phe Thr Leu Phe Ala Tyr Leu
155     160     165
Met Cys Leu Tyr Gly Asn His Lys Thr Ser Ala Phe Leu Gly Phe
170     175     180
Cys Gly Phe Met Phe Arg Gln Thr Asn Ile Ile Trp Ala Val Phe
185     190     195
Cys Ala Gly Asn Val Ile Ala Gln Lys Leu Thr Glu Ala Trp Lys
200     205     210
Thr Glu Leu Gln Lys Lys Glu Asp Arg Leu Pro Pro Ile Lys Gly
215     220     225

```

WO 02/34783

PCT/US01/49670

```

Pro Phe Ala Glu Phe Arg Lys Ile Leu Gln Phe Leu Leu Ala Tyr
230 235 240
Ser Met Ser Phe Lys Asn Leu Ser Met Leu Leu Leu Thr Trp
245 250 255
Pro Tyr Ile Leu Leu Gly Phe Leu Phe Cys Ala Phe Val Val Val
260 265 270
Asn Gly Gly Ile Val Ile Gly Asp Arg Ser Ser His Glu Ala Cys
275 280 285
Leu His Phe Pro Gln Leu Phe Tyr Phe Phe Ser Phe Thr Leu Phe
290 295 300
Phe Ser Phe Pro His Leu Leu Ser Pro Ser Lys Ile Lys Thr Phe
305 310 315
Leu Ser Leu Val Trp Lys Arg Arg Ile Leu Phe Phe Val Val Thr
320 325 330
Leu Val Ser Val Phe Leu Val Trp Lys Phe Thr Tyr Ala His Lys
335 340 345
Tyr Leu Leu Ala Asp Asn Arg His Tyr Thr Phe Tyr Val Trp Lys
350 355 360
Arg Val Phe Gln Arg Tyr Glu Thr Val Lys Tyr Leu Leu Val Pro
365 370 375
Ala Tyr Ile Phe Ala Gly Trp Ser Ile Ala Asp Ser Leu Lys Ser
380 385 390
Lys Ser Ile Phe Trp Asn Leu Met Phe Phe Ile Cys Leu Phe Thr
395 400 405
Val Ile Val Pro Gln Lys Leu Leu Glu Phe Arg Tyr Phe Ile Leu
410 415 420
Pro Tyr Val Ile Tyr Arg Leu Asn Ile Pro Leu Pro Pro Thr Ser
425 430 435
Arg Leu Ile Cys Glu Leu Ser Cys Tyr Ala Val Val Asn Phe Ile
440 445 450
Thr Phe Phe Ile Phe Leu Asn Lys Thr Phe Gln Trp Pro Asn Ser
455 460 465
Gln Asp Ile Gln Arg Phe Met Trp
470

```

```

<210> 4
<211> 223
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 71426238CD1

```

```

<400> 4
Met Ser Trp Met Phe Leu Arg Asp Leu Leu Ser Gly Val Asn Lys
1 5 10 15
Tyr Ser Thr Gly Ile Gly Trp Ile Trp Leu Ala Val Val Phe Val
20 25 30
Phe Arg Leu Leu Val Tyr Met Val Ala Ala Glu His Val Trp Lys
35 40 45
Asp Glu Gln Lys Glu Phe Glu Cys Asn Ser Arg Gln Pro Gly Cys
50 55 60
Lys Asn Val Cys Phe Asp Asp Phe Phe Pro Ile Ser Gln Val Arg
65 70 75
Leu Trp Ala Leu Gln Leu Ile Met Val Ser Thr Pro Ser Leu Leu

```

WO 02/34783

PCT/US01/49670

Val Val Leu His	80	85	90
Val Ala Tyr His Glu Gly Arg Glu Lys Arg His			
Arg Lys Lys Leu Tyr Val Ser Pro Gly Thr Met Asp Gly Gly Leu	95	100	105
Trp Tyr Ala Tyr Leu Ile Ser Leu Ile Val Lys Thr Gly Phe Glu	110	115	120
Ile Gly Phe Leu Val Leu Phe Tyr Lys Leu Tyr Asp Gly Phe Ser	125	130	135
Val Pro Tyr Leu Ile Lys Cys Asp Leu Lys Pro Cys Pro Asn Thr	140	145	150
Val Asp Cys Phe Ile Ser Lys Pro Thr Glu Lys Thr Ile Phe Ile	155	160	165
Leu Phe Leu Val Ile Thr Ser Cys Leu Cys Ile Val Leu Asn Phe	170	175	180
Ile Glu Leu Ser Phe Leu Val Leu Lys Cys Leu Ile Lys Cys Cys	185	190	195
Leu Gln Lys Tyr Leu Lys Lys Pro Gln Val Leu Ser Val	200	205	210
	215	220	

<210> 5

<211> 1553

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7475123CD1

<400> 5

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

WO 02/34783

PCT/US01/49670

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

WO 02/34783

PCT/US01/49670

```

Ile Gln Ser Ala Asp Arg Leu Thr Gly Leu Asp Arg Glu Thr Leu
620 625 630
Gln Glu Asn Leu Glu Asn Leu Met Asp Ile His Val Phe His Arg
635 640 645
Arg Arg Pro Pro Val Ser Thr Pro Cys Ala Met Glu Asn Gly Gly
650 655 660
Cys Ser His Leu Cys Leu Arg Ser Pro Asn Pro Ser Gly Phe Ser
665 670 675
Cys Thr Cys Pro Thr Gly Ile Asn Leu Leu Ser Asp Gly Lys Thr
680 685 690
Cys Ser Pro Gly Met Asn Ser Phe Leu Ile Phe Ala Arg Arg Ile
695 700 705
Asp Ile Arg Met Val Ser Leu Asp Ile Pro Tyr Phe Ala Asp Val
710 715 720
Val Val Pro Ile Asn Ile Thr Met Lys Asn Thr Ile Ala Ile Gly
725 730 735
Val Asp Pro Gln Glu Gly Lys Val Tyr Trp Ser Asp Ser Thr Leu
740 745 750
His Arg Ile Ser Arg Ala Asn Leu Asp Gly Ser Gln His Glu Asp
755 760 765
Ile Ile Thr Thr Gly Leu Gln Thr Thr Asp Gly Leu Ala Val Asp
770 775 780
Ala Ile Gly Arg Lys Val Tyr Trp Thr Asp Thr Gly Thr Asn Arg
785 790 795
Ile Glu Val Gly Asn Leu Asp Gly Ser Met Arg Lys Val Leu Val
800 805 810
Trp Gln Asn Leu Asp Ser Pro Arg Ala Ile Val Leu Tyr His Glu
815 820 825
Met Gly Phe Met Tyr Trp Thr Asp Trp Gly Glu Asn Ala Lys Leu
830 835 840
Glu Arg Ser Gly Met Asp Gly Ser Asp Arg Ala Val Leu Ile Asn
845 850 855
Asn Asn Leu Gly Trp Pro Asn Gly Leu Thr Val Asp Lys Ala Ser
860 865 870
Ser Gln Leu Leu Trp Ala Asp Ala His Thr Glu Arg Ile Glu Ala
875 880 885
Ala Asp Leu Asn Gly Ala Asn Arg His Thr Leu Val Ser Pro Val
890 895 900
Gln His Pro Tyr Gly Leu Thr Leu Leu Asp Ser Tyr Ile Tyr Trp
905 910 915
Thr Asp Trp Gln Thr Arg Ser Ile His Arg Ala Asp Lys Gly Thr
920 925 930
Gly Ser Asn Val Ile Leu Val Arg Ser Asn Leu Pro Gly Leu Met
935 940 945
Asp Met Gln Ala Val Asp Arg Ala Gln Pro Leu Gly Phe Asn Lys
950 955 960
Cys Gly Ser Arg Asn Gly Gly Cys Ser His Leu Cys Leu Pro Arg
965 970 975
Pro Ser Gly Phe Ser Cys Ala Cys Pro Thr Gly Ile Gln Leu Lys
980 985 990
Gly Asp Gly Lys Thr Cys Asp Pro Ser Pro Glu Thr Tyr Leu Leu
995 1000 1005
Phe Ser Ser Arg Gly Ser Ile Arg Arg Ile Ser Leu Asp Thr Ser
1010 1015 1020
Asp His Thr Asp Val His Val Pro Val Pro Glu Leu Asn Asn Val
1025 1030 1035

```

WO 02/34783

PCT/US01/49670

Ile Ser Leu Asp Tyr Asp Ser Val Asp Gly Lys Val Tyr Tyr Thr
 1040 1045 1050
 Asp Val Phe Leu Asp Val Ile Arg Arg Ala Asp Leu Asn Gly Ser
 1055 1060 1065
 Asn Met Glu Thr Val Ile Gly Arg Gly Leu Lys Thr Thr Asp Gly
 1070 1075 1080
 Leu Ala Val Asp Trp Val Ala Arg Asn Leu Tyr Trp Thr Asp Thr
 1085 1090 1095
 Gly Arg Asn Thr Ile Glu Ala Ser Arg Leu Asp Gly Ser Cys Arg
 1100 1105 1110
 Lys Val Leu Ile Asn Asn Ser Leu Asp Glu Pro Arg Ala Ile Ala
 1115 1120 1125
 Val Phe Pro Arg Lys Gly Tyr Leu Phe Trp Thr Asp Trp Gly His
 1130 1135 1140
 Ile Ala Lys Ile Glu Arg Ala Asn Leu Asp Gly Ser Glu Arg Lys
 1145 1150 1155
 Val Leu Ile Asn Thr Asp Leu Gly Trp Pro Asn Gly Leu Thr Leu
 1160 1165 1170
 Asp Tyr Asp Thr Arg Arg Ile Tyr Trp Val Asp Ala His Leu Asp
 1175 1180 1185
 Arg Ile Glu Ser Ala Asp Leu Asn Gly Lys Leu Arg Gln Val Leu
 1190 1195 1200
 Val Ser His Val Ser His Pro Phe Ala Leu Thr Gln Gln Asp Arg
 1205 1210 1215
 Trp Ile Tyr Trp Thr Asp Trp Gln Thr Lys Ser Ile Gln Arg Val
 1220 1225 1230
 Asp Lys Tyr Ser Gly Arg Asn Lys Glu Thr Val Leu Ala Asn Val
 1235 1240 1245
 Glu Gly Leu Met Asp Ile Ile Val Val Ser Pro Gln Arg Gln Thr
 1250 1255 1260
 Gly Thr Asn Ala Cys Gly Val Asn Asn Gly Gly Cys Thr His Leu
 1265 1270 1275
 Cys Phe Ala Arg Ala Ser Asp Phe Val Cys Ala Cys Pro Asp Glu
 1280 1285 1290
 Pro Asp Ser Gln Pro Cys Ser Leu Val Pro Gly Leu Val Pro Pro
 1295 1300 1305
 Ala Pro Arg Ala Thr Gly Met Ser Glu Lys Ser Pro Val Leu Pro
 1310 1315 1320
 Asn Thr Pro Pro Thr Thr Leu Tyr Ser Ser Thr Thr Arg Thr Arg
 1325 1330 1335
 Thr Ser Leu Glu Glu Val Glu Gly Arg Cys Ser Glu Arg Asp Ala
 1340 1345 1350
 Arg Leu Gly Leu Cys Ala Arg Ser Asn Asp Ala Val Pro Ala Ala
 1355 1360 1365
 Pro Gly Glu Gly Leu His Ile Ser Tyr Ala Ile Gly Gly Leu Leu
 1370 1375 1380
 Ser Ile Leu Leu Ile Leu Val Val Ile Ala Ala Leu Met Leu Tyr
 1385 1390 1395
 Arg His Lys Lys Ser Lys Phe Thr Asp Pro Gly Met Gly Asn Leu
 1400 1405 1410
 Thr Tyr Ser Asn Pro Ser Tyr Arg Thr Ser Thr Gln Glu Val Lys
 1415 1420 1425
 Ile Glu Ala Ile Pro Lys Pro Ala Met Tyr Asn Gln Leu Cys Tyr
 1430 1435 1440
 Lys Lys Glu Gly Gly Pro Asp His Asn Tyr Thr Lys Glu Lys Ile
 1445 1450 1455

WO 02/34783

PCT/US01/49670

Lys Ile Val Glu Gly Ile Cys Leu Leu Ser Gly Asp Asp Ala Glu
 1460 1465 1470
 Trp Asp Asp Leu Lys Gln Leu Arg Ser Ser Arg Gly Gly Leu Leu
 1475 1480 1485
 Arg Asp His Val Cys Met Lys Thr Asp Thr Val Ser Ile Gln Ala
 1490 1495 1500
 Ser Ser Gly Ser Leu Asp Asp Thr Glu Thr Glu Gln Leu Leu Gln
 1505 1510 1515
 Glu Glu Gln Ser Glu Cys Ser Ser Val His Thr Ala Ala Thr Pro
 1520 1525 1530
 Glu Arg Arg Gly Ser Leu Pro Asp Thr Gly Trp Lys His Glu Arg
 1535 1540 1545
 Lys Leu Ser Ser Glu Ser Gln Val
 1550

<210> 6

<211> 1718

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7481952CD1

<400> 6

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

WO 02/34783

PCT/US01/49670

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

WO 02/34783

PCT/US01/49670

	650	655	660
Lys Trp Tyr Gln	Pro Ile Pro Val His	Leu Leu Gln Asp Ser	Asn
	665	670	675
Thr Phe Arg Trp	Gly Leu Gly Asn Gly	Ile Ser Ile His His	Gly
	680	685	690
Glu Glu Asn His	Arg Pro Ser Val Asp	His Thr Gln Asn Thr	Thr
	695	700	705
Asp Gly Trp Tyr	Leu Tyr Ala Asp Ser	Ser Asn Gly Lys Phe	Gly
	710	715	720
Asp Thr Ala Asp	Ile Leu Thr Pro Ile	Ile Ser Leu Thr Gly	Pro
	725	730	735
Lys Cys Thr Leu	Val Phe Trp Thr His	Met Asn Gly Ala Thr	Val
	740	745	750
Gly Ser Leu Gln	Val Leu Ile Lys Lys	Asp Asn Val Thr Ser	Lys
	755	760	765
Leu Trp Ala Gln	Thr Gly Gln Gln Gly	Ala Gln Trp Lys Arg	Ala
	770	775	780
Glu Val Phe Leu	Gly Ile Arg Ser His	Thr Gln Ile Val Phe	Arg
	785	790	795
Ala Lys Arg Gly	Ile Ser Tyr Ile Gly	Asp Val Ala Val Asp	Asp
	800	805	810
Ile Ser Phe Gln	Asp Cys Ser Pro Leu	Leu Ser Pro Glu Arg	Lys
	815	820	825
Cys Thr Asp His	Glu Phe Met Cys Ala	Asn Lys His Cys Ile	Ala
	830	835	840
Lys Asp Lys Leu	Cys Asp Phe Val Asn	Asp Cys Ala Asp Asn	Ser
	845	850	855
Asp Glu Thr Thr	Phe Ile Cys Arg Thr	Ser Ser Gly Arg Cys	Asp
	860	865	870
Phe Glu Phe Asp	Leu Cys Ser Trp Lys	Gln Glu Lys Asp Glu	Asp
	875	880	885
Phe Asp Trp Asn	Leu Lys Ala Ser Ser	Ile Pro Ala Ala Gly	Thr
	890	895	900
Glu Pro Ala Ala	Asp His Thr Leu Gly	Asn Ser Ser Gly His	Tyr
	905	910	915
Ile Phe Ile Lys	Ser Leu Phe Pro Gln	Gln Pro Met Arg Ala	Ala
	920	925	930
Arg Ile Ser Ser	Pro Val Ile Ser Lys	Arg Ser Lys Asn Cys	Lys
	935	940	945
Ile Ile Phe His	Tyr His Met Tyr Gly	Asn Gly Ile Gly Ala	Leu
	950	955	960
Thr Leu Met Gln	Val Ser Val Thr Asn	Gln Thr Lys Val Leu	Leu
	965	970	975
Asn Leu Thr Val	Glu Gln Gly Asn Phe	Trp Arg Arg Glu Glu	Leu
	980	985	990
Ser Leu Phe Gly	Asp Glu Asp Phe Gln	Leu Lys Phe Glu Gly	Arg
	995	1000	1005
Val Gly Lys Gly	Gln Arg Gly Asp Ile	Ala Leu Asp Asp Ile	Val
	1010	1015	1020
Leu Thr Glu Asn	Cys Leu Ser Leu His	Asp Ser Val Gln Glu	Glu
	1025	1030	1035
Leu Ala Val Pro	Leu Pro Thr Gly Phe	Cys Pro Leu Gly Tyr	Arg
	1040	1045	1050
Glu Cys His Asn	Gly Lys Cys Tyr Arg	Leu Glu Gln Ser Cys	Asn
	1055	1060	1065
Phe Val Asp Asn	Cys Gly Asp Asn Thr	Asp Glu Asn Glu Cys	Gly

WO 02/34783

PCT/US01/49670

1070 1075 1080
Ser Ser Cys Thr Phe Glu Lys Gly Trp Cys Gly Trp Gln Asn Ser
1085 1090 1095
Gln Ala Asp Asn Phe Asp Trp Val Leu Gly Val Gly Ser His Gln
1100 1105 1110
Ser Leu Arg Pro Pro Lys Asp His Thr Leu Gly Asn Glu Asn Gly
1115 1120 1125
His Phe Met Tyr Leu Glu Ala Thr Ala Val Gly Leu Arg Gly Asp
1130 1135 1140
Lys Ala His Phe Arg Ser Thr Met Trp Arg Glu Ser Ser Ala Ala
1145 1150 1155
Cys Thr Met Ser Phe Trp Tyr Phe Ile Ser Ala Lys Ala Thr Gly
1160 1165 1170
Ser Ile Gln Ile Leu Ile Lys Thr Glu Lys Gly Leu Ser Lys Val
1175 1180 1185
Trp Gln Glu Ser Lys Gln Asn Pro Gly Asn His Trp Gln Lys Ala
1190 1195 1200
Asp Ile Leu Leu Gly Lys Leu Arg Asn Phe Glu Val Ile Phe Gln
1205 1210 1215
Gly Ile Arg Thr Arg Asp Leu Gly Gly Gly Ala Ala Ile Asp Asp
1220 1225 1230
Ile Glu Phe Lys Asn Cys Thr Thr Val Gly Glu Ile Ser Glu Leu
1235 1240 1245
Cys Pro Glu Ile Thr Asp Phe Leu Cys Arg Asp Lys Lys Cys Ile
1250 1255 1260
Ala Ser His Leu Leu Cys Asp Tyr Lys Pro Asp Cys Ser Asp Arg
1265 1270 1275
Ser Asp Glu Ala His Cys Ala His Tyr Thr Ser Thr Thr Gly Ser
1280 1285 1290
Cys Asn Phe Glu Thr Ser Ser Gly Asn Trp Thr Thr Ala Cys Ser
1295 1300 1305
Leu Thr Gln Asp Ser Glu Asp Asp Leu Asp Trp Ala Ile Gly Ser
1310 1315 1320
Arg Ile Pro Ala Lys Ala Leu Ile Pro Asp Ser Asp His Thr Pro
1325 1330 1335
Gly Ser Gly Gln His Phe Leu Tyr Val Asn Ser Ser Gly Ser Lys
1340 1345 1350
Glu Gly Ser Val Ala Arg Ile Thr Thr Ser Lys Ser Phe Pro Ala
1355 1360 1365
Ser Leu Gly Met Cys Thr Val Arg Phe Trp Phe Tyr Met Ile Asp
1370 1375 1380
Pro Arg Ser Met Gly Ile Leu Lys Val Tyr Thr Ile Glu Glu Ser
1385 1390 1395
Gly Leu Asn Ile Leu Val Trp Ser Val Ile Gly Asn Lys Arg Thr
1400 1405 1410
Gly Trp Thr Tyr Gly Ser Val Pro Leu Ser Ser Asn Ser Pro Phe
1415 1420 1425
Lys Val Ala Phe Glu Ala Asp Leu Asp Gly Asn Glu Asp Ile Phe
1430 1435 1440
Ile Ala Leu Asp Asp Ile Ser Phe Thr Pro Glu Cys Val Thr Gly
1445 1450 1455
Gly Pro Val Pro Val Gln Pro Ser Pro Cys Glu Ala Asp Gln Phe
1460 1465 1470
Ser Cys Ile Tyr Thr Leu Gln Cys Val Pro Leu Ser Gly Lys Cys
1475 1480 1485
Asp Gly His Glu Asp Cys Ile Asp Gly Ser Asp Glu Met Asp Cys

WO 02/34783

PCT/US01/49670

1490 1495 1500
Pro Leu Ser Pro Thr Pro Pro Leu Cys Ser Asn Met Glu Phe Pro
1505 1510 1515
Cys Ser Thr Asp Glu Cys Ile Pro Ser Leu Leu Cys Asp Gly
1520 1525 1530
Val Pro Asp Cys His Phe Asn Glu Asp Glu Leu Ile Cys Ser Asn
1535 1540 1545
Lys Ser Cys Ser Asn Gly Ala Leu Val Cys Ala Ser Ser Asn Ser
1550 1555 1560
Cys Ile Pro Ala His Gln Arg Cys Asp Gly Phe Ala Asp Cys Met
1565 1570 1575
Asp Phe Gln Leu Asp Glu Ser Ser Cys Ser Glu Cys Pro Leu Asn
1580 1585 1590
Tyr Cys Arg Asn Gly Gly Thr Cys Val Val Glu Lys Asn Gly Pro
1595 1600 1605
Met Cys Arg Cys Arg Gln Gly Trp Lys Gly Asn Arg Cys His Ile
1610 1615 1620
Lys Phe Asn Pro Pro Ala Thr Asp Phe Thr Tyr Ala Gln Asn Asn
1625 1630 1635
Thr Trp Thr Leu Leu Gly Ile Gly Leu Ala Phe Leu Met Thr His
1640 1645 1650
Ile Thr Val Ala Val Leu Cys Phe Leu Ala Asn Arg Lys Val Pro
1655 1660 1665
Ile Arg Lys Thr Glu Gly Ser Gly Asn Cys Ala Phe Val Asn Pro
1670 1675 1680
Val Tyr Gly Asn Trp Ser Asn Pro Glu Lys Thr Glu Ser Ser Val
1685 1690 1695
Tyr Ser Phe Ser Asn Pro Leu Tyr Gly Thr Thr Ser Gly Ser Leu
1700 1705 1710
Glu Thr Leu Ser His His Leu Lys
1715

<210> 7

<211> 224

<212> FRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 382654CD1

<400> 7

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

WO 02/34783

PCT/US01/49670

```

Arg Ser Gly Glu Gln Arg Lys Asn Pro Met Ala Leu Tyr Ile
110 115 120
Leu Lys Asp Lys Asp Ser Pro Glu Thr Glu Glu Asn Pro Ala Pro
125 130 135
Glu Pro Arg Ser Ala Thr Glu Pro Gly Pro Gly Tyr Ser Val
140 145 150
Ser Pro Ala Val Pro Gly Arg Ser Pro Gly Leu Pro Ile Arg Ser
155 160 165
Ala Arg Arg Tyr Pro Arg Ser Pro Ala Arg Ser Pro Ala Thr Gly
170 175 180
Arg Thr His Ser Ser Pro Pro Arg Ala Pro Ser Ser Pro Gly Arg
185 190 195
Ser Arg Ser Ala Ser Arg Thr Leu Arg Thr Ala Gly Val His Ile
200 205 210
Ile Arg Glu Gln Asp Glu Ala Gly Pro Val Glu Ile Ser Ala
215 220

```

<210> 8

<211> 570

<212> FRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1867351CD1

<400> 8

```

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

```

WO 02/34783

PCT/US01/49670

```

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

```

```

<210> 9
<211> 423
<212> PPT
<213> Homo sapiens

<220>
<221> misc_feature

```

WO 02/34783

PCT/US01/49670

<223> Incyte ID No: 3323104CD1

<400> 9

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

WO 02/34783

PCT/US01/49670

```

          395          400          405
Gln Ile Ile Ser His Tyr Lys Glu Glu Pro Leu Thr Glu Arg Ile
          410          415          420
Lys Tyr Asp

```

```

<210> 10
<211> 388
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4769306CD1

```

```

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

```

WO 02/34783

PCT/US01/49670

```

Asn Leu Ser Thr Arg Asp Ala Ser Ile Leu Thr Glu Met Pro Thr
  305          310          315
Gln Pro Gly Lys Pro Leu Ile Pro Pro Met Asn Arg Arg Asn Ile
  320          325          330
Leu Val Met Lys His Asn Tyr Ser Gln Asp Ala Ala Asp Ala Cys
  335          340          345
Asp Ile Asp Glu Ile Glu Glu Val Pro Thr Thr Ser His Arg Leu
  350          355          360
Ser Arg His Asp Lys Ala Val Gln Arg Phe Cys Leu Ile Gly Ser
  365          370          375
Leu Ser Lys His Glu Ser Glu Tyr Asn Thr Thr Arg Val
  380          385

```

```

<210> 11
<211> 231
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2720058CD1

```

```

<400> 11
Met Ala Phe Val Pro Phe Leu Leu Val Thr Trp Ser Ser Ala Ala
  1           5           10          15
Phe Ile Ile Ser Tyr Val Val Ala Val Leu Ser Gly His Val Asn
  20          25          30
Pro Phe Leu Pro Tyr Ile Ser Asp Thr Gly Thr Thr Pro Pro Glu
  35          40          45
Ser Gly Ile Phe Gly Phe Met Ile Asn Phe Ser Ala Phe Leu Gly
  50          55          60
Ala Ala Thr Met Tyr Thr Arg Tyr Lys Ile Val Gln Lys Gln Asn
  65          70          75
Gln Thr Cys Tyr Phe Ser Thr Pro Val Phe Asn Leu Val Ser Leu
  80          85          90
Val Leu Gly Leu Val Gly Cys Phe Gly Met Gly Ile Val Ala Asn
  95          100         105
Phe Gln Glu Leu Ala Val Pro Val Val His Asp Gly Gly Ala Leu
  110         115         120
Leu Ala Phe Val Cys Gly Val Val Tyr Thr Leu Leu Gln Ser Ile
  125         130         135
Ile Ser Tyr Lys Ser Cys Pro Gln Trp Asn Ser Leu Ser Thr Cys
  140         145         150
His Ile Arg Met Val Ile Ser Ala Val Ser Cys Ala Ala Val Ile
  155         160         165
Pro Met Ile Val Cys Ala Ser Leu Ile Ser Ile Thr Lys Leu Glu
  170         175         180
Trp Asn Pro Arg Glu Lys Asp Tyr Val Tyr His Val Val Ser Ala
  185         190         195
Ile Cys Glu Trp Thr Val Ala Phe Gly Phe Ile Phe Tyr Phe Leu
  200         205         210
Thr Phe Ile Gln Asp Phe Gln Ser Val Thr Leu Arg Ile Ser Thr
  215         220         225
Glu Ile Asn Gly Asp Ile
  230

```

WO 02/34783

PCT/US01/49670

<210> 12
<211> 293
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7481255CD1

<400> 12
Met Asp Arg Ala Lys Gln Gln Gln Ala Leu Leu Leu Leu Pro Val
1 5 10 15
Cys Leu Ala Leu Thr Phe Ser Leu Thr Ala Val Val Ser Ser His
20 25 30
Trp Cys Glu Gly Thr Arg Arg Val Val Lys Pro Leu Cys Gln Asp
35 40 45
Gln Pro Gly Gly Gln His Cys Ile His Phe Lys Arg Asp Asn Ser
50 55 60
Ser Asn Gly Arg Met Asp Asn Asn Ser Gln Ala Val Leu Tyr Ile
65 70 75
Trp Glu Leu Gly Asp Asp Lys Phe Ile Gln Arg Gly Phe His Val
80 85 90
Gly Leu Trp Gln Ser Cys Glu Glu Ser Leu Asn Gly Glu Asp Glu
95 100 105
Lys Cys Arg Ser Phe Arg Ser Val Val Pro Ala Glu Glu Gln Gly
110 115 120
Val Leu Trp Leu Ser Ile Gly Gly Glu Val Leu Asp Ile Val Leu
125 130 135
Ile Leu Thr Ser Ala Ile Leu Leu Gly Ser Arg Val Ser Cys Arg
140 145 150
Ser Pro Gly Phe His Trp Leu Arg Val Asp Ala Leu Val Ala Ile
155 160 165
Phe Met Val Leu Ala Gly Leu Leu Gly Met Val Ala His Met Met
170 175 180
Tyr Thr Thr Ile Phe Gln Ile Thr Val Asn Leu Gly Pro Glu Asp
185 190 195
Trp Lys Pro Gln Thr Trp Asp Tyr Gly Trp Ser Tyr Cys Leu Ala
200 205 210
Trp Gly Ser Phe Ala Leu Cys Leu Ala Val Ser Val Ser Ala Met
215 220 225
Ser Arg Phe Thr Ala Ala Arg Leu Glu Phe Thr Glu Lys Gln Gln
230 235 240
Ala Gln Asn Gly Ser Arg His Ser Gln His Ser Phe Leu Glu Pro
245 250 255
Glu Ala Ser Glu Ser Ile Trp Lys Thr Gly Ala Ala Pro Cys Pro
260 265 270
Ala Glu Gln Ala Phe Arg Asn Val Ser Gly His Leu Pro Pro Gly
275 280 285
Ala Pro Gly Lys Val Ser Ile Cys
290

<210> 13
<211> 526
<212> PRT
<213> Homo sapiens

WO 02/34783

PCT/US01/49670

<220>
<221> misc_feature
<223> Incyte ID No: 1510242CD1

<400> 13
Met Leu Thr Tyr Gly Val Tyr Leu Gly Leu Leu Gln Met Gln Leu
1 5 10 15
Ile Leu His Tyr Asp Glu Thr Tyr Arg Glu Val Lys Tyr Gly Asn
20 25 30
Met Gly Leu Pro Asp Ile Asp Ser Lys Met Leu Met Gly Ile Asn
35 40 45
Val Thr Pro Ile Ala Ala Leu Leu Tyr Thr Pro Val Leu Ile Arg
50 55 60
Phe Phe Gly Thr Lys Trp Met Met Phe Leu Ala Val Gly Ile Tyr
65 70 75
Ala Leu Phe Val Ser Thr Asn Tyr Trp Glu Arg Tyr Tyr Thr Leu
80 85 90
Val Pro Ser Ala Val Ala Leu Gly Met Ala Ile Val Pro Leu Trp
95 100 105
Ala Ser Met Gly Asn Tyr Ile Thr Arg Met Ala Gln Lys Tyr His
110 115 120
Glu Tyr Ser His Tyr Lys Glu Gln Asp Gly Gln Gly Met Lys Gln
125 130 135
Arg Pro Pro Arg Gly Ser His Ala Pro Tyr Leu Leu Val Phe Gln
140 145 150
Ala Ile Phe Tyr Ser Phe Phe His Leu Ser Phe Ala Cys Ala Gln
155 160 165
Leu Pro Met Ile Tyr Phe Leu Asn His Tyr Leu Tyr Asp Leu Asn
170 175 180
His Thr Leu Tyr Asn Val Gln Ser Cys Gly Thr Asn Ser His Gly
185 190 195
Ile Leu Ser Gly Phe Asn Lys Thr Val Leu Arg Thr Leu Pro Arg
200 205 210
Ser Gly Asn Leu Ile Val Val Glu Ser Val Leu Met Ala Val Ala
215 220 225
Phe Leu Ala Met Leu Leu Val Leu Gly Leu Cys Gly Ala Ala Tyr
230 235 240
Arg Pro Thr Glu Glu Ile Asp Leu Arg Ser Val Gly Trp Gly Asn
245 250 255
Ile Phe Gln Leu Pro Phe Lys His Val Arg Asp Tyr Arg Leu Arg
260 265 270
His Leu Val Pro Phe Phe Ile Tyr Ser Gly Phe Glu Val Leu Phe
275 280 285
Ala Cys Thr Gly Ile Ala Leu Gly Tyr Gly Val Cys Ser Val Gly
290 295 300
Leu Glu Arg Leu Ala Tyr Leu Leu Val Ala Tyr Ser Leu Gly Ala
305 310 315
Ser Ala Ala Ser Leu Leu Gly Leu Leu Gly Leu Trp Leu Pro Arg
320 325 330
Pro Val Pro Leu Val Ala Gly Ala Gly Val His Leu Leu Leu Thr
335 340 345
Phe Ile Leu Phe Phe Trp Ala Pro Val Pro Arg Val Leu Gln His
350 355 360
Ser Trp Ile Leu Tyr Val Ala Ala Ala Leu Trp Gly Val Gly Ser
365 370 375
Ala Leu Asn Lys Thr Gly Leu Ser Thr Leu Leu Gly Ile Leu Tyr

WO 02/34783

PCT/US01/49670

390 385 390
Glu Asp Lys Glu Arg Gln Asp Phe Ile Phe Thr Ile Tyr His Trp
395 400 405
Trp Gln Ala Val Ala Ile Phe Thr Val Tyr Leu Gly Ser Ser Leu
410 415 420
His Met Lys Ala Lys Leu Ala Val Leu Leu Val Thr Leu Val Ala
425 430 435
Ala Ala Val Ser Tyr Leu Arg Met Glu Gln Lys Leu Arg Arg Gly
440 445 450
Val Ala Pro Arg Gln Pro Arg Ile Pro Arg Pro Gln His Lys Val
455 460 465
Arg Gly Tyr Arg Tyr Leu Glu Glu Asp Asn Ser Asp Glu Ser Asp
470 475 480
Ala Glu Gly Glu His Gly Asp Gly Ala Glu Glu Glu Ala Pro Pro
485 490 495
Ala Gly Pro Arg Pro Gly Pro Glu Pro Ala Gly Leu Gly Arg Arg
500 505 510
Pro Cys Pro Tyr Glu Gln Ala Gln Gly Gly Asp Gly Pro Glu Glu
515 520 525
Gln

<210> 14

<211> 348

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 162131CD1

<400> 14

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

WO 02/34783

PCT/US01/49670

Thr Ala Gly Ile Ala Ser Ser Ile Val Glu Asn Thr Tyr Thr Arg
185 190 195
Ser Ala Glu Leu Thr Ala Ser Arg Leu Thr Ala Thr Ser Thr Asp
200 205 210
Gln Leu Glu Ala Leu Arg Asp Ile Leu His Asp Ile Thr Pro Asn
215 220 225
Val Leu Ser Phe Ala Leu Asp Phe Asp Glu Ala Thr Lys Met Ile
230 235 240
Ala Asn Asp Val His Thr Leu Arg Arg Ser Lys Ala Thr Val Gly
245 250 255
Arg Pro Leu Ile Ala Trp Arg Tyr Val Pro Ile Asn Val Val Glu
260 265 270
Thr Leu Arg Thr Arg Gly Ala Pro Thr Arg Ile Val Arg Lys Val
275 280 285
Ala Arg Asn Leu Gly Lys Ala Thr Ser Gly Val Leu Val Val Leu
290 295 300
Asp Val Val Asn Leu Val Gln Asp Ser Leu Asp Leu His Lys Gly
305 310 315
Glu Lys Ser Glu Ser Ala Glu Leu Leu Arg Gln Trp Ala Gln Glu
320 325 330
Leu Glu Glu Asn Leu Asn Glu Leu Thr His Ile His Gln Ser Leu
335 340 345
Lys Ala Gly

<210> 15

<211> 520

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1937725CD1

<400> 15

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

WO 02/34783

PCT/US01/49670

```

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

```

<210> 16

<211> 534

<212> PPT

<213> Homo sapiens

<220>

WO 02/34783

PCT/US01/49670

<221> misc_feature

<223> Incyte ID No: 3643847CD1

<400> 16

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

WO 02/34783

PCT/US01/49670

Tyr Leu Gly Leu Phe Ser Phe Leu Ala Ala Leu Ala Phe Ser Thr
 395 400 405
 Met Ile Cys Ala Gly Pro Arg Ala Trp Lys His Phe Thr Thr Arg
 410 415 420
 Asn Arg Glu Tyr Asp Asp Ile Asp Asp Asp Asp Pro Phe Asn Pro
 425 430 435
 Gln Ala Trp Arg Met Ala Ala His Ser Pro Pro Arg Gly Gln Leu
 440 445 450
 His Ser Phe Cys Ser Tyr Ser Ser Gly Leu Gly Ser Gln Thr Ser
 455 460 465
 Leu Gln Pro Pro Ala Gln Thr Ile Ser Asn Ala Pro Val Ser Glu
 470 475 480
 Tyr Met Asn Gln Ala Met Leu Phe Gly Arg Asn Pro Arg Tyr Glu
 485 490 495
 Asn Val Pro Leu Ile Gly Arg Ala Ser Pro Pro Thr Tyr Ser
 500 505 510
 Pro Ser Met Arg Ala Thr Tyr Leu Ser Val Ala Asp Glu His Leu
 515 520 525
 Arg His Tyr Gly Asn Gln Phe Pro Ala
 530

<210> 17

<211> 820

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6889872CD1

<400> 17

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

WO 02/34783

PCT/US01/49670

```

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

```

WO 02/34783

PCT/US01/49670

```

        605                610                615
Arg Gln Leu Ser Ala Asp Ala Ala Val Thr Arg Lys Thr Cys Ser
        620                625                630
Val Ser Ser Ser Gly Ser Ile Lys Ser Ala Lys Val Phe Ser Leu
        635                640                645
Asp Val Pro Asp His Pro Ala Ala Thr Gly Leu Ala Lys Gly Asp
        650                655                660
Ser Lys Tyr Ile Glu Lys Gly Ser Pro Leu Asn Ser Pro Leu Asp
        665                670                675
Arg Leu Pro Leu Val Pro Ala Gly Ser Gly Gly Gly Ser Gly Gly
        680                685                690
Gly Gly Gly Ile His His Leu Glu Val Lys Pro Ala Tyr His Cys
        695                700                705
Ser Glu His Arg His Ser Phe Pro Ala Leu Tyr Tyr Glu Glu Gly
        710                715                720
Ala Asp Ser Leu Ser Gln Arg Val Ser Phe Leu Lys Pro Leu Thr
        725                730                735
Arg Ser Lys Arg Asp Ser Thr Tyr Ser Gln Leu Ser Pro Arg His
        740                745                750
Tyr Tyr Ser Gly Tyr Ser Ser Ser Pro Glu Tyr Ser Ser Glu Ser
        755                760                765
Thr His Lys Ile Trp Glu Arg Phe Arg Pro Tyr Lys Lys His His
        770                775                780
Arg Glu Glu Val Tyr Met Ala Ala Gly His Ala Leu Arg Lys Lys
        785                790                795
Val Gln Phe Ala Lys Asp Glu Asp Leu His Asp Ile Leu Asp Tyr
        800                805                810
Trp Lys Gly Val Ser Ala Gln Gln Lys Leu
        815                820

```

<210> 18

<211> 2653

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6431478CB1

<400> 18

```

gctgcggctg agcccagcgc tcgaggcgcg aggcagccag gagggcccggt gcggcgcggg 60
gagccagcga gcgcgccttc ggcattggcc gcgcgatgt cagctcagtg ctgtgcgggc 120
cagctggcct gctgctgtgg gctgcaggc tgcctctctct gctgtgattg ctgccccagg 180
attcggcagt cctcagcac ccgcttcagt tacgccctct acttcattct ggtcgtcgtc 240
ctctgctgca tcatgatgtc aacaaccgtg gctcacaga tgaagagca cattcctttt 300
tttgaagata tgtgtaaagg cattaagct ggtgacacct gtgagaagct ggtgggatat 360
tcgtccgtgt atagagtctg ttttggaatg gctgtttct tctttatctt ctgtctactg 420
acctgaaaaa tcaacaacag caaaagtgtg agagctcata ttcaaatgg cttttggttc 480
tttaaacctt tgctgttggg ggccatgtgc tcaggagcctt tcttcattcc agatcaggac 540
acctttctga acgcctggcg ctatgtggga gcgctcgag gcttctcttt cattggcatc 600
cagctcctcc tgctcgtgga gtttgacat aagtggaaac agaaactggac agcaggcaca 660
gccagtaaca agctgtggta cgcctccctg gccctgggtga cgctcatcat gtattccatt 720
gccactggag gcttgggttt gatggcagtg ttttatacac agaaagcagc ctgcatggaa 780
aacaataatc tgctgggagt aaatggaggc ctgtgcctgc ttatatcatt ggtagccatc 840
tcaccctggg tccaaaatcg acagccacac tcggggctct tacaatcagg ggtcataagc 900
tgctatgtca cctacctcac cttctcagct ctgtccagca aacctgcaga agtagttcta 960

```

WO 02/34783

PCT/US01/49670

```

gatgaacatg ggaaaaaatgt tacaatctgt gtgcctgact ttggtcaaga cctgtacaga 1020
gatgaaaact tggtagactat actggggacc agcctcttaa tcggatgtat ctgtattca 1080
tggtttgacat caacaacaag atcaggttct gacgtctctg aggggggata cgcagctcct 1140
gaattggaga tagctcgtcg ttgtttttgc ttcatgtctg gtggagagga cactgaagag 1200
cagcagccgg ggaaggagggg accacgggtc atttatgacg agaagaaagg caccgtctac 1260
atctactctc acttccactt cgtgttcttc ttagcttccc tgtatgtgat gatgacgtc 1320
accaacttgt toaactacga aagtgcacac atcgagagct tcttcacgg gagctggtcc 1380
atctcttggg toaagatggc ctctcgtctg atatcggtgc tgttgtacct gtgtacgtcg 1440
gtcgctctccc tctgctgccc caccggggag ttctctgtgt gatgatatcg gcgtccccc 1500
gggcttttgt ggctacacgo ctggaaagtg coactctttt aacagtgtcc ccggggcagg 1560
gactggcgcc ctgtgcctga gtgggtctga aaagcctttg agagagaaaa aaaaaaatct 1620
cctgattagc tttttacttt tgaattctaa aaagaaacta ccagtttgtc ccaagggaat 1680
tgaattttc aaccaacttg atcatggttg aaatatctta cccctaggaa ctggatacca 1740
gttatgttga ctctctctcg catgtttttg ccaaaacaga atttggggca cagcatcttt 1800
tcacagggat aaaaatatct cgtggggcca gtcatctca tcttcggaat agaaaaacat 1860
gccaaactct tgagtcccca gcgcctaaca gaatccagac cctctcact cacttccgcc 1920
tcttagagcc ttgtcccccag ggggctttga ggacaggact cagcctcag ggcctcgtt 1980
atttataggg tcaagagga ggcactgtct tttcaactgc accctcagtg ctgcctcttc 2040
acggccccc aacgtttccc tttaggttg tgatgtgtgg aatccagac ttcactctct 2100
gcctgcctcc ttcccccagg tctcatcttt tctgggtccc acatctttgt aataatgtg 2160
aaagccacca ttgtctgat cacccccag gtggttccc accttattat cactacctga 2220
tcaggattac tgcaataagt acgggcctta ttatgtgtg tagtcacatg attatagaac 2280
aagattcatg ttctctctgc ctaagcaatg gagggtctat attcttactt gtttgtgtg 2340
ttgataatga taatactttt aggaacctaa ctgaaagct gcttcgtgtt gaagcctgt 2400
gcactccctc ctctctcagt tgttgaggtc agccctcag ttttttctcc acctgaggc 2460
ctttgaact gtanaagcgg aagtctgttt gtgttttggg tctgtaacgt gaccataccg 2520
ttcaggttca tgcctgcato ctltggagtag atttctaat gtgagaattt ctgaggtgag 2580
gactctcagc acactgacca gaagaagctt gttaggcaat gtgtggaagt ggccgaatat 2640
acttaaaaag agg 2653

```

<210> 19

<211> 3531

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3584654CB1

<400> 19

```

ggaagcgggtc ggggctgcac actcggatcg gggggccgg ctccggggcc cggcgggtg 60
gaggaggagag ggaaggaggc gggaggagac gagcggagcc atgggtgcgc acgtacgcc 120
cagcgtctggg atttatcggc tcgcgaggag agcggagcag gcgcgcggcc caggcggagg 180
agcgcgcact ctggagcagc cggagctgga agaggaggag gaggagagcc ggccgggaag 240
gaggaggagag gggagagtcg ctccgcggg gcgagcatgg ggcgcctggc ctgcaggccg 300
ctgctcgtgg cgtctcgtgc gttggctctt tgccgagggc gtgtgtgtg agtcccccac 360
gcgacccctgg ttctagtggt gggcactgag ctggtcatcc cctgcaactg cagtactat 420
gatggcccca gcgagcaaaa ctttgactgg agcttctcat ctttggggag oagctttgtg 480
gagcttgcaa gcaactggga ggtggggttc ccagcccgac tgtaccagga cggcgtgcag 540
aggggcgaga tctctgtaag gcgactgccc aacgacgcgc tggagctcca cataaagaac 600
gtccagcctt cagaccaagg ccaactcaaa tgttcaaccc ccagacagga tgccactgtc 660
cagggaact atgaggacac agtcaggtt aaagtgtcgg ccgactccct gcacgtgggc 720
ccagcgcgc ggcccccgc gagcctgagc ctgcgggagg gggagccctt cgagctgcgc 780
tgccacgcgc cctccgcctc gccctgcac acgcacctgg cgtctgctgt ggagtgcaac 840
cgccgcgcgc ccaggcgag cgtctcgcct ctgacccacg agggcaggtt ccaaccgggc 900
ctgggttagc agcagcgcta ccacagtggg gacgtgcgc tcgacacgtt gggcagcgac 960

```

WO 02/34783

PCT/US01/49670

```

gctacccgac totcagtgct ccgggctctg totgcggacc agggctccta cagggtgtatc 1020
gtcagcgagt ggtatcgcca gcggggcaac tggcaggaaa tccaggaaaa ggcgtggaa 1080
gttgccaccg tggatgacca gccgacagtt ctgcaggcag ctgtgcccaa gaatgtgtct 1140
gtggctgaag gaaaggaaat ggacctgacc tgaacatca caacagaccg agccgatgac 1200
gtccggcccg aggtgacgtg gtccctcagc aggatgcctg acagcacctc acctggctcc 1260
cgctgtgtgg ccgggcttga ccgtgatctc ctgggtgcaca gctcgctca tgttgccttg 1320
agtcattgtg atgcacgctc ctaccattta ctgggtcggg atgttagcaa agaaaactct 1380
ggctactatt actgccacgt gtccctgtgg gcacccggac acaacaggag ctggcacaac 1440
gtggcagagg ccgtgtcttc cccagctggg gtgggtgtga cctggctaga accagactac 1500
cagggtgtacc tgaatgcttc caaggctccc gggtttgtcg atgacccca accagctggc 1560
tgccgggtgg tggacacgaa gagtggggag gogaatgtcc gattcacggt ttcgtgttac 1620
tacaggatga accggcgtag cgaacaatgt gtgaccagcg agctgcttgc agtcagtac 1680
ggggactgga cgttaaaata tggagagagg agcaagcagc gggccaggga tggagacttt 1740
attttttcta aggaacatca agcacgttcc aatttcagg tccaaaggac tacagaggaa 1800
cacagaggca attattactg tgttgtgtct gctggacca aacagaggaa caacagctgg 1860
gtgaaagcaa aggatgtctt ctccaaagct gttacatat tttgggcat agaagattcc 1920
gtgcttgttg tgaaggcgag gcagccaaag cctttctttg ctgcccgaac tacatttgag 1980
atgacttgca aagtatcttc caagaatatt aagctcgacc gctactctgt tctcatcatg 2040
gtcgagaagc ctgtcgcgga cctctccagt cccaatgaaa cgaagtacat catctctctg 2100
gaccaggatt ctgtgttgaa gctggagagt tggacagat catcacgggt ggtggcggtt 2160
gttttagaaa aagtgcagga ggtgagttc cgtatcgaa tgtaccagac tcaggtctca 2220
gacgcagggc tgtaccgctg catggtgaca gctggtctc ctgtcagggg cagcctttgg 2280
cgagaagcag caaccagttc ctccaatctc attgagatag acttccaaac ctacagttcc 2340
atatattaat ctctgtgca ttcagacaca ccatcagtaa ttccgggaga tctgatcaaa 2400
ttgttctgta tcatcactgt cgaggagaca gcactggatc cagatgacat ggcctttgat 2460
gtgtcctggt ttgcggtgca cctctttggc ctggacaggg ctctctgtct cctgtcttcc 2520
ctggatcgga agggcactgt gacccactcc cggaggagct ggaagagcga cctcagctcc 2580
gagcggtgga gtgtgtgtga attctgtgtg caagtgcagt gctccaggga ccaggacttt 2640
ggcaactact actgttccgt gactccatgg gtgaagtac caacaggttc ctggcagaag 2700
gaggcaagga tccactccaa gccggttttt ataactgtga agatggatgt gctgaacycc 2760
ttcaagatcc ccttgctgat cggcgctggc ctgtccacgg tcatcgggct cctgtctgt 2820
ctcatcgggt actgcagctc ccactgggtg tgaagaagg aggttcagga gacacggcgc 2880
gagcgccgca ggtcatgtc gatggagatg gactaggctg gccggggagg ggaagtacag 2940
agggacgttc taggagcaat tggggcaaga agaggacagt gatattttaa aacaaagtgt 3000
gttacacata aaaccagtcc tctctaactc caggtgggac ttggcgctct ctcttttctg 3060
catgtcaagt tctgagcgcg gacatgttta ccagcacacg gctcttcttc ccacggcact 3120
ttctgatgta acaatcgagt gtgtgttttc ccaactgcag ctttttaagt gttacacttc 3180
atctaatttt ttttctccca ctggtttata gatcctctga cttgtgtgtg tttatagctt 3240
ttgtttcgcg ggggtgttgt gaggaagggg tgatggcatg cggagttctt tgtcttcagt 3300
gagaatgtgc ctgcccgcct gagggccagc ttccgcgttg gaggcacgtg ttcagagagc 3360
tgtgtgagcg caccctctac ccggctgaca gacaacacag acctgtgccc aaggctaatt 3420
tgtggctttt acgacccctc cccacccctc gttttcaggg gtttagacta catttgaat 3480
ccaaacttgg agtatataac ttcttattga gcccaactgc tttttttttt t 3531

```

<210> 20

<211> 2280

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3737084CB1

<400> 20

```

ggcgcccat ttccagccca agtttccagg tcgggtttcc aggtccagaa ttttccagga 60
gtaggtttctt gggcagtggt tgtgggagct ggaatggcgc agctggaggg ttactatttc 120

```

WO 02/34783

PCT/US01/49670

```

toggcgccct tgaagctgtac ctttttagta tctgcctcc tcttctccgc cttcagccgg 180
gcgttgccgag agccctacat ggacgagato ttccacctgc ctcaaggcga gogctactgt 240
gaggggccatt tctcccttcc ccaagtggtat cccatgatta ctacattacc tggctgtgac 300
ctgggtgtcaa ttggagtgat caaacctgcc atttggtatc ttggatgggc tgaacatggt 360
gtctgtccca ttggagtgct cagatttggc aatcttctct tcaagtgtgg caactcttat 420
ttactatatt tgcctttctg caaggtacaa cccagaacaa aggtgcctc aagtatccag 480
agagctctgtt caacattaac actagcagta ttccaacac ttatttttt taacttctct 540
tattatacag aagcaggatc tatgtttttt actctttttg cgtatttgat gtgtctttat 600
ggaaatcata aaactccagc cttctctgga ttttgggtct tcaatgtttc gcaaacaaat 660
atcatctggg ctgtctctctg tgcaggaaat gtcattgcac aaaagttaac ggaaggttgg 720
aaaactgagc tacaaaagaa ggaagacaga cttccacctc taaaggaacc atttgacaga 780
ttcagaaaaa ttcttcagtt tcttttggct tattccatgt ctttaaaaaa cttgagtatg 840
cttttgcctc tgacttggcc ctacatcctt ctgggatttc tgttttggc tttttagta 900
gttaatgggt gaattgttat tgggatcgg agtagtcag aagcctgtct tcaatttctc 960
caactattct actttttttt atttactctc ttttttctt ttccctatct cctgtctctc 1020
agcaaatcta agacttttct ttctctagtt tggaaacgta gaattctgtt ttttgggttt 1080
acottagtct ctgtgttttt agtttggaaa ttcccttatg ctcaataata'cttgctagca 1140
gacatagac attatacttt ctatgtgtgg aaaagagttt ttcaagata tgaactgta 1200
aaatatttgt tagttccagc ctataatatt gctgggtgga gtatagctga ctcatgaaa 1260
tcaaaagcaa ttttttggaa tttaattggt ttcatatgct tgttcaactg tatagtctct 1320
cagaaactgc tggaaatttcg ttacttctat ttacttatg tcaattatag gcttaacata 1380
cctctgcctc ccacatccag actcatttgt gaactgagct gctatgcagt tgttaatttc 1440
ataacttttt tcatctttct gaacaagact tttcagtgcc caaatagta ggaacttcaa 1500
aggtttatgt ggtaatatca gtgatatttc gaactgtgaa aatggactta ataattagac 1560
cattctcaca aagaacaact gaataggtgg aaacatgga atttctttta ggtgcagtg 1620
tgggtctcaa attacattag ttttttttat atatttttta aacatattga agaaattaa 1680
tggcaaaagaa ctgagaagac ttaagacctg cttcaaaagc ctgaanaaat gaaaaataaa 1740
attgttttca gatatctcat atcaactctc taatgttggc ccttaaaaaa ccttggggaat 1800
gttttggtag tacaagttaa ttaaaactgg gtatgcttca aaaaaaaaaa aaaagggggg 1860
gggttccccc ccccaattcc gaaactgga aaagcgttcc cccggggaaa attttttacc 1920
cccccaatt ccccaaaaaa ttggggcccg ggaagcctaaa ggtactaccc cggggggccc 1980
taaggggtgg gccccccccc attaatggg gtggcccaaa tgcocccgtt ccaattggga 2040
aacotttttg tccccccccc ttattattt ggccaacccc cggggaaaag ggggtttctc 2100
tttgggggccc ttccctgttc cccggccaat aaacgggttc ccccggtttt tgggttccg 2160
ggaaggggtt ccagccccc aaaggggggt aaacgggttt ccccaaaatt cggggggaaa 2220
ccccggaaaa aacatttttg cccaaagggc ccccaaaagg ccagggccct taaaaaggcc 2280

```

<210> 21

<211> 1104

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 71426238CB1

<400> 21

```

taagagaggt tttgcctctt tttagacctt agtcatgagt tggatgttcc tcaagatct 60
cctgagtgga gtaataaat actccactgg gattggatgg atttggctgg ctgtctgttt 120
tgtcttccgt ttgctggtct acatgggtgc agcagagcac gtgtggaaag atgacagaaa 180
agagtttgag tgcaacagta gacagccggc ttgcacaaat gtgtgttttg atgacttctt 240
ccccatttcc caagtacagc ttggggcctt acaactgata atggtctcca caccttcaact 300
totggtgggt ttacatgtag cctatcatga gggtagagag aaaaggcaca gaaagaaact 360
ctatgtcagc ccaggtacaa ttgatggggg cctatggtac gcttatctta tcaagctcat 420
tgttaaaact ggttttgaaa ttgcttctct tgttttattt tataagctat atgatggctt 480
tagtgttccc tacctataaa agtgtgattt gaagccttgt cccaaactgt tggactgctt 540

```

WO 02/34783

PCT/US01/49670

```

catctccaaa cccactgaga agacgatctt catctcttct ttggtcatca cctcatgctt 600
gtgtattgtg ttgaatttca ttgaactgag ttttttgggt ctcaagtgcc ttattaatg 650
ctgtctccaa aaattattta aaaaacctca agtctcaggt gtgtgagtgc cacagcctca 720
gatagtgtga atgtggttag agagggaacc ctccctactt ccgaatctt cacacttggc 780
cataaaacaa ctccctctac ctgaagcaaa gctactctgt gacacacaag aggyttaaac 840
aaagaaaacc tgcattccct ctcaagcaag cctaagctga gttggaagac aaagcaatc 900
agcttttagt taatttggga ggaatttttt tacattgtca atatgcttcc agttatgagc 960
tctagacaga ggtctcattg ttttgttcta gggttctcca gtatgtggat aacattagtt 1020
gttttagaat aggtaatgac aaattagctt gaagaaatct aacaggatcc ttttaagagc 1080
ttagattttt cagggaacaa aaaa

```

<210> 22

<211> 4966

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7475123CB1

<400> 22

```

ggcgcccgag ggcatgtgg gggcgccgag gccggtgca cccgggacgc ttccctcgg 60
ggacctcccg cgggtttctc cgcgcgcgcg tcggcgggga gccggcgggg cccggggcga 120
cggcgccggg cggcaccatg aggcggcagt gggcgcgctg gctgtttggc gccctgctct 180
ggcgcacagg cctggccagc agccccaggt gtgtttgtgg tcggagccac ttccatgtgt 240
cagtgagtgc tcttgagagc tgtacctgca tccctgccca gtggcagtgt gatggagaca 300
atgactcggg ggaccacagc gatgaggatg gatgtatact acctacctgt tccctcttgg 360
actttcactg tgacaatggc agtgcatcc gccctcctgt ggtgtgtgac ggggacaagc 420
actgtgagga tgactgggat gaggcaggact gtcccccccg ggagtgtgag gaggacgagt 480
ttccctgccca gaatggctac tgcacccgga gtctgtggca ctgcgatggt gacaatgact 540
gtggcgacaa cagcgatgag cagtgtgaca tggcgaagt ctccgacaag gagtccgct 600
gtagtgcagg aagctgcatt gctgagcatt ggtactgca cggtgacacc gactgcaaa 660
atggtccgga tgaggagaac tgtccctcag cagtgcacgc gcccccctgc aacctggagg 720
agtccagtg tgccatgga cgtgcaccc tcgacatcta ccactgcgat ggcgacgatg 780
actgtgagga ctggtcagac gagtctgact gctgtgagta ctctggccag ctggggagct 840
ccccaccgcc ctgcgcctct ggggagttca tgtgtgacag tggcctgtgc atcaatgcag 900
gctggcgctg cgatggtgac ggggactgtg atgaccagtc tgatgagcgc aactgcaact 960
ggcgagcaaa gtcaatccag cgtgttgaca aatactcagg ccggaacaag gaggcagtc 1020
tggcaaatgt ggaaggatc atggatatca tctgtgttcc cctcagcggc cagacaggga 1080
ccaatgcctg tgggtgtaac aatgtgtgct gcacccacct ctgcttggcc agagcctcgg 1140
acttctgatg tgccctgctt gacgaacctg atagccggcc ctgctccctt gtgctgtgcc 1200
tggtaaccac agctcctagg gctactggca tgagtgaaaa gagccagtg ctacccaaca 1260
caccacctac caccctgtat tcttcaacca cccggacccg cagctctctg gaggagggtg 1320
aagggaaggt ggacatccgt cgaatcagct ttgaacaga ggacctgtct gatgatgtca 1380
tcccactggc tgacgtgcgc agtgcgtggt ccttgcactg ggaactccgg gatgaccacg 1440
tgtactggac agatgtcagc actgatacca tcagcagggc caagtgggat ggaacaggac 1500
aggaggtggt agtgatacc agtttggaga gccagctgg cctggccatt gattgggtca 1560
ccaacaaact gtaactggca gatgcaggta cagaccggt tgagtagacc aacacagatg 1620
gcagcatgag aacagtactc atctgggaga acctgtatgc tccctgggac atcgtgtgtg 1680
aacccatggg cgggtacatg tattggactg actggggtgc gagccccag attgaacgag 1740
ctggcatgga tgccctcaggc cgcgaagta ttatctcttc taatctgaac tggcctaagt 1800
ggttagctat tgattatggg tcccagctc tatactgggc tgacgcgggc atgaagacaa 1860
ttgaatttgc tggactggat ggcatgaaga ggaaggtgct gattggaagc cagctcccc 1920
acctatttgg gctgacctc tatggagagc gcactatatt gactgactgg cagaccaaga 1980
gcatacagag cgtgaccgg ctgacagggc tggaccggga gactctgcag gagaacctgg 2040
aaaacctaat ggacatccat gtcttcacc gccgcgggcc cccagtgtct acaccatgtg 2100

```

WO 02/34783

PCT/US01/49670

```

ctatggagaa tggcggctgt agccacctgt gtcttaggtc cccaaatcca agcggattca 2160
gtgttaacctg cccacacaggc atcaacctgc tgtctgatgg caagacctgc tcaaccagga 2220
tgaaacagttt cctcatctctc gccaggagga tagacatctg catggctctc ctggacatcc 2280
ctatttttgc tgatgtggtg gtccaatca acattaccat gaagaaacac attgccattg 2340
gagtagaaccc ccaggaaagga aagggtgtact ggtctgacag cacactgcac aggatcagtc 2400
gtgccaatct ggtatggtcca cagcatgagc acatcatcac cacagggcta cagaccacag 2460
atgggctcgc ggttgatgcc attggccgga aagtatactg gacagacacg ggaaccaacc 2520
ggattgaagt gggcaacctg gacgggtcca tgcggaaagt gttggtgtgg cagaaccttg 2580
acagtcctccg ggcacatcta ctgtacctag agatgggggtt tatgtactgg acagactggg 2640
gggagaatgc caagttagag cgttccggaa tggatggctc agaccgcgcg gtgctcatca 2700
acaacaacct aggatggccc aatggactga ctgtggacaa ggccagctcc caactgctat 2760
gggcagatgc ccacaccagc cgaattgagg ctgctgacct gaatgggtgcc aatcggcata 2820
cattggtgtc accggtgcag caccocatag gctcaccct gctcagctcc tatatctact 2880
ggactgactg gcagactcgg agcatccacc gtgctgacaa ggttactggc agcaatgtca 2940
tctctgtcag gtccaacctg ccaggccctca tggacatgca ggtgtgtggc cgggacacag 3000
cactagggtt taacaagtgc ggtcgagaa atggcggtgt ctccacctc tgcctgcctc 3060
ggcctctcgg ctctcctgtg gcttgcctca ctggcatcca gctgaaggga gatgggaaga 3120
cctgtgatcc ctctcctgag acctaccctg tctctccag cgttggctcc atccggccta 3180
tctcactgga caccagtgc caccaccgat ccatcctgag tgttctgag ctcaacaatg 3240
tcatctcctt ggcctatgac agcgtggatg gaaaggtcta ttacacagat gtgttctctg 3300
atgttatcag gcgagcagac ctgaacggca gcaacatgga gacagtgtac gggcgaggcg 3360
tgaaagccac tgacgggctg gcagtggact ggttggccag gaacctgtac tggacagaca 3420
caggtcgaaa taccattgag gcttccaggc tggatgggtc ctgccgcaaa gtactgtaca 3480
acaatagcct ggtatgagcc cggcccatgt ctgttttccc caggaaagggt tactctctct 3540
ggacagactg gggccacatt cccaagatcg aacgggcaaa ctgtgatggt tctgagcgga 3600
aggtcctcat caacacagac ctgggttggc ccaatggcct tactcctgag tatgatcccc 3660
gcaggtacta ctgggtggat ggcctctgag accggatcga gagtgtgac ctcaatggga 3720
aactcgggca ggtcttggtc agcctatgtg cccacccttt tgccctcaca cagaagaca 3780
gttggtacta ctggacagac tggcagacca agtcaatoca gctgttggc aaatactcag 3840
ggcggaacaa ggaagcagtg ctggcaaatg tggaaagact catggatctc atcgtgtgtt 3900
ccctcagcg gcagacaggg accaatgctt gtggtgtgaa caatggtggc tgcaccaccc 3960
tctgtcttgc cagagcctcg gacttcgtat gtgctgtctc tgacgaacct gatagccagc 4020
cctgtctcct tgtgcttggc ctggtaccac cagctcctag ggtcactggc atgagtgaaa 4080
agagcccatg gctacccaac acaccaccta ccacctgtta ttcttcaacc acccggaacc 4140
gcacgtctct ggaaggagtg gaaggaagat gctctgaaag ggaatgccag ctgggctctc 4200
gtgacgttcc caatgacgtt gttcctgtcg ctccaggggg aggaacttcat atcagctacg 4260
ccattggtgg actcctcagt attctgtcta ttttgggtgt gattgcagct ttgatgtctg 4320
acagacacaa aaaatccaag ttccactgac ctggaatggg gaacctcacc tacagcaacc 4380
cctcctaccc aacatccaca caggaaatga agattgaagc aatccccaac ccagccatgt 4440
acaaccagct gtgctataag aaagaggagc ggcctgacca taactacacc aaggagaaga 4500
tcaagatcgt agaggggaatc tgcctcctgt ctggggatga tgcctgagtg gatgacctca 4560
agcaactcgc aagctcacgg gggggcctcc tccgggatca tgtatgcatg aagacagaca 4620
cgtgtctcat ccagccacagc tctgtctccc tggatgacac agagacggag cagctgttac 4680
aggaagagca gtctgagtgt agcagcgtcc atactgcagc cactccagaa agacgaggct 4740
ctctgcagca cctgggtctg aaacatgaac gcaagctctc ctacagagag caggtctaaa 4800
tgcccaactt ctcttccctg cctgctctgt ccttctcctt tatggagctc tagtctctgt 4860
gtctgcttcc accgcaggcc cgtctctgtg gtgcttctcc tctctcctcc ccaaccocat 4920
aactgttctc aagccttccc cggagctgtt taccacgtga gtcata 4966

```

```

<210> 23
<211> 5401
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

```

WO 02/34783

PCT/US01/49670

<223> Incyte ID No: 7481952CB1

```

<400> 23
atggaccaga gcatcagcat tacctgggaa cttagtggaa atgcagaacc tcaggccctg 60
gcccagccct acagaaccaa aagctacatg gaacaagcaa agcatctcac ctgtgecttt 120
gaatcggtgt tctgcggttg ggagccattt ctcacagaag atccacactg gaagtgtatg 180
aaaggattga ataattggaga gcaccacttt cctgcagctg atcacacagc aaacataaat 240
catggatcgt ttttttattt ggaggcacag cgtcccccgc ggggtggcaa gcttggaaat 300
cctgttctta caaaattgct cactgcctct accccatgtc aggtgcagtt ttggtatcat 360
ttgtctcaac attcaaatct cctcgttttt acaagaacgt ctctagatgg aaacttgcaa 420
aagcagggca aataaatcag attctccgaa tctcagtggg gccacgcaaa aattgatctc 480
attgcagaag cgggagaatc tactctacct tttcagttaa ttttgggaag tactgttttg 540
tcgtcaaatg ctaccgttgc tctagatgac atcagtgtgt cccaggaatg tgaattttcc 600
tataaatcac taccaaggac cagtacacaa agcaagtttt ccaagtgtga ctttgaagca 660
aacagctgtg attggtttga agtaattagt ggtgaccatt ttgactggat acggagctct 720
cagagtgaac tttctgtgta ttttgagcac caggctccac ctggggatca tagttccaac 780
gcctcgaag ggcattttat gtctattctg aagaaaagca gcagcttgtg gcaagttgct 840
aagcttcaga gcccaacttt caggcagaca ggacctggat gcatacttto cttctggttc 900
tataactatg gcctgtcagt gggagcagct gagctgcagc tacatatgga aaattctcat 960
gactcaacag tgatttggag agtattatac aatcagggca aacaatggtt ggaggcaacc 1020
attcagctgt ggccgttttc gcagcccttc catttgcac tagataaagt cagtctgggc 1080
atttatgatg ggtctcagc tattgatgac atccgatttg aaaaattgta tctcctcttt 1140
cctgctcaga gctgtgaagg cctggatcat ttctgtgttc gccacaccag ggttgcata 1200
gaaaagcttc ggttatgtga tctgtggatg gactgtgtgt atcgtactga tgaagtcaac 1260
tgtgcacctg agctgcagtg taactttgaa actggaatct gtaactggga acaagatgca 1320
aaagatgact ttgattggac caggacaacg ggtccaactc caaacttaaa ccaaggccca 1380
atgaaagata acaactctgg cacagctaaa ggacactatc tctacataga attctcagag 1440
cccaggcttt ttcaagcagc tgctgcctta ctcagcccaa tcttaatgca cactgatata 1500
aaaggctgca cctccgcttt ctattaccac atgtttgaaa agcgcattta taggttgaca 1560
atctaccaac gaactctggg tgaactaagg ggaacagctg tgtggcagat atttgggaat 1620
caaggcaaca gatggattag gaaacacctc aacatttcca gcaggcagcc ctttcagata 1680
ttgggtggag ctccagtggt agatggcttc actggagata ttgcgattga tgatctgtca 1740
tttatggact gcacccctta ccttggtaat ttgccagcag acctcccaac tccaccagaa 1800
acgtcagttc ctgttaacatt acctccacac aactgcacag acagtgaatt tatctgcaag 1860
tctgatgttc actgcatgta aaaaatgcag aaatgtgatt ttaaatatga ctgcctgac 1920
aaatcagatg aagcatcctg tgttatggaa gtttgcagct ttgagaaaag aagcctgtgt 1980
aaatgttatc aaaccaatccc agtcatcttg cttoagattt caaacacatt caggtggggg 2040
cttgggaacg ggaatcagcat tcatcatctg gaagaaaacc acaggccatc agtggatcat 2100
acacaaataa ccatctatgg ctggtacctg tatgctgaca gttctaatg gaatttgg 2160
gacacgctg acattctcac tctatcattt tcaactcagg gacaaaaatg taacttgg 2220
ttctggacac atatgaatgg ggccacgctt ggttctctcc aggtgtctat caagaaagat 2280
aacgttactt ctaaaattgt ggctcaaaact ggacagcaag gtgcacagt gaagagagca 2340
gaagtgtttt taggcatttc ttacatatac cagattgtct taagagccaa acgtggtatc 2400
agttacatag gagatgttag agtggatgat atttctctcc aagattgtct cccttggct 2460
agcccagaga gaaagtgtac tgatcatgaa ttcatgtgtg ctaataagca ctgcattgct 2520
aaagacaagc tgtgtgattt tgtgaatgat tgtgtgata attcagatga gactacttcc 2580
atttgccgta cctccagtgg cgcgtgtgat ttogaatttg atcttggctc ctggaagcag 2640
ggaaagatg aggaacttga ctggaaactg aaagctagca gcctccctgc agcaggccca 2700
gagccagcag cagatcacac tttgggaaat tcatctgttc attacatctt tataaagagt 2760
ttgttctctc agcagcccat gagagctgcc agaatttcaa gtccagttat aagtaagaga 2820
agcacaactc gcaagattat ttttctattt cacatgtatg gaaatggcat tggggcaact 2880
aocctaatgc aggtgtcagt cacaacaaca aagaaggttc tacttaacct cactgtagaa 2940
caaggcaact tctggcggag aagaagactg tcaatgtttg gtgtgaaga cttccaactc 3000
aaatttgagc gttaggttgg gaaaggtcag cgtggagaca ttgcaattga tgacattgtg 3060
cttacagaaa attgtctatc actccatgat tccgtgcaag aagaactgga agtgccctct 3120
ccaacagggt tctgcccact tggctatagg gaatgtcata atggaaaatg ctataggctg 3180

```

WO 02/34783

PCT/US01/49670

```

gaacaaagct gtaacttcgt agataactgt ggagataata ctgatgaana tgaagtgtgt 3240
agctctgtga cttttgaaaa aggtctgtgt ggctggcaaa actccagggc tgacaacttt 3300
gattgggttt taggggttgg ctctcatcaa agcttaagac ctcccaaaag ccacacactt 3360
ggaaatgaaa atggggcactt catgtatctg gaagctactg cagtggggctc tcgggggtgac 3420
aaagcacact tcaggaggtac catgtggcga gaatccagtg cagcctgcac catgagcttc 3480
tggtatttca tatctgcaaa ggccacagga tccattcaga ttctcatcaa gacagagaaa 3540
ggactatcaa aagtatggca agaaagtaag cagaaccctg gtaatcattg gcaaaaggct 3600
gacatctctg taggaagatt aaggattttt gaagtcattt tccaagggtt cagaacaagg 3660
gacctgggag gaggagctgc aattgatgat attgaattta aaaactgcac aactgtggga 3720
gagatctctg agctttgtcc ggaatcactt gattttttgt gccgggacaa gaagtgcatt 3780
gcctccaccc ttctttgtga ctataagcca gactgctctg ataggtctga tgaagctcac 3840
tgtgcacatt atacaagcac aacaggaagc tgcaattttg aaacaagtcc aggaactcgg 3900
accacagcct gcagctttac tcaagactct gaggatgact tggactgggc cattggcagc 3960
agatctctg ccaagcattt aattccagac tctgatcaca cgcagggtag tggtcagcac 4020
ttctgtcag tcaactcacc tggctccaag gaaggatccg ttgcagaaat tactacttcc 4080
aaatccttcc cagcaagcct tgggaatgtg actgttcggt tctggttcta catgtatgat 4140
ccacggagta tgggaatatt aaaggtgtat accattgaag aatcggggct aaacatctcg 4200
gtgtgtccag tgatttgaana taaaagaacg gtagtgacat atggctctgt gcctctctcc 4260
agtaacagtc cgtttaaggc ggcatctgaa gctgatttgg atggaaatga ggacatcttt 4320
atgtctcttg atgacatctc ttttacccca gactgtgtga ctggagggtcc tgtccagctg 4380
cagccctcac cctgtgaagc tgatcagttt tctgtatctt acacactcca atgtgtccct 4440
ctctcagggg aatgtgatgg acatgaagac tgcataagat gatctgatga aatgtatgat 4500
ccctccagcc ccacccctcc actctgtagt aacatggagt tccgtgctc tacagacgag 4560
tgtatacctt ccctctgtct atgctaggga gtgcacgact gccactttaa tgaagatgag 4620
ctcatctgct ccacacaaag ctgttctaag ggagctctgg tgtgtgctc ctccacagc 4680
tgtatccagc ccacacagcg ctgtgatgtt ttgcccagc gcattggatt ccagcttgat 4740
ggctccagct gctcagaatg tccattaaat tactgcagaa atggtgggac ttgtgtagt 4800
gagaaaaatg gtccatgtg tcgatgtaga caaggctgga aggaataatc atgcacatc 4860
aagttaatc ctctctgtac agacttcaca tacgctcaga ataatacatg gactctctg 4920
ggatcttgat tagcatctct gatgactcac atcacagttg cagtctctgt ttttcttga 4980
aacagaaagg taccaataag gaaaaccgag ggaagtgtga actgtgctt tgtcaatcca 5040
gtttacggga actggagcaa cccagagaaa acagagagtt ctgtctatcc cttctcaaac 5100
ccattatagt gcaacaacac aggaagcctg gagaccctgt cacatcatct caaatagcag 5160
catcgagacc aagtctgac caacatgtgt agtttctaga aaattgaagt ctccacaac 5220
tgtatgaaac tcactctcta caatgttaaa aaggaagagg attgtaaat ccagtgtat 5280
tataacattt atgaatgaat tttcttgag aatatagaga atgtttatat ggaatcagaa 5340
tcagtacatt atcttctactg aacatctgaa tattttaata aaatttctat ttaatcaaaa 5400
a

```

<210> 24

<211> 1949

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 382654CB1

<400> 24

```

aggcagggg ctagggtgaa aaagcattga aggccatgag atggctgtga gagagaacca 60
agggcgagaa gtgcacagag ctactctggg ggaggagata gcccccagcc ttaagaagcc 120
aggattggca gggagtgaag agccagagag gcgaagcttt ggagatcag agggcttaaa 180
gttggagtg gtctaaggag cccaggccct gatgcttccc tcttctcat gggcctctgt 240
tcacagaccc cctggagggg gtgaacatca ccagccocgt gcgcctgac catggcaccc 300
tggggagtc ggtctctgtt tctgtgagat acagcagtac cagcagcagc aggcctgtag 360
tgaagtggca gctgaagcgg gacaagccag tgaccgtggt gcagtccatt ggcacagagg 420

```

WO 02/34783

PCT/US01/49670

```

tcatcgccac cctcgccgct gactatcgag accgtatccg actctttgaa aatggctccc 480
tgctctccag cgactcgag ctggccgatg agggcaaccta tgaggtcgag atctccatca 540
ccgacgacac ctctactggg gagaagacca tcaaccttac ttagatgtg cccatttoga 600
ggccacagat gttgggtggt tcaaccactg tgcgtgagct cagcgaggcc ttacacttga 660
actgctcaca tgagaatggc accaagccca gctacactg gctgaaggat ggcaagcccc 720
tctccaatga ctcgagaatg ctctgtccc ccgacaaaa ggtgctccac atcaccocgg 780
tgctcatgga ggtgacgac ctgtacagct gcatgggtga gaaccccatc apccagggcc 840
gcagcctgcc tgtcaagatc accgtatata gaagaagctc cctttacatc atctgttota 900
caggagagat ctctctcctt gtgacctgg tgacagtctg tgcctgctgg aaacctotca 960
aaaggaaaca gaagaagcta gaaaagcaaa actccttgga atacatggat cagaatgatg 1020
accgctgaa accagaagca gacacctcc ctgcaagtgg tgagcaggaa cggaaagaaac 1080
ccatggcact ctatatcctg aaggacaagg actccccgga gaccgaggag aaccgggcc 1140
cggagcctcg aagcgcgagc gaggccggcc cgcocggcta ctccgtgtct ccccgctgc 1200
ccgcccctc gccggggctg cccatccgct ctgcccgcg ctaccgcgcg tcccacgcg 1260
gctcccagc caccggccgg acacactcgt cgcgccagc ggccccgagc tcgcccggcc 1320
gctcgccagc cgcctcgccg acactcgga ctgcccgtg gcacataatc cgcgagcagc 1380
acgagcccg cccggtggag atcagcgctc gaggcgcctc gggatccctc gagagggccc 1440
cgcggtctgc ggcagtggtc cgggggaaa gctggggctg ggaagcccg ggcggcgcg 1500
ctggggacga ggggaggtcc cgggggggag ctggtgtctc ggtgtgaa gtgtatgagc 1560
atgcgcagac ggggaggggt gcgcggaggc ggcagtgtg atatgtgaa accgggtgcg 1620
attgctccc ggtttactgg ctgtgtctc acttggtata ggtgtgcaa tggggttctt 1680
ccgtctctgc taccacttcc gagggaggtc gtctgtctc ggtttcaggc ggtcatcatt 1740
ctgactccatg tattccaggc agtttcgctt ttctagctc ttctgtttcc ttttgaggcg 1800
tttccagcg gcacagactg tcccaaggt cacaaggag aagatgcctc ctgtagacaa 1860
gatgatgac agggagcttc tcttagggag agagagaagc gagagcaggc gggcctccc 1920
ggccagatg tgtgacctc gccctacta 1949

```

<210> 25

<211> 2133

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Inocyte ID No: 1867351CBI

<220>

<221> unsure

<222> 2110

<223> a, t, c, g, or other

<400> 25

```

cactgcggc ctccgggta cccactgcgc gctccgcgc taccggcgc cagcgcgga 60
gtccactgga agctcctgag gaggccgcgc cagtgcgcgc aggccggag gccacccttg 120
aggctccgtg gtcgcgctgc ttgcgcgtgt ccgccttcgc agaagagctc cggcgctct 180
tggctcctgc tggcccgcg ttcttggttc agctgatggt gttcctgac agcttcataa 240
gtccctgtgt ctgtggccac ctgggcaagc tggagctgga tgcagtccac ctggcaatcg 300
cggttatcaa tgtcactggg gtctcagtgg gattcggctt atctctgcg tgtgacaccc 360
toatctccca gacgtacggg agccagaacc tgaagcagct gggcgctgac ctgcagcgga 420
gtgcgctcgt cctgctcctc tgcgtcttcc cctgctgggc gctctttctc aacacccagc 480
acatcctgct gctcttcagg caggaccacc atgtgtccag gcttaccacc acctatgtca 540
cgatcttcat tccagctctt cctgcaacct ttctttatat gttacaagtt aaatatattg 600
toaacccagg aattgtactg cccagatcgt taactggagt tgcagccaac ctgttcaatg 660
ccctgcacaa ctatctgttt ctccatccac tgcactcttg ggtgataggc ctgtcactgg 720
caaatctgat ttcccagtac accctggctc tactctctt ttctacatc ctggggaaaa 780
aactgcatac agctacatgg ggaggtggt cctcagtg cctgaggac tggcctcct 840

```

WO 02/34783

PCT/US01/49670

```
tctctcgccct ggccatcccc agcatgetca tgcgtgcat ggagtggtgg gcctatgagg 900
tcggggagcct ccccatgggc atcctcggca tggtagagct ggccgctcag tccatcgtgt 960
atgaactggc catcattgtg tacatggtcc ctgcagactt cagtggtggt gccagtgtcc 1020
ggtagaggaaa cgtctctgggt gctggagaca tggagcaggc acggaagtc tctacgtgtt 1080
ccctcgtgat tccagtgtcc ttgtctgtag ccttcaggtg cctgtgttta agctgttaag 1140
atccgttggg gtacattttt actacggacc gagacatcat taactgtgtg gctcaggttg 1200
ttccaattta tgcgtgtttc caectctttg aagctcttgc ttgcacaggt ggtggtgttc 1260
tggggggggg tggaaatcag aagggtggag ccattgtgaa taccattggg tactatgttg 1320
ttggcctccc catoggggac ggcgtgatgt ttgcaaccac acttgagtg atgggtctgt 1380
ggctcaggat catcatctgt acagctcttc aagctgtgtg ttttctaggc tttattattc 1440
agctaaattg gaaaaaagcc tgcagcagg ctcaggatca cgcgaattg aaagtaaca 1500
acgtgcctgg gagtgggaat tctgtctctc ctcaggatcc gcttcaccca ggtgcccctg 1560
aaaaacctga aggaatttta acgaacgatg ttggaaagac aggcggagct cagtcagatc 1620
agcagatgct ccaagaagaa cctttgcogg aacatccaca ggacggcgtc aaattgtcca 1680
ggaaacagct ggtgctgogg cgagggtctc tgcctctggg ggtctcttta atctgtctg 1740
tggggatttt agtgagattc tatgtcagaa ttcatgagc tggtaggaaa gaaagtcaag 1800
tcaagtgatg cttttgagct tacacacaat tcacaggccc accagtgaac atttactgtg 1860
agttaattgtc attcaggtgt gcccatggat tttgagggct ggaatgcaa agacacattt 1920
ttctataaaa agaaaaagca actaaggtta aaagctatat tgtggcccaa gacactgttt 1980
gtgaaagatg ccatgattag taattcacca ctatcttgaa ccaagccagc gatcaatgtg 2040
ctgactgcat cggccaatgg ctttgatact totgtctatt ttttagacac aaccoataac 2100
tacggggatn actagtctta agcgcggcca cgg 2133

<210> 26
<211> 2090
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3323104CB1

<400> 25
ggcgaagcgg agaccgtcca tccagaggaa ggcaagtttt tggctcgggc ggcctgagaag 60
accgcgcggg gctggagaca ggttagcagta cggggggcgg gcttcacgcc ggcctgtgata 120
gtctgcagtc gtttcgggtg gcagcctggc gggtagggaga tgcggcggcc acctgctgca 180
aagaaccgaa gggaaaggtta gaagtacgaa ggcagtttgg agctggggct aagcagctgt 240
cgcaacgtca gatcatgggc tccaccaagc actggggcga attgctcctg aacttgaagg 300
tggctccagc cggcgtcttt ggtgtggcct ttctagccag agtcgccctg gtttctatg 360
gcgtcttcca ggaccggacc ctgcacgtga ggtatacggg catcgactac caggtcttca 420
ccgacgcgcg gcgcttcgtc accggagggc gctgccttta cctgagagcc acgtacogtt 480
acaccgcgct gctgggttgg ctcctcactc ccaacatcta cctcagcgag ctctttggaa 540
agtctctctt catcagctgc gacotcctca ccgctttcct cttataccgc ctgctgtctc 600
tgaaggggct ggggcgcgcg caggcttctg gctactgtgt cttttggctt cttaaccccc 660
tgctctatgg agtatccagc cgcggtaatg cggactctat tgcgcctcc ctggtctcga 720
tggctctcta cttgataaag aaaagactcg tcgcgtgtgc agctgtatto tatggtttcg 780
cggtgcatat gaagataatc ccaagtactt acatccctcc cataaccctc caactgtctc 840
cagatcgoga caatgacaaa agcctccgto aattccggtc cactttccag gcttgtttgt 900
acgagctcct gaaaaagctg tgtaatcggg ctgtctgtct gttttagtca gttgtggac 960
tcaacttttt tgccctgagc tttgggtttt actatgagta cggctgggaa ttttggaaac 1020
acactacttt ttatccactg actagcgggg atatccgta caacttttct ccgtacttat 1080
actgtctgta tttgactgca gagagcaagt ggagtttttc cctgggaatt cctgcaattc 1140
tgccacagct catottgott toagctgtgt ctttgccta ttacagagac ctggtttttt 1200
gttgttttct tcatacgtcc atttttgtga cttttaacaa agtctgcacc tcccagtaac 1260
ttcttttgta cctctgttta ctgcctcttg tgatgcact agtcagaatg ccttgaaaaa 1320
gagctgtagt tctctaatg ttatggttaa tagggcaggc catgtggctg gctcctgcct 1380
```

WO 02/34783

PCT/US01/49670

```

atgttctaga gtttcaagga aagaacacct ttctgtttat ttgggttagct gggtttgtct 1440
ttttctctat caattgttcc atctcgattc aaatttattc ccattacaaa gaagaacccc 1500
tgacagagag aatcaaatat gactagtgtt tttccacac cctctgctac tgtgttaaat 1560
tctgatgttc tigtatggac cagaagagag ctttgggaca tttttctga acattctaag 1620
cattctagtg aaagtcccca tgttccaaa gaacttaaaa gcaatgtttg ccttatatat 1680
aaaagggaca caataattga ggtccacctt ctaggaaatc ctaggactcg tttatttggg 1740
acatgggtgg aataaagytc acatattgga aaatggaaag gctgatgaaa ctatcagata 1800
ctaaaaacatt cttaaaatag aggaatatag tttagagacat caggtttaag ccagtatttg 1860
ttctgttttt acaatgtctc tgtcttaagc tgtgtcttaa cttttaaac ccattctttc 1920
ttttctaaagc tttctgaca gctgtgaaaa tccaaaaaat attcttaaac tgtgtatggt 1980
ggcccttgcc tgtagtctca gcactttggg aggctgaggt gggagggtcg cttgagtcca 2040
ggagtgtcag acccaacctg ggcaagatgg tgagacctag tctcaaaaaa 2090

```

<210> 27

<211> 1618

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4759306CB1

<400> 27

```

agaatctatg ggattttcag ctcgatacaa tttcacacct gatcctgaat ttaaggacct 60
tggagctttg aaaccattac cagcgtgtga gtttgagatg ggcggttccg aaggaaattgt 120
ggagtctata caaattatga aggaagcaa agctactgct agcgaggctg ttgattgcaa 180
gtgtgtacatc cagacacctc agcgttccaa gatttactta cgattcttgg actatgagat 240
gcgaattcca aatgagtcca agaggaattt tbtggctgtg tatgatggaa gcaagtccgt 300
ggaggatttg aaagctaagt tctgtagcac tbtggctaag gatgtcactg tacgcacggg 360
tcttgggggt atccgcactg ggcagatga ggcagtcga aacagccgat ttcagatgct 420
cttcacatcc tttcaagaac ctcttgtga aggaacaca ttctctgac atagttaact 480
gtgtattaat aatactttgg tctgcaatgg actccagAAC tbtgtgtatc cttgggatga 540
aaatcactgt aaagagaaga ygaaaacag cctgctggac cagctgaaca acaccagtgg 600
gactgtcatt ggcgtgactt cctgcatcgt gatcatctc attatcatct ctgtcatcgt 660
acagatcaaa cagcctcgtt aaaaagtatg ccaagggaat tcagactttg accagacagt 720
tttccaggag gtattttgac ctctcatta tgagttatgc actctcagag ggacaggagc 780
tacagctgac tttgcagatg tggcagatga ctttgaaatc taccataaac tgcggagggtc 840
atcttccaaa tgcattcatg accatcaact tggatcacag ctgtccagca ctaaggccag 900
cgcgagtaac ctacgacaaa gagatgcttc tatcttgaca gagatgcccc cacagccagg 960
aaaaacccct atcccccca tgaacagang aaatatcctt gtcattgaaa acnaactatc 1020
gcaagatgct gcaagtgctt gtgacataga tgaatcgaaa gaggtgcgga ccaccagtca 1080
caggctgtcc agacacgata aagccgtcca gcggttctgc ctcatgggtt cttctagcaa 1140
acatgaatct gaatacaaca caactagggt ctagaagaaa aattcaagac agcttgagaa 1200
tagtgtgttc ctgaatgatt ttgaacatgc tacagtgaaa agtgacagtg tggaccatgg 1260
aatcaccagc tagagatgag gaaactgaag agtttttagt acttttttaa gattacacaa 1320
taaacatga tgaatcaagc tttgaagcca acctaccaa ccacaagatc aaccaacact 1380
cttcaccaat gtgtaataa accacgttaa tattcaaat agtacgtact gotgaagaaa 1440
gttgataact attcatatta acccgtagt tttgtgttcc ctcatctgta aaagtatgta 1500
ttataacaco tttctctcac cttacagcgt gtgagggtca aatgaccatt cattggagaa 1560
tattttttat atcctataat gcattataaa aataaatcat ttttctaaa aaaaaaaa 1618

```

<210> 28

<211> 3269

<212> DNA

<213> Homo sapiens

WO 02/34783

PCT/US01/49670

<220>

<221> misc_feature

<223> Incyte ID No: 2720058CB1

<400> 28

```

cctgaggga atggctttcg tcccttccct ctgggtgacc tgggtgtag ccgccttcat 60
tatctctctac gtggtgoccg tgctctccgg gcacgtcaac ccttctctcc cgtatatcag 120
tgatcacggga acaacacctc cagagagtagg tatttttggg tttatgataa acttctctgc 180
atttcttgggt gcagccacga tgtatacaag atacaaaata gtacagaagc aaaaacaaac 240
ctgctatttc agcactcccg ttttttaactt ggtgtcttta gtgcttggat tgggtggatg 300
tttcggaatg ggcattgtcg ccaattttca ggagttagct gtgccagtg ttcagtgcag 360
gggcgtcttt ttggcctttg tctgtgtgtg ctgtacacg ctccacaggt ccatcatctc 420
ttacaaatca tgtccccagt ggaacagctc ctgcacatgc cacatacgga tggatcatct 480
tgccgtttct tgcgcagctg tcatcccatc gattgtctgt gcttcaacta ttccataaac 540
caactctggag tggatcccaa gagaaaagga ttatgtatat cactgtatga gtgcgactcg 600
tgaattggaca gtggcctttg gttttatttt ctacttctca acttctatcc aagatttcca 660
gagtgtaacc ctaagcatat ccacagaaat caatggtagt atttgaagaa agaagaattc 720
agtcctcaact agtgaatgtc gcaggccatt tctaaaagtg ctacagagga cagacagggg 780
ttttagggcca ccttgattat tgggatgcac ctgcagcaca tccaggactt gaatttcatc 840
acgagttctc aatagttgta tttctaaaga tgtgtttcct agagaatgta cagccttatg 900
acactctagtg gatgttttta taatttttca agtagatttt tttatattaa caaattcata 960
tacagaaaaa ataagggtgt acaaaaaatg gagagctctt atttttgtac agattctgtc 1020
gttttgtgtt tatttgtgtg agattttatg aaatacacta aatgagtaat tcaggttccg 1080
taactttatt acaaaagtga acaggggatg attcatttgt aaattttatt cttagtgaat 1140
gaactgtata atttttttta tcaggagagc acttataaaa tcaattttat aaagatcata 1200
taaccaaatc ataaagattt agttgatata ttaacactaa gatactctga ttttttagccg 1260
aaataaacac agtgccttca ctgagaggcc tttataccac catgtacagt aactotaagt 1320
gaatccggaa gcccttggtt ttgaatttct gccaccttgt ttctccctgc tcatgaggtc 1380
gcaccttttg ctctgtctgc taattgccca ttgtagtggg gtgtaatgoc aggtggaaag 1440
gtttccacaa gtccagtgaa aaccatcctt tattgttgct ggcaacactt gatatatagt 1500
ctgactcaga actgaagctc acatctcaaa ttcatttcat gccagtaaat gtggcaaaag 1560
gaagaaggcc ccaagagcga gacaagaaga atggagaagg gggcagccaa gaagaacttc 1620
tgggttccag gtactgttta ttgtctctt ctcttctatc ctgtggctgg atgtccacca 1680
acactataag aaataagta agccttttgt gtttagcag aactacagac tcatctttt 1740
caccacaatc atgaatgacc aataaaaagc aagtatttcc agaggaagaa gcagcccttg 1800
aaatgttaag gcttaggctt gaaaggtgaa gagcaggaaat tctctcttcc aaatccatga 1860
gcataaacc atgtgtggcc aagtgtagat agccctcaag ggcacatgcc aaggggcagag 1920
cagcccatgt agacagcttc ggaaggcatg ggggtgtagg gagttcgagg tagctcccca 1980
ttaactattt gttgggtgag taaagggttg aggtctagtg gcaggtacct ctgcaatgac 2040
aagctgcctc cctctatgtt gtttagcata tgttattaga acatgtccga cccccctacc 2100
gtgcacattt gggcccttta ataaagccaa gttagagaaat ctggcaataa aaggcaaatg 2160
taagcatgct ttctttaaga cgcatacata atggttttct ttaagtgaat ggaagagttt 2220
gacagagata cacttttgta agaaaacatt aagaatgctg gctggctgtg gtggctcaca 2280
cotgtattcc cagcactttg ggaggcctag gcaggaggat tgcttgagcc tgggaacttcg 2340
agaccagact gggaaacatg gcaaaatccc atctctacaa caaaaataca aaattatgac 2400
aagtgcgggt gtgtgcctgt agtcttagtt acttgggagg ctgaggtggg agaataacct 2460
gagccacgga ggtggaggct gcagttagcc atgccaatgc actccagtct gggcaacaga 2520
gtgagaccct gtctcaaaaa taataaataa aataaatgaa taagagaatg gctaactatt 2580
ctcgggttca ctgcgactca ctgtatgctt ggggatcccc ctgttaacac tggaaactgaa 2640
agacagtgat gaaagctatg tcaagcatc attattctga agaggaggag aatgccaaca 2700
taoctttccc atgggacctg tgggtgaatg aatccatact tctgcctcac ttgagcaga 2760
cttttgtctt cgggcctctc cagcatggag ttctatgctt cattttcaca tctctctgca 2820
caattagatt gggagctctc tgagggcaga gtacgtgcct taattcttat ctttgtaatg 2880
ccacaatgaa cagagtgctc cctggtacac tgtaggagct taagaataac tcactgaatg 2940
catgaatgaa tgaatgaaca aatgaaggaa tgactaagga tgtttgtagt gctataaat 3000
agaatgggat ttactctgct ttaccagtta gtttcataat aaacaaatag ctgttaacag 3060

```

WO 02/34783

PCT/US01/49670

```

aacattotgt acctgccata caggctcatg ttcattccaa ttcttcttag agccaaataa 3120
ataaagactt agggggggcc ccgnaaaaaa gggccggcgc cccgggggtt atctccggcc 3180
cgggcctgaa gcgcacacgc gttccccaag gytacagctt tcccctgggg ggaactcagg 3240
gaacgggggtt ccccggggca atttttacc 3269

```

```

<210> 29
<211> 1227
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7481255CB1

```

```

<400> 29
atggacaggg ccaagcagca gcaggcgctg ctctctctcc ctgtctgctt cgcctcacc 60
ttctccctca ccgcgtgggt cagcagccac tgggtgtgag ggaccggacg ggtggtgaag 120
ccactgtgcc aggaccagcc gggaggccag cactgcattc acttcaaacy ggacacacgc 180
agcaatggca ggaatggaca caatagccag gctgtcctgt acatttggga gctgggtgat 240
gacaaagtta ttacagcggtg ttccatgttg gggctctggc agtctctgga ggagagcttc 300
aacggtgaag atgaaaagtg taggagtttc cggagtgtag tgccagctga agaacaaggt 360
gttttctggc tgtccatcgg gggcgaggto ctggatatcg ttctgatact gacaagcgcc 420
atctctctgg gctccagagt gagtgtgcgc agccctgggt tccactggct cagggtggtat 480
gccttggtag ccatcttcat ggtgtggca gggcttctag gcactgtggc ccacatgatg 540
tacacaacca tttttcaaat cactgtgaac ctggaccag aagattgga gctccagacc 600
tgggactatg gctgtctata ttgccttgcc tggggttctt tgcctctctg cctggctgtg 660
togggtctcg coatgagcag gttcaaggca gccgcctagg aattcacga gaagcagcag 720
gcacgaagcg gcagtcggca ctctcaacac agtctctggt aaccggaggg ttccggagagc 780
atttggaaaa caggagctgc tccttgccct gotgaacaag ccttcaggaa tgtttctgga 840
cactctccac caggcgcccc agycaaggty tccatatgct agccagtgtc catggctgcc 900
acatccgcac agycaaacaa gccaggcact gacactcaca atgtacacc tgcctctggg 960
ttggacttca ggagattgtt gtccaggga agcttccat cccacccctc cacatctcac 1020
ctttacttaa acaccttttg ctttagcctt tgattctctg taaatgcca gtacottgaa 1080
gtgagataat gcttactgaa gatatacaac attgacactc tagtataaaa gagagcttct 1140
taatgacagt gaatttgata aggatacca aagaacaggg aggatgccag tactaaggga 1200
agagaaggtg aagaagagg aaagcaa 1227

```

```

<210> 30
<211> 2618
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1510242CB1

```

```

<400> 30
agcctaatac cttctcaagt tgatctcccc ccaggcacag ccttgtccca gcggaagac 60
tcaaatatta aaatttcgaa ttctgaatag ttattctatg tatataagtt actgacacag 120
taagagggat ctttttttta tcttacaaga cctaaaaatt acttaatacc ttgaaataa 180
aatgttttat ttctgcocaa tggcattatg aatataataa gacttaagag caccaaaagt 240
tagttactac agcaagatac actagtatac gtatatctat ttatattaag aaactcaagg 300
caottgtcta taattcacia gttaccaatc ttaaacattt aaggcgacgc ccgcagatcc 360
gcagtagttc gggccatgga ggcggagccg ccgctctacc cgatggggg gctgcgggg 420
ccgcagggcg acgagagcct gctcggggtc ccggaagggc ccgaggcccc gctggacgag 480
ctgtgtgggg cytaccccaa ctacaacgag gaggaggagg agcgccgcta ctaccggcg 540

```

WO 02/34783

PCT/US01/49670

```

aagcggcctg ggcgtgtca agaactgct ggcgtccagc gccgggggca tgcctacota 600
cggcgtctac ctgggctctc tcagatgca gctgacctg cactacgacg agaactaccg 660
cgaggtgaag tatggcaaca tggggctgac cgacatcgac agcaaatgc tgaatggcat 720
caactgtact cccatgcgcg cctgctcta cacactgtg ctcatcaggt tttttgganc 780
gaagtggatg atgttctctg ctgtgggcat ctacgacctc tttgtctcca ccaactactg 840
ggagcgctac tacacgcttg tgcctgggc tgtggccctg ggcattggca tctgtcctct 900
ttgggtcttc atgggcaact acatcaccg gatggcgag aagtaccatg agtactccca 960
ctacacggag caggatgggc aggggatgaa gcagcggcct ccgccccgct cccacgcgc 1020
ctatctctg gtcttccaag ccattctcta cagcttctc catctgagct tgcctgcgc 1080
coagctgccc atgatttatt tcttgaacca ctacctgtat gacctgaacc aacgcgtgta 1140
caatgtgcag agctgcggca ccaacagcca cgggacctc agcggcttca acaagacggt 1200
tctgcggagc ctcccgcgga gcggaaacct cattgtggtg gagagcgtgc tcatggcagt 1260
ggcctctctg gccatgctgc tgggtctggg tttgtgggga gccgttacc gcccaacgga 1320
ggagatcgat ctgcgcagcg tgggtctggg caacatcttc cagctgcctc tcaagcacgt 1380
gcgtgactac cgcctgcgcc acctcgtgac tttctttatc tacagcggct tcaaggtgct 1440
ctttgcctgc actgttatcg ccttgggcta tggcgtgtgc tgggtggggt tggagcggct 1500
ggcttacctc ctgctggctt acagcctggg cgcctcagcc gcctcactcc tggcctgct 1560
gggctctggt ctgcacgcgc cgttccccct ggtggccgga gcaggggtgc acctgctgct 1620
cacctctatc ctctttttct gggccccgtg gctcgggtc ctgcaacaca gctggatcct 1680
ctatgtggca gctgcctctt ggggtgtggg cagtgcctct aacaagactg gactcagcac 1740
actcctggga atctgtgtac agacaagga gagacaggac tctcatctta ccatctacca 1800
ctgtgtggca gctgtgtcca tcttcaacct gtacctgggc tgcagcctgc acatgaaggc 1860
taagctggcg gtgctgtggt tgacgtggtt ggccggcgcg gtctcttacc tgcggatgga 1920
gcgaagctg cgcgcggcgg tggccccgcg ccagccccgc atccccgcgc cccagcacaa 1980
gggtgcgggt tacgcgtact tggaggagga caactcggac gagagcgacg cggagggcga 2040
gcctggggac ggcgcggagg aggagcgccc gcccgaggc cccagggcgt gccccgagcc 2100
cgctggactc ggcgcgcggc cctgcgcgta cgaacaggcg caggggggag acgggcccga 2160
ggagcagtga ggcgcgcctt ggtccccgga ctacgctccc ctctgcgcc gccctcagtt 2220
accactctcg aggtcggggg gaaccccccc gactccccgc ctgtcttcaa aggcctctgt 2280
ctccccccc ccactgtggg gacgcccccc ccagagccca ggtcacctcc gggcttcgcg 2340
agccccccc aaggcggagt ggcccttgg gaaccccccc gccaaacaca ggggttcgaa 2400
atacagctg aaaaacggcg ggcccttagc agcgccccca gcgcgggagc acggtcaggg 2460
tctcttctcg acccgcccg ctccagatcc ccacagctct cggccgcgga cccgggcgcg 2520
gtgtgagcgc actttgcacc tcttatcccc agggctccgc gagagccacg attttttaca 2580
gaaaatgagc aataaagaga tttgtactg tcaaaaaa 2618

```

<210> 31

<211> 2188

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 162131CB1

<400> 31

```

ccggaattga ccaactggta gactgccta gaggggacgc attgtgtcct agttgaggct 60
aacagtcagt atccagcctc aacattcagc agaggcccca gatcagcgtc tgagccaggc 120
caacaatgac caaggaggat gggatcctgt gtgcagctca tcacaagcgt cggggtgcag 180
caaaaccatc caggctggac agtggctgga cagttcaaag aaagaagaac cttcaatgaa 240
gaagtcatgg aatacttcca gaagaaagtt agccccagtc atctgaaatc cctgctgact 300
agcgtatgaag cctggaagag attcgtgctg gtggctgaaat tgcccaggga agaagcagat 360
gctctctatg aagctctgaa gaattctaca ccatatgttg ctattgagga caaagacatg 420
cagcaaaaag aacagcagtt tagggagtgg tttttgaaag agtttctcca aatcagatgg 480
aagattccag agtccataga aaggctctgt gtcatgcaa atgagattga aaaggtccac 540
agaggctcgc tcatcgccaa tgtgtgtctt ggctccactg gcactcctgt tgtcattggc 600

```

WO 02/34783

PCT/US01/49670

```

gttatgttgg caccatttac agcagggtcg agcctgagca ttactgoagc tggggtaggg 660
ctgggaatag catctgccac ggctgggacg gcctccagca tctgtgagaa cacatacaca 720
aggtcaggag aactcacagc cagcagggctg actgcaacca gcactgacaa attggaggca 780
ttaaggagaa ttctgcatga catcacccc aatgtgcttt cctttgcaat tgattttgac 840
gaagccacaa aatgatttgc gaatgatgtc catcacctca gtagatctaa agccactgtt 900
ggaagccctt tgattgtctg gcgatagtga cctataaatg ttgttgagac actgagaaca 960
cgtggggccc ccaccgggat agtgagaaaa gtagcccgga acctgggcaa ggccacttca 1020
gggtgctcctg ttgtgctgga ttagtcaaac cttgtgcaag actcactgga cttgcacaag 1080
ggggaataat ccgagctctg ttagtctctg aggcagtgga ctcaggagct ggaggagaat 1140
ctcaatgagc tcacccatat ccatcagagt ctaaaagcag gctaggccca attgtgtcgg 1200
gaagtcaggc accccaaacy gagggaactg ctgaagccat ggcagaagaa cgtggattgt 1260
gaagatttca tggacattta ttagttcccc aaattaatac ttttataatt tcctatgcct 1320
gtctttaccg caatctctaa acacaaattg tgaagatttc atggacactt atcacttccc 1380
caatcaatac ccttggtgatt tcttatgctt gtctttactt taatctccta atcctgtcag 1440
ctgaggaggga tgtatgtcac ctcaggacca tgtgataatt cgtttaactg cacaattgt 1500
agagcagtgt tgtttgaaca atatgaatc tgggcaacct gaaaaaagaa caggataaca 1560
gcaattgttc aggaataaag agagataacc ttaactctg accaactg agcctgggtg 1620
aacagagbca tatttctctt ctttcaaaag caaatgggag aatatcgct gaattctttt 1680
tctcagcaag gaacatccct gagaaagaga atgcacccct gagggtgggt ctataaatg 1740
cctccttggg ttgtggccatc ttctatggtc gagactgtag ggatgaataa aacccagtc 1800
tcccatagtg ctcocaggctc tattaggaag aggaatttcc cgctaataa attttggta 1860
gacoggttgc tctcaaaacc ctgtctcctg ataagatttt atcaatgaca atggtgcctg 1920
aaactcattt agcaatttta atttctcccc ggtcctgtgg tctgtgtatc tcacctgtcc 1980
tccactgtcc ttgtgatatt ctattacctt gtgaagtagg tgatctttgt gacccacacc 2040
cacacctat tcatacactc cctccctttt tgaagtcoc taataaaaaa ttgtgtgttt 2100
tcagctctgt gaggcatcac ggaacctacc gatgtgtgat gtctccctgt gacacctagc 2160
tttaaaattt ctaaaaaaaa aaaaaaaa
2188

```

<210> 32

<211> 1969

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1837725CB1

<400> 32

```

gtagcagcgg cggctcagtc gtagcccgcc cgcgcgcgcc tgtccggtcc ggtccggcca 60
cggaggcagc gcagcggcgg gactccgagc ctaccccgcc gactgagctg cgcgcgcgcc 120
tgccgtccca ccgggcaccc accagtcgga tggggcgcca gcgcgcgctg tccctgtccg 180
gggcgcgcct actctggggc ttctgtctcc agctgacagc cgtcaggaa gcaatcttgc 240
atgcgtctgtg aatgggcaca accaaggact actgcatgct ttataacccc tattggacag 300
ctcttccaag taccttagaa aatgcaactt ccattagttt gatgaatctg acttccacac 360
cactatgcaa cctttctgat attctcctg ttggcataaa gagcaagcaa gttgtgggtc 420
catgggggaag ctgccatttt cttgaaaaag ccagaattgc acagaagga ggtgctgaag 480
caatgttagt tgtcaataac agtgtcctat ttctccctc aggtaaacaga totgaatttc 540
ctgatgtgaa aatactgatt gcattttataa gctacaaaga ctttagagat atgaaccaga 600
ctcaggagaa taacattact gtgaaaatgt attctccatc gtggcctaac ttgtattata 660
ctatggttgtt tatttttgtt attgcggtgt tcaactgtgc attagtgga tactggagtg 720
gactagttag atttgaaaaa ttgaagcag tgacaactga agatagagaa atgaggaaaa 780
agaaaggaga atatttaact tttagtcttc ttacagttgt aatattttgt gtcactgtct 840
gtgttatgat ggtcttactt tatttctctc acaaatgttt ggtttatgtt atgatagcaa 900
ttttctgcat agcatcagca atgagctgtt acaactgtct tgcgtgacta attcataaga 960
taccatatgt acaatgcacg attgcatgtc gtggcaaaaa catggaaagt agacttattt 1020
ttctctctgg actgtgcata gcagtagctg ttgtttgggc tgtgtttcga aatgaagaca 1080

```

WO 02/34783

PCT/US01/49670

```

gggtgggcttg gattttacag gatattcttg ggattgcttt ctgtctgaat ttaattaaaa 1140
cactgaagtt gcccaacttc aagtcatttg tgatacttct aggccttctc ctctctatg 1200
atgtattttt tgttttcata acaccattca tcaacaagaa tggtagagtg atcatgggtg 1260
aactgcagcg tggacctttt ggaataatg aaagttgccc agtagtcatc agagtaccaa 1320
aactgatcta ttctccagta atgagtggtg gctcatgccc tgtttcaata ttgggttttg 1380
gagacattat tbtaccaggg ctgttgattg catactgtag aagatttgat gttcagactg 1440
gttcttctta catatactat gtttcgtcta cagttgccta tgcctattggc atgatactta 1500
catttctgtt tctggtgctg atgaaaaagg ggcacactgc tctctctat ttagtacctt 1560
gcacacttat tactgctcca gtgttgctct ggagacgtaa ggaatgaaa aagttctgga 1620
aeggtaacag ctatcagatg atggaccatt tggattgtgc acaaatgaaa gaaacccctg 1680
tgatatctgg tgaacagatt gtccagcaat aatattatgt ggaactgcta taatgtgtca 1740
ttgattttct acaaatagac ttgcactttt taaattgact ttgaattga caatctgaaa 1800
gagttctcaa tgatatgctt gcaaaaatat atttttatga gctggtactg acagttacat 1860
cataaataac taaaacgctt tgcttttaat gttaagtttg tgcttcaca ttaataaaaa 1920
catatggtct gtgtagtctt cgaaaaaaaa aaaaaaaa aaaaaaaa 1969

```

<210> 33

<211> 3006

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3643847CB1

<400> 33

```

gcatgcaggg cggcgcgctg ggaactacat gctccctggt gggtcgtgtg gctgcacagc 60
gtcccgacag tggcctctgc cctgcagccc gtgaacagca ccttcagccc cggcgacgag 120
agtacacagg agtcctctgt gttcctgggg ctggtggcgc cgtctctgct gggcctgaac 180
ctcatctccc ttgtggctta cctgctctgt gcatgcccact gccggcgagg cgtatcggtg 240
cagacccaagc agcaccactc ctgctgcatc aactggacgg cgtggttggc cgggctcctc 300
tgctgtgctg cgttggggct tggtttctat ggaaacagcg agaccaagca tggggcgtag 360
cagctgtagt actccttggc cgaatgcaac cacacttctc ctgggatoga tgcctcggtt 420
tcoggaacta cccagaagat gaagtyggac ctatagcagc acctggcccg gctcagtag 480
atctttgtct cccggggcga ttacctgag accctgaagt tcatacagca gatggcgagg 540
agcgtttgtt ttacgtcttc aggaactgcc gtgtggaggg aggtcaccat ggagctgacc 600
aagctatccg accagactgg ctacgtggag tactacaggt ggctctccta cctcctgctc 660
tttatcctgg accctggtcat ctgctctatt gctgctctgg gactggccaa gcctccaa 720
tgtctcctgg cctcgatgct gtgctgtggg gcaactgagc tgctcctcag ttgggcatcc 780
ctggccgctg atggctctgc ggcagtgccc accagtgaat tctgtgtggc tcttgacacc 840
ttcatcctga acgtcacgga gggccagatc agcacagagg tgactcgtca ctactgtat 900
tgacgcaga gtgggaagcag ccccttcag cagacccctg caaccttcca gcgcgcactc 960
accacatgc agatccaggt cggggggctg ctgcagtttg cgtgcccctc cttctccact 1020
gcagagggaag accctgttgc aatccagctc ctgctgaact cctcagagtc cagccttcac 1080
cagctgactg ccatgttggc ctggcggagg ctgcacaagg attatctgga cgtcttctgt 1140
ggcatctgct acgacggcct ccagggttg ctgtaccttg gctctctctc cttcctggcc 1200
gcctcgcctc tctccacct gatctgtgct gggccaaggg cctggaagca cttcacacc 1260
agaacacagc aatacgaatga cattgatgat gatgacccct ttaaccccca agcctggcgc 1320
atggcggtgc acagtcccc gaggggacag cttcacagct tctgcagcta cagcagtggc 1380
ctgggaagtc agaccagcct gcagccccc gcccagacca tctccaagc cctgtctctc 1440
gagtaacatga accaagccat gctcttttgt aggaacccac gctacagaaa cgtgccacta 1500
atcgggagag cctcccccct cctacgttac tctccagcaa tgagagccac ctactgtct 1560
gtggcggtat agcaccctgag gcaactcggg aatcagtttc cagcctaaca gactttcggg 1620
ggttccctgc tcctttttcc gttctggttt ttaaktatgt caaatacaag ctgctttct 1680
ttaatagaaa ccaaagccat ctggagcccg agaggcctcc tgctgtggca gaggagcagc 1740
tgggattccc gaccaagcc ccagggggtg cagaagactc accacgcggg ccagcctctc 1800

```

WO 02/34783

PCT/US01/49670

```
tcttttgccc tgetctccac accagaaatg cccccaggtg cttggctgoc tcagaggtac 1860
catccctgag ctggctgcoct ggccctgctc acccctacgc ctggcccttg ccaggagggg 1920
agtggcagtg agggagggggc caggctcaggc accaccatca agagagctgt gtgttctctc 1980
tggctccaca acgatgactc tgccctctgt cagccacgc agagccacag acgacccctc 2040
tgtccctgct cctctgctcct gttccctgca ggttaacatga gaagggtctga tcaggagatg 2100
ctctttaaga agttcgcacc cctgctgaca ccagaacagc ccaaatcaga gttccoaagg 2160
ccagacagggc tottccctggg ccacagagggg gaggcatcag gaaagctctg cagtgggggg 2220
ctgggtggctc cggggctggg ggtacacagg ctggtgaacc ccgggtggaa ccagagtgaa 2280
agcctgcccac attccgcctg totccctaac cctccattgc ctggcctcta ttccagaatc 2340
aatgctgcag aatgtgttag ctgcagatag gcatggtctc aggtatgaac agacactttg 2400
aaacgacttt aggtctttct tttctccagt gttttaacaa tgttgattat ccaagaatc 2460
gaactccta gcacatccag tttttacaac agatttgag ctoattcctt accctgggta 2520
ggtcactact tttgcagatt ttgctggcac tgatctggag atctgcagat ctggaggaga 2580
cgggaaggag togtatttta aataaggatc agtgaggcat cctgtcccaa gctactgttt 2640
ggtggggatc tgggttcac toaccacag agggaggatc ttaagagga gaaaaagcc 2700
aagagggaac gccagagttc cctgttctag gggactagcc aaatgcctac atcagctgtc 2760
cctccctgtt tgtctccag taagtgttgc agaaaaggtt ttagcaagt gctacaactg 2820
tgtctttata ggaggatagg cctctgccct gcccccaccc caccacctgt cccacccag 2880
tgtcccaggc cacaggagct tatggccag gagggaataa tgtcccccac tactgcctgt 2940
tgaggggacca gagttggggt ctttgggtgct tccaacctcc tgccaacctg gagtccaca 3000
caccaag 3006
```

<210> 34

<211> 2884

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6889872CB1

<400> 34

```
tcactcctgg ctacgtgagg cactctccag cctcctgtgg gaatcatctg aagttctgag 60
cccggagagc aaggaggaag acgaggagga gggaggagg gaggagagg agggaggagg 120
agaggagagc aagccctgag aacccctgca cttccttagc aggagacag gagcaacgct 180
ggcgtggggg gcaggctgtg gggcccccac cccagccct agccaggcct agtgccctgt 240
gtagcaccct agaagatccc cagcagttgg cactagctgt acccaacctg cctggggccc 300
cgtgctgggg ggtgcgcccc aagatggtgg cggcccagg gaggactgta ctgccagccc 360
cagcctctgg ccgtcaggca ccccctgctc tgccctggcc cctcaactcc agggccagcc 420
catgctgcgc ctggggctgt ggcggcgcc gctgctgtgc gtgtgccgc cgggtgccgt 480
gcgtgcgcac tgcgtgctca ttgagggcga caagggtac gtgtggctgg ccactctgag 540
ccagaaccag ccgcccacag agaccatccc gcagcacatc aatagcaccc tgacgacat 600
gcggctcaac gagaacaagc tcaaaagcgt gctctactcc tcgtcaacc gcttgggaa 660
cctcaaccag ctcaacctca ccaagaacga gatctcctac atcgaggagc gtgccttct 720
gggcccagtc agcctgcagg tccctgcagct gggctacac aagctcagca acctgaagga 780
gggcatgctg cgaggcatga gccgcctgca gttcctcttt gtccagcaca acctcatga 840
ggtggtgacg cccaccgctc tctccagtg cccagcctc atcagcatcg acctgtctc 900
caaccgctc agccgcctgg acggtgccac ctttgccagc ctgccagcc tgatggtgtg 960
tgagctggcc ggcacccct tcaactgtga gtgcgacctc ttggcttcc tggcgtggt 1020
ggtggtcttc aacaacgtca ccaagaacta cgaacgctg cagtgtgagt cggccggga 1080
gtttgcggcc taccctgtgc tgggtccccc gccctaccac agcctcaacg ccatcaacgt 1140
actccaggcc aagtgtcgga atggtctgct gcccgccgg ccogtgagcc accccagcc 1200
ctactccacc gacgcccaga gggagccaga cgagaactcg ggttcaacc ccagcagat 1260
cctttcggtg gagcgcgcgg cctcgtccac cacygatgcg tcggcagggc cagccatcaa 1320
gctgccacc gtcacgttca cctcgccac cctggtgtgc atcattccac accctacag 1380
caagatgtac atcctgtgc agtacaacaa cagctacttc tccgagctca tgaccctcaa 1440
```

WO 02/34783

PCT/US01/49670

```

gaacaagaag gagatcgtga cgctggacaa actgcgggag cactactgagt acacottctg 1500
cgtgacctcg ctgcgcaaca gccgcgctt caaccacacc tgctgacct tcaccacgag 1560
ggaccccgct ccgcgagact tgggcgccag cactcacacc accaccact acatcatgac 1620
catcctgggc tgcctcttct gcatggttat cgtgctggga gccgtgtact actgcctgac 1680
caagcggcgc atgcaggagg agaagcagaa gtctgtcaac gtcaagaaga ccattcctga 1740
gatcgctac ggggctgatg tggatgcgg ctccattgtg cagcccgccc agaagctggg 1800
cgaacctccc gtgctgcccg tatctcgcat ggcctccatc cctccatga tcggggagaa 1860
gctgcccacc gccaaaggggt tggggccgg gctggacaca cccaaggtag ccaaccaagg 1920
caactatata gaggtagcga caggcgccg cggggacggt ctggctcggc ccagggatga 1980
cctcccggac ctgcgagacg gccagggttc ggctgcagag atctccacca ttgccaagga 2040
ggtagacaag gtcaaccaga tcattaacaa ctgcctcgat gctctcaagc tggactgggc 2100
ctcttttctg ggaggcgga gcagcagtg ggaccccgag ctggccttgc agtgccagtc 2160
cctccctgca gctgctgccc cctcctcagc cactggcccc ggggccttgc agcggccag 2220
cttcccttgc cctccctaca aggagagctc ccaccacca ctacagcgcc agctgagcgc 2280
cgacgcggcc gtgacctgca agacctgcag cgtgtcgtcc agtggttcca tcaagagcgc 2340
caaggtcttt agcctggacg tgcccaccca tcgggcggcc acagggtcgg ctaagggcga 2400
ctccagctac atcgagaaag gcagccccc caacagcccg ctggaccgga tcccgctggt 2460
gccggcgggc agcggcgggg gcagcggcg gggcgggggc atccaccaco tggagggtgaa 2520
ctccagctac cactgcagcg agcaccggca cagctttccc gccctgtaot acgaggaggg 2580
tgccagcagc ctgagcagcg cgtgtcctt cctcaagcgg ctgacccgct ccaagcgtga 2640
ctccagctac tcgagcctct ccccagaca ctactactca ggtactcct ccagcccgga 2700
gtactcatcc gagagcagc acaagatctg ggagcgttc cggccctaca agaagccaca 2760
ccgggaggag gtgtacatgg ccgcgggtca cgcctgcgc aagaaggtcc agttcgccaa 2820
ggacgaggat ctgcatgaca tccttgatta ctggaagggg gtctccgccc agcagaagct 2880
tgga

```

【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
2 May 2002 (02.05.2002)

PCT

(10) International Publication Number
WO 02/034783 A3(51) International Patent Classification: C07K 14/705,
C12N 15/12, C07K 16/28, G01N 33/50, C12Q 1/68

(21) International Application Number: PCT/US01/49670

(22) International Filing Date: 26 October 2001 (26.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/244,017 27 October 2000 (27.10.2000) US
60/252,855 22 November 2000 (22.11.2000) US
60/251,825 7 December 2000 (07.12.2000) US
60/255,085 12 December 2000 (12.12.2000) US(71) Applicant (for all designated States except US): INCVTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): WARREN, Bridget,
A. [US/US]; 10130 Parkwood Drive #2, Cupertino, CA
95014 (US). XU, Yuming [US/US]; 1739 Walnut Drive,
Mountain View, CA 94040 (US). YUE, Henry [US/US];
826 Lois Avenue, Sunnyvale, CA 94087 (US). BATRA,
Sajeev [US/US]; 555 El Camino Real, #709, San Leand-
ro, CA 94087 (US). BURFORD, Neil [GB/US]; 105
Wildwood Circle, Durham, CT 06422 (US). GANDHI,
Ameena, R. [US/US]; 705 5th Avenue, San Francisco, CA
94118 (US). CHAWLA, Narinder, K. [US/US]; 33 Union
Square, #712, Union City, CA 94587 (US). ARVIZU,
Chandra [US/US]; 490 Sherwood Way #1, Menlo Park,
CA 94025 (US). TANG, Y., Tom [US/US]; 4230 Ranwic
Court, San Jose, CA 95118 (US). LU, Dyung, Aina, M.
[US/US]; 233 Goy Drive, San Jose, CA 95123 (US). DUG-
GAN, Brendan, M. [AU/US]; 243 Buena Vista Avenue
#306, Sunnyvale, CA 94086 (US). BAUGHN, Mariah, R.
[US/US]; 14244 Sanliago Road, San Leandro, CA 94577
(US). LEE, Ernestine, A. [US/US]; 624 Kains Street, Al-
bany, CA 94706 (US). KHAN, Farrah, A. [IN/US]; 9445
Harrison Street, Des Plaines, IL 60016 (US). NGUYEN,Danniel, B. [US/US]; 1403 Ridgewood Drive, San Jose,
CA 95118 (US). AZIMZAI, Valda [US/US]; 5518 Boul-
der Canyon Drive, Castro Valley, CA 94552 (US). YAO,
Monique, G. [US/US]; 1189 Woodgate Drive, Carmel,
IN 46033 (US). LAL, Preeti, G. [IN/US]; P.O. Box 5142,
Santa Clara, CA 95056 (US). THIANGAVELU, Kavitha
[IN/US]; 1950 Montecito Avenue #23, Mountain View, CA
94043 (US). RAMKUMAR, Jayalaxmi [IN/US]; 34359
Maybird Circle, Fremont, CA 94555 (US). TRAN, Bao
[US/US]; 750 Salberg Avenue, Santa Clara, CA 95051
(US). DING, Li [CN/US]; 3353 Alma Street #146, Palo
Alto, CA 94306 (US). AU-YOUNG, Janice [US/US]; 233
Golden Eagle Lane, Brisbane, CA 94005 (US).(74) Agents: HAMLET-COX, Diana et al.; Incey Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DL, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MY, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.(84) Designated States (regional): ARIPO patent (GI, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, HU,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG).Published:
with international search report(88) Date of publication of the international search report:
30 October 2003For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 02/034783 A3

(54) Title: TRANSMEMBRANE PROTEINS

(57) Abstract: The invention provides human transmembrane proteins (TMP) and polynucleotides which identify and encode TMP.
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 TMP.

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/49670
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C07K14/705 C12N15/12 C07K16/28 G01N33/50 C12Q1/68		
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 G01N C12Q		
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) SEQUENCE SEARCH, MEDLINE, BIOSIS, EPO-Internal, WPI Data, PAJ, EMBL		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 06846 A (HAWKINS PHILLIP R ; INCYTE PHARMA INC (US); AU YOUNG JANICE (US)) 19 February 1998 (1998-02-19) 82% identity in 461 aa overlap to SEQ ID NO:2 of present application references to antibodies	11,31, 32,34, 37,38, 40-43
X	DATABASE EMBL 'Online! 17 September 1999 (1999-09-17) "Homo sapiens chromosome 5 clone CTC-45812, complete sequence" Database accession no. AC010260 XP002221760 99.7% identity in 1322 nt overlap to SEQ ID NO:18 of present application -/-	13
☒ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "Z" document member of the same patent family		
Date of the actual completion of the international search 21 November 2002		Date of mailing of the international search report 13.03.03
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel: (+31-70) 340-2040, Tx: 31 651 epo nl, Fax: (+31-70) 340-3046		Authorized officer Herrmann, K

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/49670
C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 00 56891 A (INCYTE PHARMA INC ;AZIMZAI YALDA (US); REDDY ROOPA (US); YUE HENRY) 28 September 2000 (2000-09-28) the whole document	1-20,23, 26-56,73
Y	SAIER M H JR: "Families of proteins forming transmembrane channels." JOURNAL OF MEMBRANE BIOLOGY, vol. 175, no. 3, 1 June 2000 (2000-06-01), pages 165-180, XP002221812 ISSN: 0022-2631 the whole document	1-20,23, 26-56,73
Y	DATABASE GENBANK 'Online! 26 October 2000 (2000-10-26) retrieved from NCBI XP002221869 500 human ESTs putatively encoding transmembrane proteins disclosed before 27.10.00 the whole document	1-20,23, 26-56,73
A	KRUEGER W H H ET AL: "TP01, a member of a novel protein family, is developmentally regulated in cultured oligodendrocytes." JOURNAL OF NEUROCHEMISTRY, vol. 69, no. 4, 1997, pages 1343-1355, XP002221759 ISSN: 0022-3042 82% identity in 461 aa overlap to SEQ ID NO:2 of present application figure 2	1-20,23, 26-56,73

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/49670

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
see FURTHER INFORMATION sheet PCT/ISA/210
2. ☒ Claims Nos.: 21, 22, 24, 25
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-55 partially; 56, 73 completely

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

International Application No. PCT/US 01 A9670

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 19, 33 and 35 are directed to a method of treatment of the human/animal body or 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.

Claim 7 also encompasses human embryonic stem cells (cf. p. 45, l. 32-33 and p. 58, l. 18-19 of present description). This subject-matter may be contrary to morality in certain PCT member states (cf. e.g. Art. 53(a) of the EPC).

Claim 8 also encompasses transgenic human beings. This subject-matter is contrary to morality in certain PCT member states (cf. e.g. Art. 53(a) of the EPC).

Continuation of Box I.2

Claims Nos.: 21, 22, 24, 25

Claims 21, 22, 24 and 25:
Compounds as such are not sufficiently defined by their mode of action. Therefore, claims 21, 22, 24 and 25 have not been searched because an agonist or an antagonist is neither disclosed nor supported within the terms of Art. 5 and 6 PCT, respectively (Art. 17(2)(a)(ii) PCT).

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 49670

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 56, 73 and partially claims 1-55

Polypeptide comprising an amino acid sequence as in SEQ ID, NO:1 and subject-matter relating to SEQ ID NO:1.
Polynucleotides encoding the polypeptide of SEQ ID NO:1 such as a polynucleotide comprising a polynucleotide sequence as in SEQ ID NO:18 and subject-matter relating thereto.

2. Claims: 57, 74 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:2 and the polynucleotide as in SEQ ID NOs:19.

3. Claims: 58, 75 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:3 and the polynucleotide as in SEQ ID NOs:20.

4. Claims: 59, 76 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:4 and the polynucleotide as in SEQ ID NOs:21.

5. Claims: 60, 77 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:5 and the polynucleotide as in SEQ ID NOs:22.

6. Claims: 61, 78 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:6 and the polynucleotide as in SEQ ID NOs:23.

7. Claims: 62, 79 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:7 and the polynucleotide as in SEQ ID NOs:24.

8. Claims: 63, 80 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:8 and the polynucleotide as in SEQ ID NOs:25.

9. Claims: 64, 81 and partially claims 1-55

International Application No. PCT/US 01 A9670

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:9 and the polynucleotide as in SEQ ID NOs:26.

10. Claims: 65, 82 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:10 and the polynucleotide as in SEQ ID NOs:27.

11. Claims: 66, 83 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:11 and the polynucleotide as in SEQ ID NOs:28.

12. Claims: 67, 84 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:12 and the polynucleotide as in SEQ ID NOs:29.

13. Claims: 68, 85 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:13 and the polynucleotide as in SEQ ID NOs:30.

14. Claims: 69, 86 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:14 and the polynucleotide as in SEQ ID NOs:31.

15. Claims: 70, 87 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:15 and the polynucleotide as in SEQ ID NOs:32.

16. Claims: 71, 88 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:16 and the polynucleotide as in SEQ ID NOs:33.

17. Claims: 72, 89 and partially claims 1-55

Idem as subject 1 but limited to the polypeptide as in SEQ ID NOs:17 and the polynucleotide as in SEQ ID NOs:34.

INTERNATIONAL SEARCH REPORT
Information on patent family membersInternational Application No.
PCT/US 01/49670

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9806846	A	19-02-1998	US 5776732 A	07-07-1998
			AU 4068897 A	06-03-1998
			EP 0918857 A1	02-06-1999
			JP 2001500007 T	09-01-2001
			US 6080841 A	27-06-2000
			WO 9806846 A1	19-02-1998
WO 0056891	A	28-09-2000	AU 3916400 A	09-10-2000
			EP 1210419 A2	05-06-2002
			WO 0056891 A2	28-09-2000

フロントページの続き

(51)Int.Cl.⁷

F I

テーマコード (参考)

A 6 1 P 1/04	A 6 1 P 1/08		4 B 0 6 5
A 6 1 P 1/06	A 6 1 P 1/10		4 C 0 8 4
A 6 1 P 1/08	A 6 1 P 1/12		4 H 0 4 5
A 6 1 P 1/10	A 6 1 P 1/14		
A 6 1 P 1/12	A 6 1 P 1/16		
A 6 1 P 1/14	A 6 1 P 1/16	1 0 1	
A 6 1 P 1/16	A 6 1 P 1/16	1 0 5	
A 6 1 P 1/18	A 6 1 P 1/18		
A 6 1 P 3/00	A 6 1 P 3/00		
A 6 1 P 3/04	A 6 1 P 3/04		
A 6 1 P 3/06	A 6 1 P 3/06		
A 6 1 P 3/10	A 6 1 P 3/10		
A 6 1 P 3/12	A 6 1 P 3/12		
A 6 1 P 5/00	A 6 1 P 5/00		
A 6 1 P 5/06	A 6 1 P 5/06		
A 6 1 P 5/08	A 6 1 P 5/08		
A 6 1 P 5/14	A 6 1 P 5/14		
A 6 1 P 5/40	A 6 1 P 5/40		
A 6 1 P 7/02	A 6 1 P 7/02		
A 6 1 P 7/04	A 6 1 P 7/04		
A 6 1 P 7/06	A 6 1 P 7/06		
A 6 1 P 9/00	A 6 1 P 9/00		
A 6 1 P 9/04	A 6 1 P 9/04		
A 6 1 P 9/08	A 6 1 P 9/08		
A 6 1 P 9/10	A 6 1 P 9/10		
A 6 1 P 9/12	A 6 1 P 9/10	1 0 3	
A 6 1 P 11/00	A 6 1 P 9/12		
A 6 1 P 11/06	A 6 1 P 11/00		
A 6 1 P 13/08	A 6 1 P 11/06		
A 6 1 P 13/12	A 6 1 P 13/08		
A 6 1 P 15/00	A 6 1 P 13/12		
A 6 1 P 15/08	A 6 1 P 15/00		
A 6 1 P 15/10	A 6 1 P 15/08		
A 6 1 P 17/06	A 6 1 P 15/10		
A 6 1 P 17/12	A 6 1 P 17/06		
A 6 1 P 19/06	A 6 1 P 17/12		
A 6 1 P 21/00	A 6 1 P 19/06		
A 6 1 P 21/04	A 6 1 P 21/00		
A 6 1 P 25/00	A 6 1 P 21/04		
A 6 1 P 25/02	A 6 1 P 25/00		
A 6 1 P 25/08	A 6 1 P 25/02		
A 6 1 P 25/14	A 6 1 P 25/02	1 0 3	
A 6 1 P 25/16	A 6 1 P 25/08		
A 6 1 P 25/18	A 6 1 P 25/14		
A 6 1 P 25/20	A 6 1 P 25/16		
A 6 1 P 25/22	A 6 1 P 25/18		
A 6 1 P 25/24	A 6 1 P 25/20		

A 6 1 P	25/28	A 6 1 P	25/22	
A 6 1 P	27/02	A 6 1 P	25/24	
A 6 1 P	27/06	A 6 1 P	25/28	
A 6 1 P	27/12	A 6 1 P	27/02	
A 6 1 P	27/16	A 6 1 P	27/06	
A 6 1 P	29/00	A 6 1 P	27/12	
A 6 1 P	31/04	A 6 1 P	27/16	
A 6 1 P	31/06	A 6 1 P	29/00	1 0 1
A 6 1 P	31/10	A 6 1 P	31/04	
A 6 1 P	31/12	A 6 1 P	31/06	
A 6 1 P	31/18	A 6 1 P	31/10	
A 6 1 P	33/00	A 6 1 P	31/12	
A 6 1 P	35/00	A 6 1 P	31/18	
A 6 1 P	35/02	A 6 1 P	33/00	
A 6 1 P	37/02	A 6 1 P	35/00	
A 6 1 P	37/08	A 6 1 P	35/02	
A 6 1 P	39/02	A 6 1 P	37/02	
A 6 1 P	39/04	A 6 1 P	37/08	
C 0 7 K	1/22	A 6 1 P	39/02	
C 0 7 K	14/705	A 6 1 P	39/04	
C 0 7 K	16/18	C 0 7 K	1/22	
C 0 7 K	16/28	C 0 7 K	14/705	
C 0 7 K	16/46	C 0 7 K	16/18	
C 1 2 M	1/00	C 0 7 K	16/28	
C 1 2 N	1/15	C 0 7 K	16/46	
C 1 2 N	1/19	C 1 2 M	1/00	A
C 1 2 N	1/21	C 1 2 N	1/15	
C 1 2 N	5/10	C 1 2 N	1/19	
C 1 2 N	15/02	C 1 2 N	1/21	
C 1 2 P	21/02	C 1 2 P	21/02	C
C 1 2 P	21/08	C 1 2 P	21/08	
C 1 2 Q	1/02	C 1 2 Q	1/02	
C 1 2 Q	1/68	C 1 2 Q	1/68	A
G 0 1 N	33/15	G 0 1 N	33/15	Z
G 0 1 N	33/50	G 0 1 N	33/50	Z
G 0 1 N	33/53	G 0 1 N	33/53	D
G 0 1 N	37/00	G 0 1 N	33/53	M
		G 0 1 N	37/00	1 0 2
		C 1 2 N	15/00	C
		C 1 2 N	15/00	F
		C 1 2 N	5/00	A
		A 6 1 K	37/02	
		A 6 1 K	37/52	

(31)優先権主張番号 60/255,085

(32)優先日 平成12年12月12日(2000.12.12)

(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,GQ,GW,ML,MR,NE,SN,TD,

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

- (72)発明者 スー、ユーミング
アメリカ合衆国カリフォルニア州94040・マウンテンビュー・ウォルナットドライブ 1739
- (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94087・サニーベイル・ルイスアベニュー 826
- (72)発明者 バトラ、サジューブ
アメリカ合衆国カリフォルニア州94087・サンレアンドロ・#709・エルカミーノリアル 555
- (72)発明者 バーフォード、ニール
アメリカ合衆国コネチカット州06422・ダラム・ワイルドウッドサークル 105
- (72)発明者 ガンディー、アミーナ・アール
アメリカ合衆国カリフォルニア州94118・サンフランシスコ・フィフスアベニュー 705
- (72)発明者 チョーラ、ナリンダー・ケイ
アメリカ合衆国カリフォルニア州94587・ユニオンシティ・#712・ユニオンスクエア 33
- (72)発明者 アービズ、チャンドラ
アメリカ合衆国カリフォルニア州94025・メンロパーク・#1・シャーウッドウェイ 490
- (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95118・サンノゼ・ランウィックコート 4230
- (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95123・サンノゼ・ゴイドライブ 233
- (72)発明者 ダガン、ブレンダン・エム
アメリカ合衆国カリフォルニア州94086・サニーベイル・#306・ブエナビスタアベニュー 243
- (72)発明者 ボーグン、マライア・アール
アメリカ合衆国カリフォルニア州94577・サンレアンドロ・サンティアゴロード 14244
- (72)発明者 リー、アーンステーション・エイ
アメリカ合衆国カリフォルニア州94706・アルバニー・ケインズストリート 624
- (72)発明者 カーン、ファラ・エイ
アメリカ合衆国イリノイ州60016・デプレイン・ハリソンストリート 9445
- (72)発明者 ニュエン、ダニエル・ビー
アメリカ合衆国カリフォルニア州95118・サンノゼ・リッジウッドドライブ 1403
- (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94552・カストロバレー・ボールダーキャニオンドライブ 5518
- (72)発明者 ヤオ、モニック・ジー
アメリカ合衆国インディアナ州46033・カーメル・ウッドゲートドライブ 1189
- (72)発明者 ラル、プリーティ・ジー
アメリカ合衆国カリフォルニア州95056・サンタクララ・ピーオーボックス 5142
- (72)発明者 サンガベル、カピサ
アメリカ合衆国カリフォルニア州94043・マウンテンビュー・#23・モンテシトアベニュー 1950
- (72)発明者 ランクマール、ジャヤラクシミ
アメリカ合衆国カリフォルニア州94555・フレモント・メイバードサークル 34359
- (72)発明者 トラン、バオ
アメリカ合衆国カリフォルニア州95051・サンタクララ・サルバーグアベニュー 750

(72)発明者 ディング、リー

アメリカ合衆国カリフォルニア州 9 4 3 0 6 ・ パロアルト ・ # 1 4 6 ・ アルマストリート 3 3 5
3

(72)発明者 オウーヤング、ジャニス

アメリカ合衆国カリフォルニア州 9 4 0 0 5 ・ ブリスベーン ・ ゴールデンイーグルレーン 2 3 3

F ターム(参考) 2G045 AA25 AA40 BB03 BB20 CB01 CB21 DA12 DA13 DA14 DA36

DA37 DA77 FB02 FB03

4B024 AA01 AA11 BA43 BA63 CA04 CA05 CA06 CA07 CA09 CA10

CA12 DA02 DA06 EA02 EA04 FA02 FA10 GA05 GA11 GA18

GA19 HA03 HA08 HA12 HA14

4B029 AA07 AA21 AA23 BB20 CC03 CC08 FA12

4B063 QA01 QA18 QQ20 QQ42 QR08 QR14 QR32 QR41 QR42 QR48

QR55 QR57 QR59 QR62 QR66 QR69 QR77 QR82 QS12 QS24

QS25 QS28 QS34 QS39 QX01 QX07

4B064 AG20 AG27 CA02 CA10 CA19 CA20 CC24 CE12 CE13 DA01

DA13

4B065 AA26X AA90X AA93Y AB01 AC14 BA02 BA25 CA24 CA44 CA46

4C084 AA02 AA06 AA07 AA17 BA01 BA08 BA22 CA53 DC25 DC50

MA13 MA24 MA52 MA56 MA57 MA59 MA60 MA63 MA66 ZA022

ZA052 ZA062 ZA082 ZA122 ZA162 ZA182 ZA222 ZA242 ZA332 ZA342

ZA362 ZA392 ZA402 ZA422 ZA452 ZA522 ZA542 ZA552 ZA592 ZA662

ZA682 ZA692 ZA702 ZA712 ZA722 ZA752 ZA762 ZA772 ZA812 ZA892

ZA942 ZA962 ZA972 ZB072 ZB082 ZB132 ZB152 ZB262 ZB272 ZB322

ZB332 ZB352 ZB372 ZB382 ZC032 ZC042 ZC062 ZC082 ZC212 ZC332

ZC352 ZC372 ZC552

4H045 AA10 AA11 AA20 AA30 BA10 BA41 CA40 DA50 DA76 DA86

EA20 EA50 FA72 FA74 GA26

專利名称(译)	跨膜蛋白		
公开(公告)号	JP2004530409A	公开(公告)日	2004-10-07
申请号	JP2002537772	申请日	2001-10-26
[标]申请(专利权)人(译)	洞察Genomics公司		
申请(专利权)人(译)	洞察基因组公司		
[标]发明人	ワレンブリジットエイ スーユーミング ユエヘンリー バトラサジープ バーフォードニール ガンディーアミーナアール チョーラナリンダーケイ アービズチャンドラ タングワイトム リュデュングアイナエム ダガンブレンダンエム ボーグンマライアアール リーアーンステイーンエイ カーンファラエイ ニュエンダニエルビー アジムザイヤルダ ヤオモニークジー ラルプリーティジー サンガベルカピサ ランクマールジャヤラクシミ トランバオ ディングリー オウヤングジャンス		
发明人	ワレン、ブリジット、エイ スー、ユーミング ユエ、ヘンリー バトラ、サジープ バーフォード、ニール ガンディー、アミーナ・アール チョーラ、ナリンダー・ケイ アービズ、チャンドラ タング、ワイ・トム リュ、デュング・アイナ・エム ダガン、ブレンダン・エム ボーグン、マライア・アール リー、アーンステイーン・エイ カーン、ファラ・エイ ニュエン、ダニエル・ビー アジムザイ、ヤルダ ヤオ、モニーク・ジー ラル、プリーティ・ジー サンガベル、カピサ ランクマール、ジャヤラクシミ トラン、バオ		

ディング、リー
オウ-ヤング、ジャニス

IPC分类号

G01N33/50 A61K38/00 A61K38/45 A61K45/00 A61P1/00 A61P1/04 A61P1/06 A61P1/08 A61P1/10
A61P1/12 A61P1/14 A61P1/16 A61P1/18 A61P3/00 A61P3/04 A61P3/06 A61P3/10 A61P3/12 A61P5
/00 A61P5/06 A61P5/08 A61P5/14 A61P5/40 A61P7/02 A61P7/04 A61P7/06 A61P9/00 A61P9/04
A61P9/08 A61P9/10 A61P9/12 A61P11/00 A61P11/06 A61P13/08 A61P13/12 A61P15/00 A61P15/08
A61P15/10 A61P17/06 A61P17/12 A61P19/06 A61P21/00 A61P21/04 A61P25/00 A61P25/02 A61P25
/08 A61P25/14 A61P25/16 A61P25/18 A61P25/20 A61P25/22 A61P25/24 A61P25/28 A61P27/02
A61P27/06 A61P27/12 A61P27/16 A61P29/00 A61P31/04 A61P31/06 A61P31/10 A61P31/12 A61P31
/18 A61P33/00 A61P35/00 A61P35/02 A61P37/02 A61P37/08 A61P39/02 A61P39/04 C07K1/22
C07K14/705 C07K16/18 C07K16/28 C07K16/46 C12M1/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10
C12N15/02 C12N15/09 C12N15/12 C12P21/02 C12P21/08 C12Q1/02 C12Q1/68 G01N33/15 G01N33
/53 G01N37/00

CPC分类号

A61P1/00 A61P1/04 A61P1/06 A61P1/08 A61P1/10 A61P1/12 A61P1/14 A61P1/16 A61P1/18 A61P3
/00 A61P3/04 A61P3/06 A61P3/10 A61P3/12 A61P5/00 A61P5/06 A61P5/08 A61P5/14 A61P5/40
A61P7/02 A61P7/04 A61P7/06 A61P9/00 A61P9/04 A61P9/08 A61P9/10 A61P9/12 A61P11/00 A61P11
/06 A61P13/08 A61P13/12 A61P15/00 A61P15/08 A61P15/10 A61P17/06 A61P17/12 A61P19/06
A61P21/00 A61P21/04 A61P25/00 A61P25/02 A61P25/08 A61P25/14 A61P25/16 A61P25/18 A61P25
/20 A61P25/22 A61P25/24 A61P25/28 A61P27/02 A61P27/06 A61P27/12 A61P27/16 A61P29/00
A61P31/04 A61P31/06 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37
/02 A61P37/08 A61P39/02 A61P39/04 C07K14/705

FI分类号

C12N15/00.ZNA.A A61K45/00 A61P1/00 A61P1/04 A61P1/06 A61P1/08 A61P1/10 A61P1/12 A61P1
/14 A61P1/16 A61P1/16.101 A61P1/16.105 A61P1/18 A61P3/00 A61P3/04 A61P3/06 A61P3/10 A61P3
/12 A61P5/00 A61P5/06 A61P5/08 A61P5/14 A61P5/40 A61P7/02 A61P7/04 A61P7/06 A61P9/00
A61P9/04 A61P9/08 A61P9/10 A61P9/10.103 A61P9/12 A61P11/00 A61P11/06 A61P13/08 A61P13/12
A61P15/00 A61P15/08 A61P15/10 A61P17/06 A61P17/12 A61P19/06 A61P21/00 A61P21/04 A61P25
/00 A61P25/02 A61P25/02.103 A61P25/08 A61P25/14 A61P25/16 A61P25/18 A61P25/20 A61P25/22
A61P25/24 A61P25/28 A61P27/02 A61P27/06 A61P27/12 A61P27/16 A61P29/00.101 A61P31/04
A61P31/06 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37/02 A61P37
/08 A61P39/02 A61P39/04 C07K1/22 C07K14/705 C07K16/18 C07K16/28 C07K16/46 C12M1/00.A
C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/02 C12Q1/68.A G01N33/15.Z
G01N33/50.Z G01N33/53.D G01N33/53.M G01N37/00.102 C12N15/00.C C12N15/00.F C12N5/00.A
A61K37/02 A61K37/52

F-TERM分类号

2G045/AA25 2G045/AA40 2G045/BB03 2G045/BB20 2G045/CB01 2G045/CB21 2G045/DA12 2G045
/DA13 2G045/DA14 2G045/DA36 2G045/DA37 2G045/DA77 2G045/FB02 2G045/FB03 4B024/AA01
4B024/AA11 4B024/BA43 4B024/BA63 4B024/CA04 4B024/CA05 4B024/CA06 4B024/CA07 4B024
/CA09 4B024/CA10 4B024/CA12 4B024/DA02 4B024/DA06 4B024/EA02 4B024/EA04 4B024/FA02
4B024/FA10 4B024/GA05 4B024/GA11 4B024/GA18 4B024/GA19 4B024/HA03 4B024/HA08 4B024
/HA12 4B024/HA14 4B029/AA07 4B029/AA21 4B029/AA23 4B029/BB20 4B029/CC03 4B029/CC08
4B029/FA12 4B063/QA01 4B063/QA18 4B063/QQ20 4B063/QQ42 4B063/QR08 4B063/QR14 4B063
/QR32 4B063/QR41 4B063/QR42 4B063/QR48 4B063/QR55 4B063/QR57 4B063/QR59 4B063/QR62
4B063/QR66 4B063/QR69 4B063/QR77 4B063/QR82 4B063/QS12 4B063/QS24 4B063/QS25 4B063
/QS28 4B063/QS34 4B063/QS39 4B063/QX01 4B063/QX07 4B064/AG20 4B064/AG27 4B064/CA02
4B064/CA10 4B064/CA19 4B064/CA20 4B064/CC24 4B064/CE12 4B064/CE13 4B064/DA01 4B064
/DA13 4B065/AA26X 4B065/AA90X 4B065/AA93Y 4B065/AB01 4B065/AC14 4B065/BA02 4B065
/BA25 4B065/CA24 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA06 4C084/AA07 4C084/AA17
4C084/BA01 4C084/BA08 4C084/BA22 4C084/CA53 4C084/DC25 4C084/DC50 4C084/MA13 4C084
/MA24 4C084/MA52 4C084/MA56 4C084/MA57 4C084/MA59 4C084/MA60 4C084/MA63 4C084/MA66
4C084/ZA022 4C084/ZA052 4C084/ZA062 4C084/ZA082 4C084/ZA122 4C084/ZA162 4C084/ZA182
4C084/ZA222 4C084/ZA242 4C084/ZA332 4C084/ZA342 4C084/ZA362 4C084/ZA392 4C084/ZA402
4C084/ZA422 4C084/ZA452 4C084/ZA522 4C084/ZA542 4C084/ZA552 4C084/ZA592 4C084/ZA662
4C084/ZA682 4C084/ZA692 4C084/ZA702 4C084/ZA712 4C084/ZA722 4C084/ZA752 4C084/ZA762
4C084/ZA772 4C084/ZA812 4C084/ZA892 4C084/ZA942 4C084/ZA962 4C084/ZA972 4C084/ZB072
4C084/ZB082 4C084/ZB132 4C084/ZB152 4C084/ZB262 4C084/ZB272 4C084/ZB322 4C084/ZB332

4C084/ZB352 4C084/ZB372 4C084/ZB382 4C084/ZC032 4C084/ZC042 4C084/ZC062 4C084/ZC082
4C084/ZC212 4C084/ZC332 4C084/ZC352 4C084/ZC372 4C084/ZC552 4H045/AA10 4H045/AA11
4H045/AA20 4H045/AA30 4H045/BA10 4H045/BA41 4H045/CA40 4H045/DA50 4H045/DA76 4H045
/DA86 4H045/EA20 4H045/EA50 4H045/FA72 4H045/FA74 4H045/GA26

優先権
60/244017 2000-10-27 US
60/252855 2000-11-22 US
60/251825 2000-12-07 US
60/255085 2000-12-12 US

外部リンク Espacenet

摘要(译)

本发明提供了鉴定和编码TMP的人跨膜蛋白 (TMP) 和多核苷酸。本发明还提供了表达载体，宿主细胞，抗体，激动剂和拮抗剂。本发明还提供了诊断，治疗或预防与TMP异常表达相关的病症的方法。

特許公開番号 (P2004-53049A) (43) 公表日 平成16年10月7日 (2004.10.7)			
(51) Int. Cl. ⁷	F I	テーマコード (参考)	
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A	2 G 0 4 5
A 6 1 K 38/00	A 6 1 K 45/00		4 B 0 2 4
A 6 1 K 38/45	A 6 1 P 1/00		4 B 0 2 9
A 6 1 K 45/00	A 6 1 P 1/04		4 B 0 6 3
A 6 1 P 1/00	A 6 1 P 1/06		4 B 0 6 4
		審査請求 未請求 予備審査請求 有	(全 279 頁) 最終頁に続く
(21) 出願番号	特願2002-537772 (P2002-537772)	(71) 出願人	301005050
(86) (22) 出願日	平成13年10月26日 (2001.10.26)		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15年4月24日 (2003.4.24)		アメリカ合衆国カリフォルニア州9430 4・パロアルト・ボータードドライブ 31 60
(86) 国際出願番号	PCT/US2001/049670	(74) 代理人	100089266
(87) 国際公開番号	W02002/034783		弁理士 大島 陽一
(87) 国際公開日	平成14年5月2日 (2002.5.2)	(72) 発明者	ワレン、ブリジット、エイ
(31) 優先権主張番号	60/244,017		アメリカ合衆国カリフォルニア州9501 4・クーパービルティノー・#2・パークウッ ドドライブ 10130
(32) 優先日	平成12年10月27日 (2000.10.27)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/252,855		
(32) 優先日	平成12年11月22日 (2000.11.22)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/251,825		
(32) 優先日	平成12年12月7日 (2000.12.7)		
(33) 優先権主張国	米国 (US)		
最終頁に続く			