

(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2004-531213

(P2004-531213A)

(43) 公表日 平成16年10月14日(2004. 10. 14)

(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 39/395	N 4 B O 2 4
A 6 1 K 39/395	A 6 1 K 45/00	4 B O 2 9
A 6 1 K 45/00	C O 7 K 1/22	4 B O 6 3
C O 7 K 1/22	C O 7 K 14/705	4 B O 6 4
	審査請求 未請求 予備審査請求 有	(全 302 頁) 最終頁に続く

(21) 出願番号	特願2002-543549 (P2002-543549)	(71) 出願人	301005050
(86) (22) 出願日	平成13年10月25日 (2001. 10. 25)		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/US2001/046055		
(87) 国際公開番号	W02002/040541	(74) 代理人	100089266
(87) 国際公開日	平成14年5月23日 (2002. 5. 23)		弁理士 大島 陽一
(31) 優先権主張番号	60/243, 989	(72) 発明者	タング、ワイ・トム
(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95118・サンノゼ・ランウィックコート 4230
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/245, 904		
(32) 優先日	平成12年11月3日 (2000. 11. 3)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/247, 673		
(32) 優先日	平成12年11月9日 (2000. 11. 9)		
(33) 優先権主張国	米国 (US)		
			最終頁に続く

(54) 【発明の名称】 輸送体及びイオンチャネル

(57) 【要約】

本発明はヒト輸送体およびイオンチャネル (T R I C H)、並びにT R I C H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、T R I C Hの異常発現に関連する疾患を診断、治療または予防する方法をも提供する。

【特許請求の範囲】

【請求項 1】

以下の (a) 乃至 (d) 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: 1-20 (配列番号 1 乃至 20) 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: 1-20 からなる群から選択したアミノ酸配列と少なくとも 90% が同一であるような天然のアミノ酸配列を含むポリペプチド

(c) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片

(d) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片 10

【請求項 2】

SEQ ID NO: 1-20 からなる群から選択されたアミノ酸配列を含むことを特徴とする請求項 1 に記載の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 4】

請求項 2 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 5】

SEQ ID NO: 21-40 (配列番号 21 乃至 40) からなる群から選択したポリヌクレオチド配列を含む、請求項 4 に記載の単離されたポリヌクレオチド。 20

【請求項 6】

請求項 3 に記載のポリヌクレオチドに機能的に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されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養するステップと、

(b) そのように発現した前記ポリペプチドを回収する過程とから成る。

【請求項 10】

前記ポリペプチドが、SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を含むことを特徴とする請求項 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。

【請求項 12】 40

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: 21-40 からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド

(b) SEQ ID NO: 21-40 からなる群から選択したポリヌクレオチド配列と少なくとも 90% が同一であるような天然のポリヌクレオチド配列を含むポリヌクレオチド

(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド

(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド

(e) (a) ~ (d) の RNA 等価物

【請求項 13】

請求項 12 に記載のポリヌクレオチドの少なくとも 60 の連続したヌクレオチドを含む単 50

離されたポリヌクレオチド。

【請求項 14】

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

(a) 前記サンプル中の前記標的ポリヌクレオチドに相補的な配列を有する少なくとも 20 の連続したヌクレオチドを含むプローブを用いて、そのプローブと前記標的ポリヌクレオチド、またはその断片との間でハイブリダイゼーション複合体が形成される条件下で前記サンプルをハイブリダイズする過程と、

(b) 前記ハイブリダイゼーション複合体の存在・不存在を検出し、前記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方法を含む。

10

【請求項 15】

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

【請求項 16】

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

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

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

20

【請求項 17】

請求項 1 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: 1-20 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】

機能的な TRICH の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 17 の組成物を投与することを含むことを特徴とする治療方法。

30

【請求項 20】

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

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝すステップと、

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

【請求項 21】

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

40

【請求項 22】

機能的な TRICH の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 21 の組成物を投与することを含むことを特徴とする治療方法。

【請求項 23】

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

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝すステップと、

(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。

50

【請求項 24】

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

【請求項 25】

機能的な T R I C H の過剰な発現に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 24 の組成物を投与することを含むことを特徴とする治療方法。

【請求項 26】

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

10

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

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

【請求項 27】

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

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

20

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

(c) 試験化合物の存在下での請求項 1 に記載のポリペプチドの活性を、試験化合物の不存在下での請求項 1 に記載のポリペプチドの活性と比較する過程とを含み、試験化合物の存在下での請求項 1 に記載のポリペプチドの活性の変化が、請求項 1 に記載のポリペプチドの活性を調節する化合物を標示することを特徴とする方法。

【請求項 28】

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

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

30

(b) 前記標的ポリヌクレオチドの変異発現を検出する過程と、

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

【請求項 29】

試験化合物の毒性を算定する方法であって、

(a) 核酸を含む生物学的サンプルを前記試験化合物で処理する過程と、

(b) 処理した前記生体サンプルの核酸と、請求項 12 のポリヌクレオチドの少なくとも 20 の連続するヌクレオチドを含むプローブをハイブリダイズさせるステップであって、このハイブリダイゼーションゼーションが、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われ、前記標的ポリヌクレオチドが、請求項 12 のポリヌクレオチドのポリヌクレオチド配列またはその断片を含むポリヌクレオチドである、前記ステップと、

40

(c) ハイブリダイゼーション複合体の収量を定量するステップと、

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

【請求項 30】

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

50

(a) 前記抗体を前記ポリペプチドと混合し、抗体とポリペプチドとの複合体が形成されるのに適した条件下で、前記生物学的サンプルを請求項 11 に記載の抗体と結合する過程と、

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

【請求項 31】

前記抗体が、

(a) キメラ抗体

(b) 単鎖抗体

(c) Fab 断片

(d) F(ab')₂ 断片

(e) ヒト化抗体 のいずれかであることを特徴とする請求項 11 に記載の抗体。

10

【請求項 32】

請求項 11 に記載の抗体と、許容できる賦形剤とを含む組成物。

【請求項 33】

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

【請求項 34】

前記抗体が標識されることを特徴とする請求項 32 に記載の組成物。

【請求項 35】

20

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

【請求項 36】

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

(a) 抗体反応を誘発する条件下で、SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列からなるポリペプチドまたはその免疫原性断片を用いて動物を免疫化する過程と、

(b) 前記動物から抗体を単離する過程と、

(c) 前記単離された抗体をポリペプチドでスクリーニングし、それによって、SEQ ID NO: 1-20 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ポリクローナル抗体を同定する過程とを含むことを特徴とする方法。

30

【請求項 37】

請求項 36 に記載の方法で産出したポリクローナル抗体。

【請求項 38】

請求項 37 に記載のポリクローナル抗体及び適切なキャリアーを含む組成物。

【請求項 39】

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

(a) 抗体反応を誘発する条件下で、SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列からなるポリペプチドまたはその免疫原性断片を用いて動物を免疫化する過程と、

40

(b) 前記動物から抗体産出細胞を単離する過程と、

(c) 不死化の細胞を用いて前記抗体産出細胞を融合して、モノクローナル抗体を産出するハイブリドーマ細胞を形成する過程と、

(d) 前記ハイブリドーマ細胞を培養する過程と、

(e) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を含むポリペプチドに特異結合するようなモノクローナル抗体を前記培養物から単離する過程とを含むことを特徴とする方法。

【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及び適切なキャリアーを含む組成物。

【請求項 42】

F a b 発現ライブラリのスクリーニングにより前記抗体を産出することを特徴とする請求項 11 に記載の抗体。

【請求項 43】

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

【請求項 44】

S E Q I D N O : 1 - 20 からなる群から選択したアミノ酸配列を含むポリペプチドを検出する方法であって、

10

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

(b) 特異結合を検出する過程とを含み、該特異結合が、S E Q I D N O : 1 - 20 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がサンプル中に存在することを示唆することを特徴とする方法。

【請求項 45】

S E Q I D N O : 1 - 20 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) 前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項 11 に記載の抗体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 - 20 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 つが請求項 13 に記載のポリヌクレオチドであることを特徴とするマイクロアレイ。

【請求項 47】

ポリヌクレオチドを含むサンプルの発現プロファイルを作製する方法であって、

a) サンプルのポリヌクレオチドを標識する過程と、

b) ハイブリダイゼーション複合体形成に適した条件下で請求項 46 に記載のマイクロアレイのエレメントをサンプルの標識されたポリヌクレオチドと接触させる過程と、

30

c) サンプルのポリヌクレオチドの発現を定量化する過程を含むことを特徴とする方法。

【請求項 48】

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

【請求項 49】

40

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

【請求項 50】

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

【請求項 51】

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

50

【請求項 5 2】

請求項 4 8 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 5 3】

請求項 4 8 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を含むヌクレオチド分子にハイブリダイズした前記の標的ポリヌクレオチドを有することを特徴とするアレイ。

【請求項 5 4】

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

【請求項 5 5】

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

10

【請求項 5 6】

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

【請求項 5 7】

SEQ ID NO : 2 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 5 8】

SEQ ID NO : 3 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

20

【請求項 5 9】

SEQ ID NO : 4 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 6 0】

SEQ ID NO : 5 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 6 1】

SEQ ID NO : 6 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 6 2】

SEQ ID NO : 7 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 6 3】

SEQ ID NO : 8 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

30

【請求項 6 4】

SEQ ID NO : 9 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 6 5】

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

【請求項 6 6】

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

【請求項 6 7】

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

【請求項 6 8】

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

40

【請求項 6 9】

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

【請求項 7 0】

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

【請求項 7 1】

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

【請求項 7 2】

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

【請求項 7 3】

50

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

【請求項 74】

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

【請求項 75】

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

【請求項 76】

SEQ ID NO: 21 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 77】

SEQ ID NO: 22 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。 10

【請求項 78】

SEQ ID NO: 23 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 79】

SEQ ID NO: 24 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 80】

SEQ ID NO: 25 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。 20

【請求項 81】

SEQ ID NO: 26 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 82】

SEQ ID NO: 27 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 83】

SEQ ID NO: 28 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 84】

SEQ ID NO: 29 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。 30

【請求項 85】

SEQ ID NO: 30 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 86】

SEQ ID NO: 31 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 87】

SEQ ID NO: 32 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。 40

【請求項 88】

SEQ ID NO: 33 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 89】

SEQ ID NO: 34 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。

【請求項 90】

SEQ ID NO: 35 のポリヌクレオチド配列を含む請求項 12 に記載のポリヌクレオチド。 50

【請求項 9 1】

SEQ ID NO: 36 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

【請求項 9 2】

SEQ ID NO: 37 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

【請求項 9 3】

SEQ ID NO: 38 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

【請求項 9 4】

SEQ ID NO: 39 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

【請求項 9 5】

SEQ ID NO: 40 のポリヌクレオチド配列を含む請求項 1 2 に記載のポリヌクレオチド。

【発明の詳細な説明】

【0001】

(技術分野)

本発明は、輸送体及びイオンチャネルの核酸配列及びアミノ酸配列に関する。本発明はまた、これらの配列を利用した輸送障害、神経の疾患、筋疾患、及び免疫疾患および細胞の異常増殖の診断・治療・予防に関する。本発明はさらに、輸送体及びイオンチャネルの核酸配列及びアミノ酸配列の発現における外来性化合物の効果についての評価に関する。

【0002】

(発明の背景)

真核細胞は、殆どの極性分子に高度に非浸透性であるような疎水性脂質二重層膜に囲まれ、この膜によって機能的に異なるオルガネラに細区画される。細胞及びオルガネラは、必須栄養素及び、 K^+ 、 NH_4^+ 、 Pi 、 SO_4^{2-} を含む金属イオン、糖、ビタミン及び種々の廃棄物を移出入するための輸送タンパク質を必要とする。輸送タンパク質は、抗生物質抵抗性、毒物分泌、イオンバランス、シナプス神経伝達、腎機能、腸管吸収、腫瘍成長及びその他の多様な細胞機能においても役割を果たす (Griffith, J. および C. Sansom (1998) The Transporter Facts Book, Academic Press, San Diego CA, 3-29 ページ)。輸送は、受動的な濃度依存メカニズムによって起こるか、或いは ATP 加水分解またはイオン勾配などのエネルギー源に結びつけることができる。輸送時に機能するタンパク質には、キャリアタンパク質及びチャネルタンパク質がある。キャリアタンパク質は、特定の溶質に結合して、あるコンフォメーションの変化を起こし、その変化によって結合した溶質が膜を通過できるようにする。チャネルタンパク質は、特定の溶質が電気化学溶質勾配に従って膜を通過して拡散することができるような疎水性ポアを形成する。

【0003】

膜の一方の側から他方の側へ単一溶質を輸送するキャリアタンパク質をユニポーターと呼ぶ。対照的に、共役輸送体は、1つの溶質の移動を第2溶質の移動に連結させ、その移動は同時または順次に起こるかのいずれかであり、移動の方向は同一方向 (シンポート) または逆方向 (アンチポート) のいずれかである。例えば、腸管及び腎臓の上皮は、原形質膜に存在するナトリウム勾配によって駆動される種々のシンポーター系を含む。ナトリウムは、電気化学勾配に従って細胞内に移動し、その移動と共に溶質を細胞内に運ぶ。溶質を取り込むための駆動力を供給するナトリウム勾配は、遍在性の Na^+/K^+ ATP 分解酵素系によって維持される。ナトリウム共役輸送体には、哺乳動物グルコース輸送体 (SGLT1)、ヨウ化物輸送体 (NIS) 及び多ビタミン輸送体 (SMVT) がある。3つの輸送体はすべて、12個の推定上の膜貫通セグメント、細胞外グリコシル化部位、

10

20

30

40

50

並びに細胞質内側にあるN末端及びC末端を有する。NISは、放射性ヨウ素甲状腺イメージング技術及び甲状腺への放射性同位元素の特定のターゲティングのための分子基準であるので、種々の甲状腺病理の評価、診断及び治療において重大な役割を果たす (Levy, O. 他 (1997) Proc. Natl. Acad. Sci. USA 94 : 5568-5573)。SMTは、腸管粘膜、腎臓及び胎盤において発現され、ビオチンやパントテン酸などの水溶性ビタミンの輸送に関与していると考えられる (Prasad, P. D. 他 (1998) J. Biol. Chem. 273 : 7501-7506)。

【0004】

MFS (major facilitator superfamily) は、輸送体の最大ファミリーの1つであり、ユニポーター-シンポーター-アンチポーターファミリーとも呼ばれる。MFS輸送体は、イオン勾配に応じて小溶質を輸送する単一のポリペプチドキャリアである。MFSのメンバーは、すべてのクラスの生物に見られ、糖、オリゴ糖、リン酸、硝酸、ヌクレオチド、モノカルボン酸及び薬剤のための輸送体を含む。真核生物に見られるMFS輸送体はすべて12個の膜貫通セグメントを含む構成となっている (Pao, S. S. 他 (1998) Microbiol. Molec. Biol. Rev. 62 : 134)。MFS輸送体の最大のファミリーは糖輸送体ファミリーであり、ヒトに見られるような、グルコース及びその他の六炭糖の輸送に必要な7つのグルコース輸送体 (GLUT1~GLUT7) を含む。これらのグルコース輸送タンパク質は、独自の組織分布及び生理学的機能を有する。GLUT1は、基底板でのグルコース要求を有する多くの細胞型に提供して、上皮及び内皮の障壁を通してグルコースを輸送し、GLUT2は、グルコースの取り込みまたは肝臓からの流出を促進し、GLUT3は、ニューロンへのグルコースの供給を調整し、GLUT4は、インシュリンによって調整されるグルコース分解に関与し、GLUT5は、骨格筋へのフルクトースの取り込みを調節する。グルコース輸送体の欠陥は、最近同定された神経症候群であって、糖原貯蔵症、ファンコニービッケル症候群、及び非インシュリン依存糖尿病のみならず、乳児てんかん及び発達障害の原因となるような神経症候群に関与している (Mueckler, M. (1994) Eur. J. Biochem. 219 : 713-725、Longo, N. および L. J. Elsas (1998) Adv. Pediatr. 45 : 293-313)。

【0005】

モノカルボン酸アニオン輸送体は、L-乳酸、ピルビン酸及びケトン体アセテート、アセトアセテート及びβヒドロキシ酪酸を含む広範な基質特異性を有するプロトン共役シンポーターである。これまでに少なくとも7つのアイソフォームが同定されている。アイソフォームは、TM6とTM7との間に大きな細胞内ループを有する12個の膜貫通 (TM) らせん状ドメインを有し、解糖中に乳酸と共に化学量論的に生成されるプロトン除去することにより、細胞内pHを維持する際に重要な役割を果たすと推定されている。最も良く特徴付けられたH⁺ モノカルボン酸輸送体は、L-乳酸及び広範囲の他の脂肪族モノカルボン酸を輸送する赤血球膜の輸送体である。その他の細胞は、異なる基質及び抑制因子選択性を有するH⁺ 共役モノカルボン酸輸送体を含む。特に心筋及び腫瘍細胞は、或る基質に対するKm値 (D-乳酸よりL-乳酸を選択する立体化学的選択性を含む) 及び抑制因子に対する感受性が異なる輸送体を有する。腸管及び肝臓上皮の管腔表面にNa⁺ モノカルボン酸共輸送体があり、それによってこれらの組織における乳酸、ピルビン酸及びケトン体の取り込みが可能になる。更に、肝臓、腸及び肺を含む器官に、有機カチオン及び有機アニオンのための特異的且つ選択的な輸送体がある。有機アニオン輸送体は、電子誘引側基を有する疎水性、荷電分子に選択的である。アンモニウム輸送体などの有機カチオン輸送体は、種々の薬剤及び内因性代謝の分泌を媒介し、細胞内pHの維持に寄与する (Poole, R. C. および A. P. Halestrap (1993) Am. J. Physiol. 264 : C761-C782、Price, N. T. 他 (1998) Biochem. J. 329 : 321-328、Martinel

le, K. および I. Haggstrom (1993) J. Biotechnol. 30 : 339-350)。

【0006】

A T P 結合カセット (A B C) 輸送体は、イオン、糖、アミノ酸、ペプチド、リン脂質のような小分子からリポペプチド、大タンパク質、及び複合体疎水性薬剤までに及ぶ基質を輸送するような、膜タンパク質のスーパーファミリーのメンバーである。A B C 輸送体は、4つのモジュール即ち2つのヌクレオチド結合ドメイン (N B D) 及び2つの膜貫通ドメイン (M S D) を含む。N B D は、A T P を加水分解して輸送に必要なエネルギーを供給し、各 M S D は、推定上の6つの膜貫通セグメントを含む。これら4つのモジュールは、単一の遺伝子または別々の遺伝子によりコードし得る。単一の遺伝子によりコードし得るのは、嚢胞性繊維症膜通過伝導制御因子 (C F T R) に対する場合などである。別々の遺伝子にコードされる場合には、各遺伝子産物には1つのN B D 及び1つのM S D が含まれる。これらの「半分子」は、小胞体での主要組織適合複合体 (M H C) ペプチド輸送系である T a p 1 及び T a p 2 などのホモ及びヘテロ二量体を形成する。幾つかの遺伝病は、A B C 輸送体の欠陥に起因する。疾患及びそれに対応するタンパク質の例としては、嚢胞性繊維症 (C F T R、イオンチャネル)、副腎白質ジストロフィー (副腎白質ジストロフィータンパク質、A L D P)、ツェルヴェーガー症候群 (ペルオキシソームの膜タンパク質-70、P M P 70) 及び高インシュリン性低血糖症 (スルホニル尿素受容体、S U R) がある。別のA B C 輸送体である多剤耐性 (M D R) タンパク質がヒトの癌細胞において過剰発現すると、化学療法に用いられる種々の細胞毒性薬剤に対して細胞が耐性を有する (T a g l i c h t, D. および S. M i c h a e l i s (1998) M e t h. E n z y m o l. 292 : 130-162)。

10

20

【0007】

鉄、亜鉛、銅、コバルト、マンガン、モリブデン、セレン、ニッケル、クロムなどの多数の金属イオンは、多数の酵素に対する補助因子として重要である。例えば、銅は、スーパーオキシドジスムターゼ、フェロキシダーゼ (セルロプラスミン) 及びリシルオキシダーゼなどのオキシドレダクターゼにおいて補助因子として作用することにより、ヘモグロビン合成、結合組織代謝及び骨格の発達に関与する。銅及びその他の金属イオンは、食餌中に与えなければならないものであり、消化管で輸送体に吸収される。血漿タンパク質は、金属イオンを肝臓及びその他の標的器官へ輸送し、ここで特異輸送体が必要に応じてイオンを細胞及び細胞のオルガネラに移動させる。金属イオンでの代謝の不均衡は、多数の疾患状態に関連してきた (D a n k s, D. M. (1986) J. M e d. G e n e t. 23 : 99-106)。

30

【0008】

原形質膜を通過する脂肪酸の輸送は、高容量で低親和性プロセスである拡散により起こる。しかしながら、通常の生理学的条件下では、脂肪酸輸送の内の有意の割合が高親和性、低容量のタンパク質媒介輸送プロセスによって起こっているように見える。脂肪酸輸送タンパク質 (F A T P) は、4つの膜貫通セグメントを有する膜内在性タンパク質であり、筋肉、心臓及び脂肪などの、高レベルの原形質膜脂肪酸流入を示す組織において発現される。F A T P の発現は、脂肪変換中に 3 T 3 - L 1 細胞において上方制御され、C O S 7 線維芽細胞における発現は、長鎖脂肪酸の取り込みを増加させる (H u i, T. Y. 他 (1998) J. B i o l. C h e m. 273 : 27420-27429)。

40

【0009】

リポカリンスーパーファミリーは、小さな疎水性分子に結合して輸送する細胞外リガンド結合タンパク質として機能する40以上のタンパク質のグループであり、系統樹的に保存されている。このファミリーのメンバーはレチノイド、臭気物質、発色団、フェロモン、アレルゲンおよびステロールのキャリアーとして機能し、また栄養輸送、細胞成長調節、免疫応答、およびプロスタグランジン合成を含むさまざまなプロセスにおいて機能する。これらのタンパク質のあるサブセットは多機能で、生合成酵素または特定の酵素インヒビターとしての役割を果たしている可能性がある。(T a n a k a, T. 他 (1997)

50

J. Biol. Chem. 272:15789-15795; および van 't Hoff, W. 他 (1997) J. Biol. Chem. 272:1837-1841.)。

【0010】

リポカリンのファミリーのメンバーは、異常に低い全体的な配列の保存レベルを示す。ペアワイズの配列の同一性はしばしば20%以下に落ちる。ファミリーメンバーの間の配列の類似性は、ジスルフィド結合を形成する保存性のシステインおよび標的細胞の認識部位として機能して並列クラスターを形成する3つのモチーフに限定されている。リポカリンは、8ストランドの逆平行ベータシートを共有し、このベータシートはそれ自身の上に折りたたんで、連続的な水素結合を持つベータバレルを形成する。バレルによって形成されるポケットは内部リガンド結合部位として機能する。7つのループ(L1~L7)が短いベータヘアピンを形成するが、L1だけは例外で、内部リガンド結合部位を部分的に閉じる蓋状構造を形成する大きなオメガグループになっている(Flowe (1996) Biochem. J. 318:1-14)。

【0011】

リポカリンは重要な輸送分子である。各リポカリンは、特定のリガンドと結合して、そのリガンドを生物体の中の適切な標的部位に送達する。最もよく研究されたリポカリンの一つであるレチノール結合タンパク質(RBP)は、肝臓内の貯蔵場所からレチノールを標的組織に輸送する。アポリポプロテインD(アポD)は高密度リポプロテイン(HDL)と低密度リポプロテイン(LDL)のコンポーネントであり、全身のコレステロールの標的を特定した収集と送達の役割を果たす。リポカリンは細胞の調節過程にも関与する。アポDは巨視的嚢胞性疾患液タンパク質(GCDFP)24と同じもので、乳房の嚢胞液内に高レベルで発現するプロゲステロン/プレグネノロン結合タンパク質である。ある種のヒトの乳癌細胞株におけるアポDの分泌は、細胞増殖の低減および細胞がより分化する表現型に進行することを伴う。同様にアポDともう一つのリポカリンである α_1 酸糖タンパク質(AGP)は神経細胞の再生に関与している。AGPは抗炎症および免疫抑制活性にも関与している。AGPは陽性急性期タンパク質(APP)の一つである。AGPの循環レベルはストレスおよび炎症刺激への応答として増加する。AGPは炎症部位に蓄積し、血小板と好中球の活性化を阻害し、食細胞活動を阻害する。AGPの免疫調節特性はグリコシル化に起因している。AGPは40%が糖質であるため、異常に酸性で可溶性である。AGPのグリコシル化パターンは急性期反応中に変化し、脱グリコシル化されたAGPは免疫抑制活性を持っていない(前出のFlowe (1994) FEBS Lett. 354:7-11; Flowe (1996))。

【0012】

リポカリンスーパーファミリーはいくつかの動物アレルゲンも含んでいて、これにはマウスの主要尿タンパク質(mMUP)、ラットの α -2ミクログロブリン(rA2U)、ウシの β ラクトグロブリン(β lg)、ゴキブリのアレルゲン(Bla g4)、ウシの鱗屑アレルゲン(Bos d2)、およびEquus caballus アレルゲン1(Equ c1)と名づけられた主要ウマアレルゲンが含まれる。Equ c1は、慢性的にウマのアレルゲンに曝される患者の抗ウマIgE抗体反応の原因の約80%を占める強力なアレルゲンである。リポカリンはIgE反応を誘発する能力を持った共通の構造を持っているように見える(Gregoire, C. 他, (1996) J. Biol. Chem. 271:32951-32959)。

【0013】

lipocalinはさまざまな疾患状態において診断および予測マーカーとして使用されている。AGPの血漿レベルは妊娠期間中に監視され、癌の化学療法、腎機能障害、心筋梗塞、関節炎および多発性硬化症を含む状態の診断と予測の際にも監視される。RBPは臨床的に腎臓における尿細管再吸収のマーカーとして使用され、アポDは巨視的嚢胞乳房疾患のマーカーである(前出のFlowe (1996))。さらに、リポカリン動物アレルゲンの使用は、ウマ(前出のGregoire)、ブタ、ゴキブリ、マウスおよ

びラットへのアレルギー反応の診断に役立つ可能性がある。

【0014】

ミトコンドリアキャリアタンパク質は、サイトゾル及びミトコンドリア基質の間でイオン及び荷電代謝産物を輸送する膜貫通タンパク質である。例として、ADP、ATPキャリアタンパク質、2-オキソグルタル酸／リンゴ酸キャリア、リン酸キャリアタンパク質、ピルビン酸キャリア、ジカルボン酸キャリア（リンゴ酸、コハク酸、フマル酸及びリン酸を輸送する）、トリカルボン酸キャリア（クエン酸及びリンゴ酸を輸送する）、及びグレース病キャリアタンパク質、即ち甲状腺機能亢進症の原因となる自己免疫異常である活性グレース病を患った患者のIgGにより認識されたタンパク質がある。このファミリーのタンパク質は、約100アミノ酸ドメインの3つのタンデムリピートからなり、各タンデムリピートは2つの膜貫通領域を有する（Stryer, L. (1995) Biochemistry, W. H. Freeman and Company, New York NY, p. 551、PROSITE PDOC00189 Mitochondrial energy transfer proteins signature、Online Mendelian Inheritance in Man (OMIM) *275000 Graves Disease）。

【0015】

このクラスの輸送体には、またミトコンドリア脱共役タンパク質も含まれる。この脱共役タンパク質はミトコンドリア内膜でプロトンリークを起こさせ、酸化的リン酸化をATP合成から脱共役させる。結果として、熱の形態でのエネルギー散逸が起こる。ミトコンドリア脱共役タンパク質は、温度調節及び代謝率のモジュレーターとみなされ、肥満症などの代謝病に対する薬剤の潜在的標的として提案されてきた（Ricquier, D. 他 (1999) J. Int. Med. 245 : 637-642）。

【0016】

イオンチャネル

細胞の電位は、原形質膜を通過するイオンの動きを制御することにより起こり、かつ維持される。イオンの動きは、膜内にイオン選択性ポアを形成するようなイオンチャネルを必要とする。イオンチャネルには2つの基本型があり、それはイオン輸送体とイオンチャネルがある。イオン輸送体は、ATP加水分解から得られるエネルギーを利用して、イオンの濃度勾配に逆らってイオンを能動的に輸送する。イオンチャネルは、制限された条件下でイオンの電気化学勾配に従ってイオンの受動的流出を起こさせる。同時に、これらのタイプのイオンチャネルは、1) 神経細胞の軸索に沿った電気インパルス伝導、2) 濃度勾配に逆らった細胞内への分子の輸送、3) 筋収縮の開始、及び4) 内分泌細胞の分泌に用いられるような電気化学勾配を発生させ、維持及び利用する。

【0017】

イオン輸送体

イオン輸送体は、細胞の静止電位を発生させ及び維持する。ATP加水分解から発生するエネルギーを利用して、イオン輸送体は、イオンの濃度勾配に逆らってイオンを輸送する。これらの膜貫通ATP分解酵素は、3つのファミリーに区分される。リン酸化（P）クラスイオン輸送体には、 Na^+/K^+ ATP分解酵素、 Ca^{2+} ATP分解酵素及び H^+ ATP分解酵素があり、リン酸化イベントによって活性化される。Pクラスイオン輸送体は、 Na^+ 及び Ca^{2+} のサイトゾル内濃度が低く、 K^+ のサイトゾル内濃度が高くなるように静止電位での分布を維持するのに関与している。液胞（V）クラスのイオン輸送体には、リソソーム及びゴルジなどの細胞内オルガネラ上の H^+ ポンプがある。オルガネラの内膜内では低いpHが、その機能に必要であり、Vクラスイオン輸送体は、その低pHを発生させるのに関与している。共役因子（F）クラスは、ミトコンドリアに H^+ ポンプを含む。Fクラスイオン輸送体は、プロトン勾配を利用してADP及び無機リン酸（Pi）からATPを生成する。

【0018】

P-ATP分解酵素は、イオン結合で役割を果たし得るような、10個の膜貫通ドメイン

及び幾つかの大きな細胞質内領域を有する100 kD サブユニットの六量体である (Scarborough, G. A. (1999) Curr. Opin. Cell Biol. 11 : 517-522)。V-ATP分解酵素は、2つの機能的ドメインから構成される。V₁ドメインは、ATP加水分解に関与する周辺複合体であり、V₂ドメインは、プロトンの膜通過移動に関与する、膜内在性の複合体である。F-ATP分解酵素は、構造的及び進化的にV-ATP分解酵素に関連する。F-ATP分解酵素F₀ドメインは、12個のコピーのcサブユニットを含んでおり、各サブユニットは高度に疎水性のタンパク質で、2つの膜貫通ドメインから成り、プロトン輸送に不可欠なTM2に1つの埋没されたカルボキシル基を含む。V-ATP分解酵素V₀ドメインは、TM4またはTM3に4つまたは5つの膜貫通ドメイン及び必須カルボキシル基を有する3つの型の相同性cサブユニットを含む。複合体の両型は、活性のpH依存性の制御に関与し得る単一サブユニットも含む (Forgacs, M. (1999) J. Biol. Chem. 274 : 12951-12954)。

10

【0019】

細胞の静止電位は、キャリアタンパク質及びイオンチャネルに関連する多くのプロセスに利用される。キャリアタンパク質は、静止電位を利用して分子を細胞内外に輸送する。多くの細胞へのアミノ酸及びグルコースの輸送は、ナトリウムイオン共輸送（シンポート）にカップルされ、Na⁺の電気化学勾配に従った動きが他の分子を濃度勾配に逆らって輸送させる。同様にして、心筋は、細胞からのCa²⁺の移動を細胞内へのNa⁺の輸送にカップルさせる（アンチポート）。

20

【0020】

イオンチャネル

イオンチャネルは、ポアの開閉を調節することによってイオンの流れを制御する。種々のゲーティング機構によってイオン流速を制御する能力は、ニューロン及び内分泌シグナル伝達、筋収縮、受精並びに、イオン及びpHバランスの調整などの多様なシグナル伝達及び恒常性機能をイオンチャネルに媒介させ得る。イオンチャネルは、開口機能の制御方法に応じて区分される。機械依存性チャネルは、機械的応力に応じてポアを開け、電位依存性チャネル（例えばNa⁺、K⁺、Ca²⁺及びCl⁻チャネル）は膜電位に応じてポアを開け、リガンド依存性チャネル（例えばアセチルコリン依存性、セロトニン依存性、及びグルタミン依存性カチオンチャネル、並びにGABA依存性及びグリシン依存性クロライドチャネル）は、特定のイオン、ヌクレオチドまたは神経伝達物質の存在下でポアを開ける。特定のイオンチャネルの開口特性（即ち開閉に対するその閾値及び持続時間）は、補助チャネルタンパク質及び／またはリン酸化などの翻訳後修飾と関連して時折調節される。

30

【0021】

機械依存性または機械刺激イオンチャネルは、触覚、聴覚及び平衡感覚のトランスデューサとして作用し、細胞容積調整、平滑筋収縮及び心リズム生成においても重要な役割を果たす。機械受容不活性化チャネル（SIC）は、最近ラットの腎臓からクローニングされた。SICチャネルは、細胞膜への圧力またはストレスによって活性化され、Ca²⁺及びNa⁺を共に伝導するようなあるグループのチャネルに属している (Suzuki, M. 他 (1999) J. Biol. Chem. 274 : 6330-6335)。

40

【0022】

電位依存性カチオンチャネルのポア形成サブユニットは、イオンチャネルタンパク質のスーパーファミリーを形成する。チャネルタンパク質の特徴的なドメインには、6つの膜貫通ドメイン（S1～S6）、S5とS6との間に配置されたポア形成領域（P）及び、細胞内にあるアミノ末端及びカルボキシ末端がある。Na⁺及びCa²⁺サブファミリーでは、このドメインは4回繰り返され、K⁺チャネルサブファミリーでは、各チャネルは、同一または非類似のいずれかであるサブユニットの四量体から形成される。P領域は、チャネルに対するイオン選択性を特定する情報を含む。K⁺チャネルの場合、GYGT

50

リペプチドがこの選択性に関与する (Ishii, T. M. 他 (1997) Proc. Natl. Acad. Sci. USA 94 : 11651-11656)。

【0023】

電位依存性 Na^+ 及び K^+ チャンネルは、神経及び筋肉細胞などの電氣的興奮性細胞の機能に必要である。活動電位は、神経伝達物質放出及び筋収縮の原因となるものであるが、 Na^+ 及び K^+ イオンへの膜の浸透性における大きな一過的变化から生じる。閾値レベルを超えた膜の脱分極は、電位依存性 Na^+ チャンネルを開く。ナトリウムイオンは、細胞へ流れ、更に膜を脱分極してより多く電位依存性 Na^+ チャンネルを開き、このようにして脱分極が細胞に沿って伝わっていく。脱分極はまた、電位依存性カリウムチャンネルを開く。結果として、カリウムイオンは外向きに流れ、それによって膜の再分極が生じる。電位依存性チャンネルは、第4膜貫通セグメント (S4) において荷電した残基を利用して電圧変化を検出する。開状態は、約1ミリ秒しか持続せず、その約1ミリ秒後に、チャンネルは膜電位と無関係に開き得ない不活性化状態に自発的に変換する。不活性化は、ポアを閉じるプラグとして作用するようなチャンネルのN末端によって媒介される。不活性化状態から閉状態へ遷移するには、静止電位に戻すことが必要である。 10

【0024】

電位依存性 Na^+ チャンネルは、2つの更に小さな補助サブユニット $\beta 1$ 及び $\beta 2$ に関連する260 kDaのポア形成 α サブユニットから構成されるようなヘテロ三量体の複合体である。 $\beta 2$ サブユニットは、細胞外Igドメインを含む複合膜糖タンパク質であり、その α 及び $\beta 1$ サブユニットとの関連は、チャンネルの機能的発現の増加、そのゲーティング特性の変化、及び膜表面面積の増加に起因する全細胞キャパシタンスの増加と関連がある (Isom, L. L. 他 (1995) Cell 83 : 433-442)。 20

【0025】

非電位依存性 Na^+ チャンネルには、アミロライド感受性 Na^+ チャンネル/degenerin (NaC/DEG) ファミリーのメンバーが含まれる。このファミリーのチャンネルサブユニットは、アミノ末端及びカルボキシル末端が細胞内にあり、長い細胞外ループに隣接する2つの膜貫通ドメインを含むと考えられる。 NaC/DEG ファミリーには、気道、遠位結腸、腎臓の皮質集合管及び外分泌腺導管を含む上皮における Na^+ 再吸収に関与する上皮性 Na^+ チャンネル (EnaC) がある。 EnaC における突然変異は、I型偽性低アルドステロン症及びリドル症候群 (偽性高アルドステロン症) を引き起こす。 NaC/DEG ファミリーには、最近特徴付けられた H^+ 依存性カチオンチャンネルまたは酸感受性イオンチャンネル (ASIC) もある。 ASIC サブユニットは、脳で発現し、ヘテロ多量体の Na^+ 浸透性チャンネルを形成する。これらのチャンネルは、活性化のために酸性pH変動を必要とする。 ASIC サブユニットは、元々は線虫から単離された機械依存性チャンネルのファミリーであるdegenerinとの相同性を示す。Degenerinでの突然変異は、神経変性を引き起こす。組織アシドーシスが痛みを発生させるものであるので、 ASIC サブユニットは、神経機能または痛みの知覚においても役割を有し得る (Waldmann, R. および M. Lazdunski (1998) Curr. Opin. Neurobiol. 8 : 418-424 ; Eglén, R. M. 他 (1999) Trends Pharmacol. Sci. 20 : 337-342)。 30 40

【0026】

K^+ チャンネルはすべての細胞種に存在し、電圧、ATP濃度、または Ca^{2+} 及びcAMPなどのセカンドメッセンジャーによって K^+ チャンネルを制御し得る。非興奮性組織では、 K^+ チャンネルが、タンパク質合成、内分泌性分泌の制御及び膜内外の浸透平衡の維持に関与している。ニューロン及びその他の興奮性細胞では、活動電位の調整及び膜の再分極に加えて、 K^+ チャンネルが静止膜電位の設定に関与する。サイトゾルは、非拡散性アニオンを含み、この負の電荷総ての平衡を保持するために、細胞には Na^+ 、 K^+ 及び Cl^- の再分布を提供する Na^+/K^+ ポンプ及びイオンチャンネルが含まれる。ポンプは、能動的に Na^+ を細胞へ輸送し、 K^+ を細胞内に輸送する ($\text{Na}^+:\text{K}^+$ の 50

割合は3:2である)。原形質膜のイオンチャンネルは、 K^+ 及び Cl^- が受動的な拡散により流出入することを許容する。サイトゾル内の電荷が高度に負であるので、 Cl^- は細胞外に流出する。 K^+ の流れは、 K^+ を細胞内に引き込む起電力及び K^+ を細胞外に押し出す濃度勾配によって均衡が保たれる。従って、静止膜電位は、主として K^+ の流れによって調整される (Salzberg, L. および T. J. J. (1995) *Neuron* 15 : 489-492)。

【0027】

Shaker様スーパーファミリーのカリウムチャンネルサブユニットはすべて6つの膜貫通ドメイン/1つのポアドメイン構造の特徴を有する。4つのサブユニットは、ホモまたはヘテロ四量体として結合し、機能的Kチャンネルを形成する。これらのポア形成サブユニットは、チャンネル不活性化動力学を変化させるような種々の細胞質βサブユニットとも関連している。Shaker様チャンネルファミリーには、QT延長症、心律動異常症候群に関連するようなHetherington-agogo関連遺伝子(HERG)などの遅延整流型チャンネルや、電圧依存性Kチャンネルがある(Curran, M. E. (1998) *Curr. Opin. Biotechnol.* 9 : 565-572, Kaczarowski, G. J. および M. L. Garcia (1999) *Curr. Opin. Chem. Biol.* 3 : 448-458)。

【0028】

K^+ チャンネルの第2のスーパーファミリーは、内向き整流型チャンネル(Kir)で構成される。Kirチャンネルは、優先的に K^+ の流れを内向きに導く特性を有する。これらのタンパク質は、1つのカリウム選択性ポアドメイン及び2つの膜貫通ドメインから構成され、これらは電位依存性Kチャンネルの第5及び第6膜貫通ドメインに対応する。Kirサブユニットは、四量体としても会合している。Kirファミリーには、腎尿細管疾患のバスター症候群の原因となる突然変異であるROMK1がある。Kirチャンネルは、心臓ペースメーカー活動の調整、発作及びてんかん、並びにインシュリン調整にも関与している(Dougnik, C. A. 他 (1995) *Curr. Opin. Neurobiol.* 5 : 268-277、前出のCurran)。

【0029】

最近認識されたTWIK K^+ チャンネルファミリーには、哺乳類のTWIK-1、TREK-1及びTASKタンパク質がある。このファミリーのメンバーは、4つの膜貫通ドメイン及び2つのPドメインを含む全体構造を有する。これらのタンパク質は、おそらく多くの細胞種における静止電位の制御に関与している(Duprat, F. 他 (1997) *EMBO J* 16 : 5464-5471)。

【0030】

電位依存性 Ca^{2+} チャンネルは、電気生理学的特性及び薬理学的特性に基づき、幾つかのサブタイプに区分されてきた。L型 Ca^{2+} チャンネルは、心筋及び骨格筋に優性に発現し、興奮収縮連関において不可欠な役割を果たす。T型チャンネルは、心臓ペースメーカー活動に重要であり、N型及びP/Q型チャンネルは、中枢及び抹消神経系における神経伝達物質放出の制御に関与している。L型及びN型電位依存性 Ca^{2+} チャンネルは精製されており、両チャンネルは機能が非常に異なるが類似のサブユニット組成を有する。チャンネルは、3つのサブユニットを含む。 α_1 サブユニットは、膜ポア及び電圧センサを形成し、 $\alpha_2\delta$ 及び β サブユニットは、チャンネルの電圧依存、ゲーティング特性及び電流振幅を調節する。これらのサブユニットは、少なくとも6つの α_1 、1つの $\alpha_2\delta$ 、4つの β 遺伝子によってコードされる。第4サブユニット γ は、骨格筋において同定されてきた(Walker, D. 他 (1998) *J. Biol. Chem.* 273 : 2361-2367, McCleskey, E. W. (1994) *Curr. Opin. Neurobiol.* 4 : 304-312)。

【0031】

生化学的特性が研究されてきた高電圧活性化 Ca^{2+} チャンネルは、約190~250 kDaのポア形成アルファ1サブユニットの複合体、アルファ2およびデルタサブユニッ

10

20

30

40

50

トの膜貫通複合体、細胞内ベータサブユニットおよび場合によっては膜貫通ガンマサブユニットを含む。さまざまなアルファ1サブユニット、アルファ2デルタ複合体、ベータサブユニットおよびガンマサブユニットが知られている。アルファ1サブユニットのCav1ファミリーはLタイプCa²⁺電流を流し、それによって筋肉の収縮、内分泌および遺伝子の転写が起こる。LタイプCa²⁺電流の調節は主にセカンドメッセンジャー活性化タンパク質リン酸化経路によって行われる。アルファ1サブユニットのCav2ファミリーはNタイプ、P/QタイプおよびRタイプのCa²⁺電流を流し、それによって急速なシナプス伝達が起こる。調節は主にGタンパク質およびSNAREタンパク質との直接の相互作用によって、そして補助的にタンパク質リン酸化によって行われる。アルファ1サブユニットのCav3ファミリーはTタイプのCa²⁺電流を流れるが、この活性化と不活性化は他のCa²⁺電流タイプと比べてより速く起こり、またより負の膜電位で起こる。これら3つのファミリーのCa²⁺チャネルの異なる構造と調節パターンにより、膜電位の変化に応答してさまざまなCa²⁺流入経路が提供され、またセカンドメッセンジャー経路および相互作用するタンパク質によってCa²⁺の流入を調節するさまざまな可能性が提供される (Catterall, W. A. (2000) Annu. Rev. Cell Dev. Biol. 16:521555)。

10

【0032】

電圧依存性Ca²⁺チャネルのアルファ2サブユニットは1つ以上のCacheドメインを含んでいる可能性がある。細胞外Cacheドメインは、細胞内触媒ドメイン (ヒスチジンキナーゼ、PP2Cホスホターゼ、GGDEF (推定ジグアニレートシクラーゼ)、HD-GYP (推定ホスホジエステラーゼ)、またはアデニルシクラーゼドメイン等) あるいは非触媒ドメイン (メチル受容DNA結合ウィング付きヘリックス・ターン・ヘリックス、GAF、PASまたはHAMP (istidine キナーゼ、デニルシクラーゼ、エチル結合タンパク質、およびホスホターゼに見られるドメイン) 等) に融合している可能性がある。小さな分子はCacheドメインを通じて結合し、このシグナルが細胞内ドメインに依存して多様な出力に変換される (Anantharaman, V. および Aravind, L. (2000) Trends Biochem. Sci. 25:535537)。

20

【0033】

カルシウムイオンチャネルの過渡的な受容体ファミリー (Trp) は容量性カルシウム流入 (CCE) を媒介すると考えられている。CCEは、イノシトール三リン酸 (IP3) や多くのホルモンや成長因子に反応する他の物質の作用によって消耗されるCa²⁺貯蔵を再供給するためのCa²⁺の細胞内への流入である。TrpおよびTrp様物質は、最初にショウジョウバエからクローンされ、S3 から S6 領域においてCa²⁺チャネルにゲート開閉をする電位に類似性を持つ。これはTrpやその関連タンパク質が、哺乳動物のCCEチャネルを形成することを示唆する (Zhu, X. 他 (1996) Cell 85:661-671; Boulay, G. 他 (1997) J. Biol. Chem. 272:29672-29680)。メラスタチン (melastatin) はマウスおよびヒトの両方において単離される遺伝子であり、黒色腫細胞におけるその発現は、生体内における黒色腫の侵襲性と逆比例している。このヒトcDNA転写物はTrpファミリーのメンバーに相同体を有する1533基のアミノ酸をもつタンパク質に相当する。メラスタチンmRNA発現状態と腫瘍の厚さを組み合わせて利用すると、転移性疾患の発症に対するリスクの低いグループと高いグループに患者のグループ分けを決定できる可能性が提示されている (Duncan, L. M. 他 (2001) J. Clin. Oncol. 19:568-576)。

30

40

【0034】

クロライドチャネルは、内分泌性分泌及びサイトゾルオルガネラのpH調整に必要である。分泌性上皮細胞では、Cl⁻はNa⁺、K⁺/Cl⁻共輸送体によって側底膜を通過して細胞に入り、電気化学的平衡濃度以上で細胞内に蓄積する。ホルモンの刺激に応じて、尖端表面からのCl⁻を分泌することにより、Na⁺及び水が分泌性ルーメンに流入

50

する。嚢胞性線維症膜通過伝導制御因子 (CFTR) は、ヒトによく見られる致命的な遺伝病である嚢胞性線維症のための遺伝子によってコードされるクロライドチャンネルである。CFTRはABC輸送体ファミリーのメンバーであり、6つの膜貫通ドメインを各々有するような2つのドメインとその後に続くヌクレオチド結合部位から構成される。CFTR機能の損失は、経上皮水分泌を減少させ、結果的に、呼吸樹、膵管及び腸を被膜する粘液の層は、脱水されて、取り除くのが困難になる。結果的にこれらの部位が閉塞することにより、膵不全、「胎便性イレウス」及び悲惨な「慢性閉塞性肺疾患」を生じさせる (Al-Awqati, Q. 他 (1992) J. Exp. Biol. 172 : 245-266)。

【0035】

10

電位依存性クロライドチャンネル (CLC) は、CBSドメインとして知られる2つの小球形ドメインの他に10~12個の膜貫通ドメインによって特徴付けられる。CLCサブユニットは、おそらくホモ四量体として機能する。CLCタンパク質は、細胞容積、膜電位安定化、シグナル伝達及び経上皮輸送の調整に関与している。骨格筋で主に発現されるようなCLC-1の突然変異は、常染色体劣性全身性ミオトニー及び常染色体優性先天性ミオトニーの原因であり、腎臓チャンネルCLC-5における突然変異は、腎臓結石を引き起こす (Jentsch, T. J. (1996) Curr. Opin. Neurobiol. 6 : 303-310)。

【0036】

リガンド依存性チャンネルがポアを開くのは、細胞外または細胞内メディエータがチャンネルに結合するときである。神経伝達物質依存性チャンネルは、神経伝達物質が細胞外ドメインに結合すると開くチャンネルである。これらのチャンネルは、神経または筋肉細胞のシナプス後膜に存在する。神経伝達物質依存性チャンネルには、2つの型がある。ナトリウムチャンネルは、アセチルコリン、グルタミン酸及びセロトニンなどの興奮性神経伝達物質に応じて開く。こうして開くことによって、 Na^+ の流入を引き起こし、電位依存性チャンネルを活性化して活動電位を開始する最初の局在脱分極を提供する。クロライドチャンネルは、 γ アミノ酪酸 (GABA) 及びグリシンなどの抑制神経伝達物質に応じて開き、膜の過分極及びそれに続く活動電位の生成を引き起こす。神経伝達物質依存性イオンチャンネルは、4つの膜貫通ドメインを有し、おそらく五量体として機能する (前出のJentsch)。第2膜貫通ドメインのアミノ酸は、チャンネル浸透及び選択性の決定において重要であるように見える (Sather, W. A. 他 (1994) Curr. Opin. Neurobiol. 4 : 313-323)。

20

30

【0037】

リガンド依存性チャンネルは、細胞内のセカンドメッセンジャーによって制御することができる。例えば、カルシウム活性化 K^+ チャンネルは、内部カルシウムイオンによってゲーティングされる。神経細胞では、脱分極中のカルシウムの流入は、 K^+ チャンネルを開いて活動電位の大きさを調節する (前出のIshii 他)。大コンダクタンス (BK) チャンネルは、脳から精製され、そのサブユニット組成が決定された。BKチャンネルの α サブユニットは、電位依存性 K^+ チャンネルとは対照的に6つよりもむしろ7つの膜貫通ドメインを有する。余分の膜貫通ドメインは、サブユニットN末端に位置する。サブユニットのC末端領域 (「カルシウムボール」領域) の28個のアミノ酸領域は、負に荷電した残基を多数含み、カルシウム結合に責任を果たす領域であると考えられる。 β サブユニットは、グリコシル化細胞外ループによって結合された2つの膜貫通ドメインを含み、細胞内にN末端及びC末端を有する (前出のKaczorowski, Vergara, C. 他 (1998) Curr. Opin. Neurobiol. 8 : 321-329)。

40

【0038】

サイクリックヌクレオチド依存性 (CNG) チャンネルは、サイトゾルのサイクリックヌクレオチドによってゲーティングされる。最も良い例は、嗅覚に関与するcAMP依存性 Na^+ チャンネルと、視覚に関与するcGMP依存性カチオンチャンネルである。両システム

50

は、Gタンパク質結合受容体のリガンド媒介活性化に関与し、Gタンパク質結合受容体は次に細胞内でのサイクリックヌクレオチドのレベルを変化させる。CNGチャネルはまた、 Ca^{2+} がニューロンに入る主要経路を示し、ニューロンの発達及び形成性において役割を果たす。CNGチャネルは、少なくとも2つのタイプのサブユニット即ち機能的ホモマーチャネルを形成し得る α サブユニットとチャネル特性を調節する β サブユニットとを含む四量体である。すべてのCNGサブユニットは、電位依存性 K^+ チャネルに類似の、6つの膜貫通ドメイン及び第5及び第6膜貫通ドメインの間のポア形成領域を有する。大きなC末端ドメインにはサイクリックヌクレオチド結合ドメインが含まれ、N末端ドメインはチャネルサブタイプ間で変異を与える (Zufall, F. 他 (1997) Curr. Opin. Neurobiol. 7 : 404-412)。

10

【0039】

別のタイプのイオンチャネルタンパク質の活性は、種々の細胞内シグナル伝達タンパク質によっても調節し得る。多くのチャネルは、1個若しくは数個のタンパク質キナーゼによるリン酸化のための部位を有する。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、チロシンキナーゼ及びカゼインキナーゼIIがあり、これらはすべて細胞におけるイオンチャネル活性を調節する。Kirチャネルは、ヘテロ三量体のGタンパク質のGpyサブユニットの結合により活性化される (Reimann, F. および F. M. Ashcroft (1999) Curr. Opin. Cell Biol. 11 : 503-508)。別のタンパク質は、細胞膜の特定部位へのイオンチャネルの局在化に関与している。このようなタンパク質にはPDZドメインタンパク質があり、これはニューロンのシナプスでイオンチャネルのクラスター形成を調整するMAGUK (膜関連グアニル酸キナーゼ) として知られている (Craven, S. E. および D. S. Bredt (1998) Cell 93 : 495-498)。

20

【0040】

疾患との相関関係

無数のヒトの疾患及び障害の原因は、膜を通過する分子の輸送における欠陥に起因し得る。膜結合輸送体及びイオンチャネルの輸送における欠陥は、嚢胞性繊維症、グルコースガラクトース吸収不良症候群、高コレステロール血症、フォンギルケ病及び或る種の糖尿病など幾つかの疾患に関連している。膜を通過して小分子を輸送できないことに起因する単一の遺伝子欠陥疾患には、例えばシスチン尿症、イミノグリシン尿症、Hartup病及びファンコーニ病がある (van 't Hoff, W. G. (1996) Exp. Nephrol. 4 : 253-262, Talente, G. M. 他 (1994) Ann. Intern. Med. 120 : 218-226、及びChillon, M. 他 (1995) New Engl. J. Med. 332 : 1475-1480)。

30

【0041】

イオンチャネル遺伝子における突然変異に起因するヒトの疾患には、骨格筋、心筋及び中枢神経系の疾患がある。ナトリウムチャネル及びクロライドチャネルのポア形成サブユニットにおける突然変異は、ミオトニーを引き起こす。ミオトニーは、随意収縮後の弛緩が遅れる筋肉障害である。ナトリウムチャネルミオトニーは、チャネル遮断薬で治療されてきた。筋肉ナトリウム及びカルシウムチャネルにおける突然変異は、周期性麻痺の形成を引き起こし、筋形質カルシウム放出チャネル、T細管カルシウムチャネル及び筋肉ナトリウムチャネルにおける突然変異は、悪性高熱症を引き起こす。QT延長症候群及び特発性心室細動などの心不整脈疾患は、カリウム及びナトリウムチャネルにおける突然変異に起因する (Cooper, E. C. および L. Y. Jan (1998) Proc. Natl. Acad. Sci. USA 96 : 4759-4766)。既知のヒト真性てんかん遺伝子は、4つがすべてイオンチャネルタンパク質をコードする (Berkovic, S. F. および I. E. Scheffer (1999) Curr. Opin. Neurology 12 : 177-182)。運動失調、片麻痺性片頭痛及び遺伝性難聴などのその他の神経疾患もまた、イオンチャネル遺伝子における突然変異

40

50

に起因し得る (Jen, J. (1999) Curr. Opin. Neurobiol. 9 : 274-280、前出のCooper)。

【0042】

イオンチャネルは、多くの薬物治療法に対する標的であった。神経伝達物質依存性チャネルは、不眠症、不安、抑うつ症及び精神分裂病の治療のための治療法において標的にされてきた。電位依存性チャネルは、不整脈、虚血性発作、頭部外傷及び神経変性疾患の治療法において標的にされてきた (Taylor, C. P. および L. S. Narasimhan (1997) Adv. Pharmacol. 39 : 47-98)。種々のクラスのイオンチャネルは、痛みの知覚においても重要な役割を果たすので、新たな鎮痛薬の潜在的標的である。このようなイオンチャネルにはバニロイド (vanilloid) 依存性イオンチャネルがあり、これは有害熱によってのみならずバニロイドカプサイシンによっても活性化される。電位依存性Na⁺チャネルを遮断するリドカインやメキシレチンなどの局部麻酔薬は、神経障害性の痛みの治療において有用であった (前出のEglen)。

10

【0043】

免疫調節のための標的として、最近では免疫系におけるイオンチャネルが示唆されている。T細胞活性化は、カルシウムシグナル伝達に依存し、多様なT細胞特異イオンチャネルは、このシグナル伝達過程に影響するように特徴付けられてきた。チャネルブロッカーは、リンホカインの分泌、細胞増殖、及び標的細胞の死滅を阻害することができる。T細胞カリウムチャネルKv1.3のペプチドアンタゴニストは、ブタにおける遅延型過敏症及び異質遺伝子型反応を抑制し、安全で効果的な免疫抑制薬としてのチャネル遮断薬の使用のアイデアを実証している (Calahan, M. D. および K. G. Chand y (1997) Curr. Opin. Biotechnol. 8 : 749-756)。

20

【0044】

新たな輸送体及びイオンチャネル及びこれらをコードするポリヌクレオチドの発見は、輸送障害、神経障害、筋肉疾患及び免疫疾患および細胞増殖異常の診断、治療並びに予防において、また、輸送体及びイオンチャネルの核酸及びアミノ酸配列の発現に対する外因性化合物の効果の評価において有用であるような新たな組成物を提供することにより、当分野における必要性を満たす。

30

【0045】

(発明の概要)

本発明は、総称して「TRICH」、個別にはそれぞれ「TRICH-1」、「TRICH-2」、「TRICH-3」、「TRICH-4」、「TRICH-5」、「TRICH-6」、「TRICH-7」、「TRICH-8」、「TRICH-9」、「TRICH-10」、「TRICH-11」、「TRICH-12」、「TRICH-13」、「TRICH-14」、「TRICH-15」、「TRICH-16」、「TRICH-17」、「TRICH-18」、「TRICH-19」、および「TRICH-20」と呼ぶ輸送体及びイオンチャネルである精製されたポリペプチドを提供する。或る実施態様において本発明は、(a) SEQ ID NO: 1-20からなる群から選択したアミノ酸配列からなるポリペプチド、(b) SEQ ID NO: 1-20からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-20からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1-20からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を含む群から選択した単離されたポリペプチドを提供する。一実施態様では、SEQ ID NO: 1-20のアミノ酸配列を含む単離されたポリペプチドを提供する。

40

【0046】

また、本発明は(a) SEQ ID NO: 1-20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: 1-20からなる群から選択したアミノ

50

酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列からなるポリペプチド、
(c) SEQ ID NO: 1-20からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-20からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群より選択されたポリペプチドをコードするような単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドはSEQ ID NO: 1-20からなる群から選択したポリペプチドをコードする。別の実施態様では、ポリヌクレオチドはSEQ ID NO: 21-40を有する群から選択される。

【0047】

本発明は更に、(a) SEQ ID NO: 1-20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: 1-20からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-20からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-20からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群から選択したポリペプチドをコードするようなポリヌクレオチドと機能的に連結したプロモーター配列を有する組換えポリヌクレオチドを提供する。一実施態様では、本発明は組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様では、本発明は組換えポリヌクレオチドを含む遺伝形質転換体を提供する。

【0048】

また、本発明は、(a) SEQ ID NO: 1-20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: 1-20からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を有するポリペプチド、(c) SEQ ID NO: 1-20からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-20からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片から構成される群から選択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) 組換えポリヌクレオチドを用いて形質転換した細胞をポリペプチドの発現に適した条件下で培養する過程と、(b) そのように発現したポリペプチドを回収する過程とを有し、組換えポリヌクレオチドはポリペプチドをコードするポリヌクレオチドに機能的に連結したプロモーター配列を有する。

【0049】

本発明は更に、(a) SEQ ID NO: 1-20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: 1-20からなる群から選択したアミノ酸配列と少なくとも90%同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-20からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-20からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片から構成される群から選択されたポリペプチドに特異結合するような単離された抗体を提供する。

【0050】

本発明は更に、(a) SEQ ID NO: 21-40からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(b) SEQ ID NO: 21-40からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c) (a) に相補的なポリヌクレオチド、(d) (b) に相補的なポリヌクレオチド、および(e) (a) ~ (d) のRNA等価物からなる群から選択された単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドは少なくとも60の連続したヌクレオチドを有する。

【0051】

本発明は更に、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは(a) SEQ ID NO: 21-40からなる群から選択した

ポリヌクレオチド配列を有するポリヌクレオチド、(b) SEQ ID NO: 21-40 からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を有するポリヌクレオチド、(c) (a) に相補的なポリヌクレオチド、(d) (b) に相補的なポリヌクレオチド、および(e) (a) ~ (d) のRNA等価物からなる群から選択されたポリヌクレオチドの配列を有する。検出方法は、(a) サンプル中の前記の標的ポリヌクレオチドに相補的な配列を含む、少なくとも20の連続したヌクレオチドを含むプローブを用いて該サンプルをハイブリダイズする過程と、(b) ハイブリダイゼーション複合体の存在・不存在を検出し、複合体が存在する場合にはオプシオンでその量を検出する過程からなり、プローブと標的ポリヌクレオチドあるいはその断片の間でハイブリダイゼーション複合体が形成されるような条件下で、プローブは標的ポリヌクレオチドに特異的にハイブリダイズする。一実施態様では、プローブは少なくとも60の連続したヌクレオチドを含む。

10

【0052】

本発明はまた、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは、(a) SEQ ID NO: 21-40 からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(b) SEQ ID NO: 21-40 からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c) (a) 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された配列のポリヌクレオチドを有する。検出方法は、(a) ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、(b) 標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプシオンでその量を検出する過程を含む。

20

【0053】

本発明は更に、有効量のポリペプチドと薬剤として許容できる賦形剤とを含む組成物を提供する。有効量のポリペプチドは、(a) SEQ ID NO: 1-20 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: 1-20 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を含むポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列の免疫原性断片からなる群から選択される。一実施例では、SEQ ID NO: 1-20 からなる一群から選択されたアミノ酸配列を含む組成物を提供する。更に、本発明は、患者にこの組成物を投与することを含む、機能的TRICHの発現の低下に関連した疾患やその症状の治療方法を提供する。

30

【0054】

本発明はまた、(a) SEQ ID NO: 1-20 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: 1-20 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群から選択されたポリペプチドのアゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドを含むサンプルを化合物に曝す過程と、(b) サンプル中のアゴニスト活性を検出する過程とを含む。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的TRICHの発現の低下に関連した疾患やその症状の治療方法を提供する。

40

【0055】

本発明は更に、(a) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列

50

を含むポリペプチド、(b) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群から選択されたポリペプチドのアンタゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドを含むサンプルを化合物に曝す過程と、(b) サンプル中のアンタゴニスト活性を検出する過程とを含む。一実施態様で本発明は、この方法によって同定したアンタゴニスト化合物と薬剤として許容できる賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的 TRICH の過剰な発現に関連した疾患やその症状の治療方法を提供する。

10

【0056】

本発明は更に、(a) SEQ ID NO: 1-20 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: 1-20 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を含む群から選択されたポリペプチドに特異結合する化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドを適切な条件下で少なくとも1つの試験化合物と混合させる過程と、(b) 試験化合物とのポリペプチドの結合を検出し、それによってポリペプチドに特異結合する化合物を同定する過程とを含む。

20

【0057】

本発明は更に、(a) SEQ ID NO: 1-20 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: 1-20 からなる群から選択したアミノ酸配列と少なくとも90%の相同性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1-20 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を含む群から選択されたポリペプチドの活性を調節する化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドの活性が許容された条件下で、ポリペプチドを少なくとも1つの試験化合物と混合させる過程と、(b) ポリペプチドの活性を試験化合物の存在下で算定する過程と、(c) 試験化合物の存在下でのポリペプチドの活性を試験化合物の不存在下でのポリペプチドの活性と比較する過程とを含み、試験化合物の存在下でのポリペプチドの活性の変化は、ポリペプチドの活性を調節する化合物であることを意味する。

30

【0058】

本発明は更に、標的ポリヌクレオチドの変異発現の有効性を確認するために化合物をスクリーニングする方法を提供する。標的ポリヌクレオチドは、SEQ ID NO: 21-40 からなる群から選択した配列を含む。スクリーニング方法は、(a) 標的ポリヌクレオチドを含むサンプルを化合物に曝す過程と、(b) 標的ポリヌクレオチドの変異発現を検出する過程とを含む。

40

【0059】

本発明は更に、(a) 核酸を含む生物学的サンプルを試験化合物で処理する過程と、(b) (i) SEQ ID NO: 21-40 からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(ii) SEQ ID NO: 21-40 からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(iii) (i) に相補的な配列を有するポリヌクレオチド、(iv) (ii) のポリヌクレオチドに相補的なポリヌクレオチド、(v) (i) ~ (iv) の RNA 等価物からなる群から選択したポリヌクレオチドの少なくとも20の連続し

50

たヌクレオチドを含むプローブを用いて、処理した生物学的サンプルの核酸をハイブリダイズする過程を含む、試験化合物の毒性の算定方法を提供する。ハイブリダイゼーションは、前記プローブと生物学的サンプル中の標的ポリヌクレオチドの間に特定のハイブリダイゼーション複合体が形成されるような条件下で発生し、前記標的ポリヌクレオチドは、
(i) SEQ ID NO: 21-40 からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、
(ii) SEQ ID NO: 21-40 からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、
(iii) (i) のポリヌクレオチドに相補的な配列を有するポリヌクレオチド、
(iv) (ii) のポリヌクレオチドに相補的なポリヌクレオチド、および
(v) (i) ~ (iv) のRNA等価物からなる群から選択する。或いは、標的ポリヌクレオチドは、上記(i) ~ (v) からなる群から選択したポリヌクレオチド配列の断片と、
(c) ハイブリダイゼーション複合体の量を定量する過程と、
(d) 処理した生物学的サンプルのハイブリダイゼーション複合体の量を、非処理の生物学的サンプルのハイブリダイゼーション複合体の量と比較する過程を含み、処理した生物学的サンプルのハイブリダイゼーション複合体の量との差は、試験化合物の毒性を意味する。

10

20

30

40

50

【0060】

(発明を実施するための形態)

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

【0061】

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

【0062】

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

【0063】

(定義)

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

【0064】

用語「アゴニスト」は、TRICHの生物学的活性を強めたり、模倣する分子を指す。このアゴニストは、TRICHに直接相互作用するか、或いはTRICHが関与する生物学的経路の成分と作用して、TRICHの活性を調節するタンパク質、核酸、糖質、小分子、任意の他の化合物や組成物を含み得る。

【0065】

用語「対立遺伝子変異配列」は、TRICHをコードする遺伝子の別の形を指す。対立遺伝子変異体は、核酸配列における少なくとも1つの突然変異から作製し得る。また、変異

R N Aまたはポリペプチドからも作製し得る。ポリペプチドの構造または機能は、変異することもしないこともある。遺伝子は、天然の対立遺伝子変異体を全く有しないか、1個若しくは数個の天然の対立遺伝子変異体を有し得る。一般に対立遺伝子変異体を生じさせる通常の突然変異性変化は、ヌクレオチドの自然欠失、付加または置換に帰するものである。これら各変化は、単独或いは他の変化と共に、所定の配列内で1回若しくは数回生じ得る。

【0066】

T R I C Hをコードする「変異」核酸配列は、様々なヌクレオチドの欠失、挿入、或いは置換が起こっても、T R I C Hと同じポリペプチド或いはT R I C Hの機能特性の少なくとも1つを備えるポリペプチドを指す。この定義には、T R I C Hをコードするポリヌクレオチド配列の正常な染色体の遺伝子座ではない位置での対立遺伝子変異配列との不適合或いは予期しないハイブリダイゼーション、並びにT R I C Hをコードするポリヌクレオチドの特定のオリゴヌクレオチドプローブを用いて容易に検出可能な或いは検出困難な多型性を含む。コードされたタンパク質も変異され得り、サイレント変化を生じT R I C Hと機能的に等価となるアミノ酸残基の欠失、挿入、或いは置換を含み得る。意図的なアミノ酸置換は、生物学的或いは免疫学的にT R I C Hの活性が保持される範囲で、残基の極性、電荷、溶解度、疎水性、親水性、及び／または両親媒性についての類似性に基づいて成され得る。例えば、負に帯電したアミノ酸にはアスパラギン酸及びグルタミン酸があり、正に帯電したアミノ酸にはリジン及びアルギニンがある。親水性値が近似している非荷電極性側鎖を有するアミノ酸には、アスパラギンとグルタミン、セリンとトレオニンがある。親水性値が近似している非荷電側鎖を有するアミノ酸には、ロイシン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。

【0067】

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

【0068】

用語「増幅」は、核酸配列の複製物を作製することに関連する。増幅は通常、当業者によく知られたポリメラーゼ連鎖反応（P C R）技術を用いて行う。

【0069】

用語「アンタゴニスト」は、T R I C Hの生物学的活性を阻害或いは減弱する分子である。アンタゴニストは、T R I C Hに直接相互作用するか、或いはT R I C Hが関与する生物学的経路の成分と作用して、T R I C Hの活性を調節する抗体、核酸、糖質、小分子、任意の他の化合物や組成物などのタンパク質を含み得る。

【0070】

「抗体」の語は、エピトープの決定基と結合することができる、そのままの免疫グロブリンやその断片、例えばF a、F（a b'）₂及びF v断片を指す。T R I C Hポリペプチドと結合する抗体は、免疫抗原として、そのままのポリペプチド、または関心のある小ペプチドを含む断片を用いて作製可能である。動物（マウス、ラット、ウサギ等）を免疫化するために用いるポリペプチドまたはオリゴペプチドは、R N Aの翻訳、または化学合成によって得られるポリペプチドまたはオリゴペプチドに由来し得るもので、好みに応じてキャリアタンパク質に抱合することも可能である。通常用いられるキャリアであってペプチドと化学結合するものは、ウシ血清アルブミン、サイログロブリン及びスカシガイのヘモシアニン（K L H）等がある。その結合ペプチドは、動物を免疫化するために用いる。

【0071】

用語「抗原決定基」は、特定の抗体と接触する分子の領域（即ちエピトープ）を指す。タンパク質またはタンパク質断片を用いて宿主動物を免疫化する場合、タンパク質の多数の

領域が、抗原決定基（タンパク質の特定の領域または3次元構造）に特異結合する抗体の産生を誘導し得る。抗原決定基は、抗体に結合するための無損傷抗原（即ち免疫応答を誘導するために用いられる免疫原）と競合し得る。

【0072】

用語「アプタマー」は、特定の分子標的に結合する核酸またはオリゴヌクレオチドを指す。アプタマーは *in vitro* の進化過程（例えば米国特許番号第5, 270, 163号に記載された *SELEX* (*Systematic Evolution of Ligands by EXponential Enrichment* の略、試験管内選択法)) から由来するもので、そのような過程は大きな組み合わせライブラリから標的特異的なアプタマー配列を選択する。アプタマー組成は、2本鎖または1本鎖であってもよく、デオキシリボヌクレオチド、リボヌクレオチド、ヌクレオチド誘導体、または他のヌクレオチド様分子を含み得る。アプタマーのヌクレオチド成分は、修飾された糖基（例えばリボヌクレオチドの2'-OH基が2'-Fまたは2'-NH₂で置換し得る）を有することが可能で、そのような糖基はヌクレアーゼへの抵抗性または血液中でのより長い寿命などの望ましい性質に改善し得る。循環系からアプタマーが除去される速度を遅くするために、アプタマーを高分子量キャリア等の分子に抱合させることができる。アプタマーは、たとえば光活性化または架橋剤によって各々のリガンドと特異的に架橋させることができる。（Brody, E. N. および L. Gold (2000) *J. Biotechnol.* 74: 5-13等を参照。）

10

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

20

【0073】

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

【0074】

本明細書において「アンチセンス」は、特定の核酸配列のセンス（コーディング）鎖と塩基対を形成し得る任意の組成物を指す。アンチセンス成分には、DNAや、RNAや、ペプチド核酸（PNA）や、ホスホロチオ酸、メチルホスホン酸またはベンジルホスホン酸等の修飾されたバックボーン結合を有するオリゴヌクレオチドや、2'-メトキシエチル糖または2'-メトキシエトキシ糖等の修飾された糖類を有するオリゴヌクレオチドや、或いは5-メチルシトシン、2'-デオキシウラシルまたは7-デアザ-2'-デオキシグアノシン等の修飾された塩基を有するオリゴヌクレオチドがある。アンチセンス分子は、化学合成または転写を含む任意の方法で製造することができる。相補的アンチセンス分子は、ひとたび細胞に導入されたら、細胞が形成した天然の核酸配列と塩基対を形成し、転写または翻訳を妨害する二重鎖を形成する。「負」または「マイナス」という表現は、ある参考DNA分子のアンチセンス鎖を意味し、「正」または「プラス」という表現はある参考DNA分子のセンス鎖を意味する。

30

40

【0075】

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

【0076】

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

【0077】

50

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

【0078】

10

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

【0079】

用語「保存的なアミノ酸置換」は、元のタンパク質の特性を殆ど変えない置換を指す。即ち、置換によってそのタンパク質の構造や機能が大きくは変わらず、そのタンパク質の構造、特にその機能が保存される。下表は、タンパク質中で元のアミノ酸と置換され得るアミノ酸と、保存アミノ酸置換と認められるアミノ酸を示している。

20

元の残基	保存的な置換	
Ala	Gly, Ser	
Arg	His, Lys	
Asn	Asp, Gln, His	
Asp	Asn, Glu	
Cys	Ala, Ser	10
Gln	Asn, Glu, His	
Glu	Asp, Gln, His	
Gly	Ala	
His	Asn, Arg, Gln, Glu	
Ile	Leu, Val	
Leu	Ile, Val	20
Lys	Arg, Gln, Glu	
Met	Leu, Ile	
Phe	His, Met, Leu, Trp, Tyr	
Ser	Cys, Thr	
Thr	Ser, Val	
Trp	Phe, Tyr	30
Tyr	His, Phe, Trp	
Val	Ile, Leu, Thr	

【0080】

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

【0081】

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

【0082】

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、ポリヌクレオチド配列の化学修飾に含まれ得る。ポリヌクレオチド誘導体は、天然分子の生物学的または免疫学的機能を少なくとも1つは保持しているポリペプチドをコードする。ポリペプチド誘導体は、グリコシル化、ポリエチレングリコール化(pegylation)、或いは任意の同様なプロセスであって、誘導起源のポリペプチドの少なくとも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って、修飾されたポリペプチドである

40

50

。

【0083】

「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或いは非共有結合するレポーター分子や酵素を指す。

【0084】

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

【0085】

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

【0086】

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

【0087】

SEQ ID NO: 21-40の断片には、固有のポリヌクレオチド配列領域が含まれる。この領域は、SEQ ID NO: 21-40を特異的に同定するものであり、例えばこの断片を得たゲノム中のその他の配列とは異なるものである。SEQ ID NO: 21-40のある断片は、例えば、ハイブリダイゼーションや増幅技術、またはSEQ ID NO: 21-40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: 21-40の断片の正確な長さ及び断片に対応するSEQ ID NO: 21-40の領域は、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

【0088】

SEQ ID NO: 1-20のある断片は、SEQ ID NO: 21-40のある断片によってコードされる。SEQ ID NO: 1-20のある断片には、SEQ ID NO: 1-20を特異的に同定する固有のアミノ酸配列領域が含まれている。例えば、SEQ ID NO: 1-20のある断片は、SEQ ID NO: 1-20を特異認識する抗体を産出するための免疫原性ペプチドとして有用である。SEQ ID NO: 1-20のある断片の正確な長さ、及びその断片に対応するSEQ ID NO: 1-20の領域は、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

【0089】

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

【0090】

10

20

30

40

50

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

【0091】

ポリヌクレオチド配列についての用語「一致率」または「一致性%」とは、標準化されたアルゴリズムを用いてアラインメントされる、2つ以上のポリヌクレオチド配列間の一致する残基の百分率のことである。このようなアルゴリズムは、2配列間のアラインメントを最適化するために比較する配列において、標準化された再現性のある方法でギャップを挿入するので、2つの配列をより有意に比較できる。

【0092】

ポリヌクレオチド配列間の一致率は、MEGALIGN version 3.12e配列
アラインメントプログラムに組込まれているようなCLUSTAL Vアルゴリズムのデ
フォルトのパラメータを用いて決定できる。このプログラムは、LASERGENE ソ
フトウェアパッケージ（一組の分子生物学的分析プログラム）（DNASTAR, Ma
dison WI）の一部である。このCLUSTAL Vは、Higgins, D. G
. 及び P. M. Sharp (1989) CABIOS 5:151-153、Hig
gins, 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2つ1組でアラインメントする際のデフォルトパラメータ
は、Ktuple=2、gap penalty=5、window=4、「diagonal saved」=4と設定する。「重み付けされた」残基重み付け表が、デフォ
ルトとして選択された。一致率は、アラインメントされたポリヌクレオチド配列間の「類
似性パーセント」としてCLUSTAL Vによって報告される。

【0093】

或いは、一般的に用いられ且つ自由に入手できる配列比較アルゴリズム一式が、国立バイ
オテクノロジー情報センター（NCBI）Basic Local Alignment
Search Tool（BLAST）から提供されており（Altschul, S. F
. 他 (1990) J. Mol. Biol. 215:403-410）、これはメリ
ーランド州ベセスダにあるNCBI及びインターネット（http://www.ncbi.nlm.nih.gov/BLAST/）を含む幾つかの情報源から入手可能である
。このBLASTソフトウェア一式には、既知のポリヌクレオチド配列と様々なデータベ
ースの別のポリヌクレオチド配列とのアラインメントに用いられる「blastn」を含
む、様々な配列分析プログラムが含まれる。「BLAST 2 Sequences」と呼
ばれるツールが入手可能であり、2つのヌクレオチド配列の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
及び blastp（以下に記載）の両方に用いることができる。BLASTプログラム
は、一般的には、ギャップ及びデフォルト設定に設定された他のパラメータと共に用いる
。例えば、2つのヌクレオチド配列を比較するために、デフォルトパラメータとして設定
された「BLAST 2 Sequences」ツールVersion 2.0.12（2
000年4月21日）を用いてblastnを実行してもよい。デフォルトパラメータの
設定例を以下に示す。

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

【0094】

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

10

【0095】

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

【0096】

ポリペプチド配列間の一致率は、MEGALIGN version 3.12e配列アラインメントプログラムに組込まれているようなCLUSTAL Vアルゴリズムのデフォルトのパラメータを用いて決定できる(既に説明したのでそれを参照されたい)。CLUSTAL Vを用いて、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と設定する。デフォルトの残基重み付け表としてPAM250マトリクスを選択する。ポリヌクレオチドアラインメントと同様に、CLUSTAL Vは、アラインメントされたポリペプチド配列対間の「類似率」として一致率を報告する。

20

【0097】

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

30

Matrix: BLOSUM62

Open Gap: 11 及び Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

40

【0098】

一致率は、ある定義された配列の全長(例えば特定のSEQ IDナンバーで定義された配列)で測定し得る。或いは、より短い長さ、例えば、定義された、より大きな配列から得られた断片(例えば少なくとも15、20、30、40、50、70、または150の連続したヌクレオチドの断片)の長さと比較して一致率を測定してもよい。ここに挙げた長さは単なる例示的なものに過ぎず、表、図及び配列リストを含めた本明細書に記載された配列に裏付けられた任意の配列長さの断片を用いて、或る長さであってその長さに対して一致率を測定し得る長さを説明し得ることを理解されたい。

【0099】

「ヒト人工染色体(HAC)」は、約6kb~10MbのサイズのDNA配列を含み得

50

る、安定した染色体複製の分離及び維持に必要な全てのエレメントを含む直鎖状の微小染色体である。

【0100】

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

【0101】

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

10

【0102】

一般に、ハイブリダイゼーションのストリンジェンシーは或る程度、洗浄ステップを実行する温度を基準にして表すことができる。このような洗浄温度は通常、所定のイオン強度及びpHにおける特異配列の融点（ T_m ）より約5～20℃低くなるように選択する。この T_m は、所定のイオン強度及びpHの条件下で、完全に一致するプローブに標的配列の50%がハイブリダイズする温度である。 T_m を計算する式及び核酸のハイブリダイゼーション条件はよく知られており、Sambrook 他（1989）Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NYに記載されており、特に2巻の9章を参照されたい。

20

【0103】

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

30

40

【0104】

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

【0105】

50

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

【0106】

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

【0107】

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

10

【0108】

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

【0109】

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。

【0110】

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

20

【0111】

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

【0112】

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

30

【0113】

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

40

【0114】

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

【0115】

「プローブ」とは、同一配列或いは対立遺伝子核酸配列、関連する核酸配列の検出に用いる、T R I C H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

50

イソトープ、リガンド、化学発光試薬及び酵素がある。「プライマー」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って標的DNA鎖に沿って延長し得る。プライマー対は、例えばポリメラーゼ連鎖反応（PCR）による核酸配列の増幅（及び同定）に用い得る。

【0116】

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

10

【0117】

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

20

【0118】

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

30

40

50

、例えばPCRまたはシーケンシングプライマーとして、マイクロアレイエレメントとして、或いは核酸のサンプルにおいて完全または部分的相補的ポリヌクレオチドを同定する特異プローブとして有用である。オリゴヌクレオチドの選択方法は、上記の方法に限定されるものではない。

【0119】

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

10

【0120】

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

【0121】

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

20

【0122】

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

【0123】

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

30

【0124】

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

【0125】

用語「特異的結合」及び「特異的に結合する」は、タンパク質若しくはペプチドと、アゴニスト、抗体、アンタゴニスト、小分子、若しくは任意の天然若しくは合成の結合組成物との間の相互作用を指す。この相互作用は、タンパク質の特定の構造（例えば抗原決定基即ちエピトープ）であって結合分子が認識するものが存在するか否かに依存していることを意味している。例えば、抗体がエピトープ「A」に対して特異的である場合、結合していない標識した「A」及び抗体を含む反応液に、エピトープAを含むポリペプチド或いは結合していない無標識の「A」が存在すると、抗体と結合する標識Aの量が減少する。

40

【0126】

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

50

【0127】

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

【0128】

用語「基板」は、任意の好適な固体或いは半固体の支持物を指し、膜及びフィルター、チップ、スライド、ウェハ、ファイバー、磁気または非磁気ビード、ゲル、チューブ、プレート、ポリマー、微小粒子、毛細管が含まれる。基板は、壁、溝、ピン、チャンネル、孔等、様々な表面形態を有することができ、基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。

【0129】

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

10

【0130】

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

20

【0131】

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

30

【0132】

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

40

50

チド配列中での変異である。多型変異配列はまた、ポリヌクレオチド配列の1つのヌクレオチドが異なる「1塩基多型性」(SNP)も含み得る。SNPの存在は、例えば特定の個体群、病状または病状性向を示し得る。

【0133】

特定のポリペプチド配列の「変異体」は、ポリペプチド配列の1本の長さ全体で特定のポリペプチド配列に対して少なくとも40%の相同性を有するポリペプチド配列として画定される。ここで、デフォルトパラメータに設定した「BLAST 2 Sequence s」ツールVersion 2.0.9(1999年5月7日)を用いてblastpを実行する。このようなポリペプチド対は、所定の長さに対して、例えば少なくとも50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の配列同一性を示し得る。 10

【0134】

(発明)

本発明は、新規のヒト輸送体及びイオンチャネル(TRICH)及びTRICHをコードするポリヌクレオチドの発見に基づき、これらの組成物を利用した輸送障害、神経の疾患、筋疾患、免疫疾患および細胞増殖異常の診断、治療、及び予防に関する。

【0135】

表1は、本発明の完全長ポリヌクレオチド配列及びポリペプチド配列の命名の概略である。各ポリヌクレオチド及びその対応するポリペプチドは、1つのIncyteプロジェクト識別番号(IncyteプロジェクトID)と相関する。各ポリペプチド配列は、ポリペプチド配列識別番号(ポリペプチドSEQ ID NO)とIncyteポリペプチド配列番号(IncyteポリペプチドID)によって表示した。各ポリヌクレオチド配列は、ポリヌクレオチド配列識別番号(ポリヌクレオチドSEQ ID NO)とIncyteポリヌクレオチドコンセンサス配列番号(IncyteポリヌクレオチドID)によって表示した。 20

【0136】

表2は、GenBankタンパク質(genpept)データベースに対するBLAST分析によって同定されたような、本発明のポリペプチドとの相同性を有する配列を示している。列1および列2はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号(ポリペプチドSEQ ID NO:)およびそれに対応するIncyteポリペプチド配列番号(IncyteポリペプチドID)を示す。列3は、GenBankの最も近い相同体のGenBankの識別番号(GenBank ID NO:)を示す。列4は、各ポリペプチドとその相同体との間の一致を表す確率スコアを示す。列5は、GenBank相同体の注釈を示し、更に該当箇所には関連する引用文も示す。これらを引用することを以って本明細書の一部とする。 30

【0137】

表3は、本発明のポリペプチドの様々な構造的特徴を示す。列1および列2はそれぞれ、本発明の各ポリペプチドのポリペプチド配列識別番号(SEQ ID NO:)およびそれに対応するIncyteポリペプチド配列番号(IncyteポリペプチドID)を示す。列3は、各ポリペプチドのアミノ酸残基数を示す。列4および列5はそれぞれ、GCG配列分析ソフトウェアパッケージのMOTIFSプログラム(Genetics Computer Group, Madison WI)によって決定された、リン酸化およびグリコシル化の可能性のある部位を示す。列6は、シグネチャ配列、ドメイン、およびモチーフを含むアミノ酸残基を示す。列7は、タンパク質の構造/機能の分析のための分析方法を示し、該当箇所にはさらに分析方法に利用した検索可能なデータベースを示す。 40

【0138】

表2及び3は共に、本発明の各々のポリペプチドの特性を要約しており、それら特性が請求の範囲に記載されたポリペプチドが輸送体及びイオンチャネルであることを確立している。例えば、SEQ ID NO: 5はショウジョウバエのナトリウム-水素交換体NHE 50

1 (GenBank ID g4894991) と 61% の同一性を有することが Basic Local Alignment Search Tool (BLAST) によって確認された。(表 2 参照)。BLAST 確率スコアは 6.0×10^{-139} であり、これは保存されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO: 5 はまた、ナトリウム／水素交換体ファミリ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確認された。(表 3 参照)。

BLIMPS 解析よりのデータは、SEQ ID NO: 5 がナトリウム／水素交換体である、さらに実証的な証拠を提供する。別の例として、SEQ ID NO: 6 はヒトのシトリン (成人発症型タイプ I シトリン血タンパク質) (GenBank ID g5052319) と約 50% 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 6.0×10^{-51} 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: 6 はまた、ミトコンドリアキャリアタンパク質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ドメインの PFAM データベースにおいて、統計的に有意な一致を検索して決定された。(表 3 参照)。BLIMPS、MOTIFS、及び PROFILESCAN 解析よりのデータは、SEQ ID NO: 6 がミトコンドリアキャリアタンパク質である、さらに実証的な証拠を提供する。別の例において、SEQ ID NO: 7 はランソウのメリビオースキャリアタンパク質 (GenBank ID g1653342) と 27% 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 1.8×10^{-16} 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。DOMO および PRODOM の解析よりの付加的な BLAST データは、SEQ ID NO: 7 がシンポータタンパク質である、さらに実証的な証拠を提供する。別の例として、SEQ ID NO: 9 はシロイヌナズナの ABC 輸送体 (GenBank ID g4262239) と 26% の同一性を有し、ヒトの sterolin-2 (GenBank ID g15146444) の残基 M1 ~ 残基 W374 と 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 確率スコアはそれぞれ 4.1×10^{-25} および 0.0 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: 9 は 2 つの膜貫通ドメインを含むことが隠れマルコフモデル (HMM) 解析によって決定され、またホワイト／スカーレット ABC 輸送体ドメインも有している。(表 3 参照)。これらのデータは、SEQ ID NO: 9 が ABC 輸送体であるという、さらに確証的な証拠を提供する。別の例として、SEQ ID NO: 12 はラットの神経グルタミン輸送体 (GenBank ID g6978016) と 93% の同一性を有することが Basic Local Alignment Search Tool (BLAST) によって確認された。(表 2 参照)。BLAST 確率スコアは 4.4×10^{-239} 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: 12 はまた、膜 7 回貫通受容体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)。

これらのデータは、SEQ ID NO: 12 がアミノ酸輸送体であるという、さらに確証的な証拠を提供する。別の例において、SEQ ID NO: 14 はマウスの多剤耐性タンパク質 (GenBank ID g387426) に 52% の同一性を有することが Basic Local Alignment Search Tool (BLAST) によって決定された。(表 2 参照)。BLAST 確率スコアは 0.0 であるが、これは観測されたポリペプチド配列アラインメントが偶然に得られる確率を示す。SEQ ID NO: 14 はまた、ABC 輸送体ドメインと ABC 輸送体膜貫通領域ドメインを有するが、

これは、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリードメインの PFAM データベースにおいて、統計的に有意な一致を検索して決定された。(表 3 参照)。BLIMPS、MOTIFS、及び PROFILESCAN 解析よりのデータは、SEQ ID NO: 14 が多剤耐性 ABC 輸送体である、さらに確証的な証拠を提供する。別の例として、SEQ ID NO: 18 はシロイヌナズナの推定膜輸送体 (GenBank ID g2289003) と 41% の同一性を有し、ヒトのプロトンマイオシトール輸送体 (GenBank ID g15211933) の残基 M20 ~ 残基 E648 と 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 確率スコアはそれぞれ 1.4×10^{-94} および 0.0 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: 18 はまた、糖 (およびその他の) 輸送体ドメインを有するが、これは、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリードメインの PFAM データベースにおいて、統計的に有意な一致を検索して決定された。(表 3 参照)。BLIMPS、MOTIFS、及び PROFILESCAN 解析よりのデータは、SEQ ID NO: 18 が糖輸送体である、さらに確証的な証拠を提供する。SEQ ID NO: 1-4、SEQ ID NO: 8、SEQ ID NO: 10-11、SEQ ID NO: 13、SEQ ID NO: 15-17 および SEQ ID NO: 19-20 は同様にして解析し、注釈付けをされた。SEQ ID NO: 1-20 の解析のためのアルゴリズム及びパラメータが表 7 で記述されている。

10

20

【0139】

表 4 に示すように、本発明の完全長ポリヌクレオチド配列は、cDNA 配列またはゲノム DNA 由来のコード (エキソン) 配列を用いて、或いはこれら 2 種類の配列を任意に組み合わせる構築した。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 (塩基対単位) を示す。列 4 は、例えば、SEQ ID NO: 21-40 を同定するため、或いは SEQ ID NO: 21-40 と関連するポリヌクレオチド配列とを区別するためのハイブリダイゼーションまたは増幅技術に有用なポリヌクレオチド配列の断片を示す。列 5 は cDNA 配列、ゲノム DNA から予想されたコード配列 (エキソン) 及び / または cDNA 及びゲノム DNA を共に有する配列集合に対応する識別番号を示している。これらの配列は、本発明の完全長ポリヌクレオチド配列を構築するのに用いた。表 4 の列 6 および列 7 はそれぞれ、列 5 の配列に対応する cDNA 配列および / 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。

30

【0140】

表 4 の列 5 の識別番号は、特に例えば Incyte cDNA とそれに対応する cDNA ライブラリを指す場合もある。例えば、6122382H1 はインサイト cDNA 配列の識別番号であり、BRAHNON05 はそれが由来する cDNA ライブラリの識別番号である。cDNA ライブラリが示されていない Incyte cDNA は、プールされている cDNA ライブラリ (例えば、72008374V1) 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、g2077361)。さらに、列 5 の識別番号は、ENSEMBL (The Sanger Centre、英国ケンブリッジ) データベースから由来した配列を同定し得る (即ち「ENST」命名を含む配列)。或いは、列 5 の識別番号は、NCBI RefSeq Nucleotide Sequence Records データベースから由来する場合もあり (即ち「NM」または「NT」の命名を含む配列)、また NCBI RefSeq Protein Sequence Records から由来する場合もある (即ち「NP」の命名を含む配列)。または列 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

40

50

結び合わせた cDNA 及び GenScan 予想エキソンの両方からなる群を意味する場合がある。例えば、FL_XXXXXX_N₁_N₂_YYYYY_N₃_N₄ はアルゴリズムが適用される配列のクラスターの識別番号が XXXXXX であり、アルゴリズムにより生成される予測の番号が YYYYY であり、(もし存在すれば) N₁, 2, 3... が解析中に手動で編集された可能性のある特定のエキソンであるような「縫合された」配列である(実施例 5 参照)。または、列 5 の識別番号は「エキソストレッチ(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、FLXXXXXXXX_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「エキソストレッチ」アルゴリズムを適用したヒトゲノム配列の GenBank 識別番号、gBBBBB は一番近い GenBank タンパク質相同体の GenBank 識別番号または NCBI RefSeq 識別番号である。(実施例 5 を参照。)ある RefSeq 配列が「エキソストレッチ」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、RefSeq 識別番号(「NM」、「NP」、または「NT」によって表される)が、GenBank 識別(即ち、gBBBBB)の代わりに使用される場合もある。

10

【0141】

或いは、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から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を列記する(実施例 4 と 5 を参照)。

20

接頭コード	解析タイプやプログラムの例
GNN, GFG, ENST	例えば、GENSCAN (Stanford University, CA, USA) または FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK)を用いたゲノム配列からのエキソン予測
GBI	手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	EST 配列のゲノムへのマッピングからの全長転写とエキソンの 予想エキソンと転写を予想するために、ゲノム位置と EST 構成 データが組み合わされる。

10

20

30

【0142】

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に示すような配列の適用範囲と重複する Incyte cDNA の適用範囲が得られたが、それに関連する Incyte cDNA 識別番号は示さなかった。

【0143】

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

40

【0144】

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

【0145】

本発明はまた、TRICHをコードするポリヌクレオチドを提供する。特定の実施例において、本発明は、TRICHをコードするSEQ ID NO: 21-40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。SEQ ID NO: 21-40のポリヌクレオチド配列は、配列表に示されているように等価RNA配列と同等の価値

50

を有しているが、窒素塩基チミンの出現はウラシルに置換され、糖のバックボーンはデオキシリボースではなくてリボースから構成されている。

【0146】

本発明はまた、T R I C H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、T R I C Hをコードするポリヌクレオチド配列と少なくとも約70%のポリヌクレオチド配列同一性、或いは少なくとも約85%のポリヌクレオチド配列同一性、更には少なくとも約95%ものポリヌクレオチド配列同一性を有する。本発明の特定の実施形態は、S E Q I D N O : 2 1 - 4 0からなる一群から選択された核酸配列と少なくとも約70%のポリヌクレオチド配列同一性、或いは少なくとも約85%のポリヌクレオチド配列同一性、更には少なくとも約95%ものポリヌクレオチド配列同一性を有するS E Q I D N O : 2 1 - 4 0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の変異配列を提供する。上記したポリヌクレオチド変異配列は何れも、T R I C Hの機能的或いは構造的特徴の少なくとも1つを有するアミノ酸配列をコードする。

【0147】

さらに、あるいは別方法において、本発明のポリヌクレオチド変異体はT R I C Hをコードするポリヌクレオチド配列のスプライス変異体である。スプライス変異体はT R I C Hをコードするポリヌクレオチド配列とかなりの相同性を有し得るが、m R N Aプロセッシング中のエキソンの異なるスプライシングによる配列ブロックの追加または削除により、通常、ポリヌクレオチドがより多くまたはより少数の塩基を有することになる。スプライス変異体はその全長についてT R I C Hをコードするポリヌクレオチド配列とのポリヌクレオチド配列相同性が約70%以下、あるいは約60%、あるいは約50%以下でありえるが、そのスプライス変異体のある部分はT R I C Hをコードするポリヌクレオチド配列のある部分とのポリヌクレオチド配列相同性が少なくとも約70%、あるいは少なくとも約85%、あるいは少なくとも95%、あるいは100%になっている。上記したスプライス変異配列は何れも、T R I C Hの機能的或いは構造的特徴の少なくとも1つを有するアミノ酸配列をコードする。

【0148】

遺伝暗号の縮重により作り出され得るT R I C Hをコードする種々のポリヌクレオチド配列には、既知の自然発生する任意の遺伝子のポリヌクレオチド配列と最小の類似性しか有しないものも含まれることを、当業者は理解するであろう。したがって本発明には、可能コドン選択に基づく組合せの選択によって産出し得るようなありとあらゆる可能性のあるポリヌクレオチド配列変異体を網羅し得る。これらの組み合わせは、天然のT R I C Hのポリヌクレオチド配列に適用される標準的なトリプレット遺伝暗号を基に作られ、全ての変異が明確に開示されているとみなす。

【0149】

T R I C Hをコードするヌクレオチド配列及びその変異配列は一般に、好適に選択されたストリンジェントな条件下で、天然のT R I C Hのヌクレオチド配列とハイブリダイズ可能であるが、非天然のコドンを含めるなどの実質的に異なった使い方のコドンを有するT R I C H或いはその誘導体をコードするヌクレオチド配列を作ること有利となり得る。宿主が特定のコドンを利用する頻度に基づいて、特定の真核宿主又は原核宿主に発生するペプチドの発現率を高めるようにコドンを選択することが可能である。コードされたアミノ酸配列を変えないで、T R I C H及びその誘導体をコードするヌクレオチド配列を実質的に変更する別の理由は、天然の配列から作られる転写物より例えば長い半減期など好ましい特性を備えるR N A転写物を作ることにある。

【0150】

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

の中に突然変異を導入することも可能である。

【0151】

更に本発明には、種々のストリンジェントな条件下で、請求項に記載されたポリヌクレオチド配列、特に、SEQ ID NO: 21-40 及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる（例えば、Wahl, G. M. 及び S. L. Berger (1987) *Methods Enzymol.* 152: 399-407、Kimmel, A. R. (1987) *Methods Enzymol.* 152: 507-511等を参照）。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。

【0152】

DNAシーケンシングの方法は当分野では公知であり、本発明のいずれの実施例もDNAシーケンシング方法を用いて実施可能である。DNAシーケンシング方法には酵素を用いることができ、例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amersham, Pharmacia Biotech, Piscataway NJ)を用いることができる。或いは、例えばELONGASE増幅システム (Life Technologies, Gaithersburg MD)において見られるように、ポリメラーゼと校正エキソヌクレアーゼを併用することができる。好適には、MICROLAB 2200液体転移システム (Hamilton, Reno, NV)、PTC 200サーマルサイクラー (MJ Research, Watertown MA) 及びABI CATALYST 800サーマルサイクラー (Applied Biosystems)等の装置を用いて配列の準備を自動化する。次に、ABI 373 或いは 377 DNAシーケンシングシステム (Applied Biosystems)、MEGABACE 1000 DNAシーケンシングシステム (Molecular Dynamics, Sunnyvale CA) または当分野でよく知られている他の方法を用いてシーケンシングを行う。結果として得られた配列を当分野でよく知られている種々のアルゴリズムを用いて分析する（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）。

【0153】

当分野で周知のPCR法をベースにした種々の方法で、部分的なヌクレオチド配列を利用して、TRICHをコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば、使用し得る方法の1つである制限部位PCR法は、ユニバーサルプライマー及びネステッドプライマーを用いてクローニングベクター内のゲノムDNAから未知の配列を増幅する方法である（例えば、Sarkar, G. (1993) *PCR Methods Applic.* 2: 318322を参照。）別の方法に逆PCR法があり、これは広範な方向に伸長させたプライマーを用いて環状化した鋳型から未知の配列を増幅する方法である。鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限酵素断片から得る（例えば、Triglia, T. 他 (1988) *Nucleic Acids Res.* 16: 8186。）第3の方法としてキャプチャPCR法があり、これはヒト及び酵母菌人工染色体DNAの既知の配列に隣接するDNA断片をPCR増幅する方法に關与している。（Lagerstrom, M. 他 (1991) *PCR Methods Applic.* 1: 111-119等を参照）。この方法では、PCRを行う前に複数の制限酵素の消化及びライゲーション反応を用いて未知の配列領域内に組換え二本鎖配列を挿入することが可能である。また、未知の配列を検索するために用い得る別の方法については当分野で知られている（Parker, J. D. 他 (1991) *Nucleic Acids Res.* 19: 3055-3060等を参照）。更

10

20

30

40

50

に、PCR、ネステッドプライマー及びPromoter Finder（商標）ライブラリ（Clontech, Palo Alto CA）を用いてゲノムDNAをウォーキングすることができる。この手順は、ライブラリをスクリーニングする必要がなく、イントロン／エキソン接合部を見付けるのに有用である。全てのPCRベースの方法に対して、市販されているソフトウェア、例えばOLIGO 4.06プライマー分析ソフトウェア（National Biosciences, Plymouth MN）或いは別の好適なプログラムを用いて、長さが約22～30ヌクレオチド、GC含有率が約50%以上、温度約68℃～72℃で鋳型に対してアニーリングするようにプライマーを設計し得る。

【0154】

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

【0155】

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

【0156】

本発明の別の実施例では、TRICH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TRICH、その断片または機能的等価物を発現させることが可能である。遺伝暗号固有の縮重により、実質的に同じ或いは機能的に等価のアミノ酸配列をコードする別のDNA配列が作られ得り、これらの配列をTRICHの発現に利用可能である。

【0157】

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

【0158】

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

10

20

30

40

50

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

10

【0159】

別の実施例によれば、TRICHをコードする配列は、当分野で周知の化学的方法を用いて、全体或いは一部が合成可能である（例えば、Caruthers, M. H. ら（1980）Nuc l. Acids Res. Symp. Ser 7: 215-223; 及びHorn, T. 他（1980）Nuc l. Acids Res. Symp. Ser. 225-232を参照）。別法として、化学的方法を用いてTRICH自体またはその断片を合成することが可能である。例えば、種々の液相または固相技術を用いてペプチド合成を行うことができる（たとえば、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: 2022-204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TRICHのアミノ酸配列または任意のその一部を用い、直接的な合成の際の変更、及び／または化学的方法を用いた他のタンパク質または任意のその一部からの配列との組み合わせにより、天然のポリペプチド配列を有するポリペプチドまたは変異体ポリペプチドを作製することが可能である。

20

【0160】

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

30

【0161】

生物学的に活性なTRICHを発現させるために、TRICHをコードするヌクレオチド配列またはその誘導体を好適な発現ベクターに挿入する。この発現ベクターは、好適な宿主に挿入されたコーディング配列の転写及び翻訳の調節に必要なエレメントを含む。これらのエレメントには、ベクター及びTRICHをコードするポリヌクレオチド配列におけるエンハンサー、構成型及び発現誘導型のプロモーター、5'及び3'の非翻訳領域などの調節配列が含まれる。このような要素は、長さ及び特異性が様々である。特定の開始シグナルによって、TRICHをコードする配列のより効果的な翻訳を達成することが可能である。このようなシグナルには、ATG開始コドンと、コザック配列などの近傍の配列が含まれる。TRICHをコードする配列及びその開始コドン、上流の調節配列が好適な発現ベクターに挿入された場合は、更なる転写調節シグナルや翻訳調節シグナルは必要なくなるであろう。しかしながら、コーディング配列或いはその断片のみが挿入された場合は、インフレームのATG開始コドンを含む外来性の翻訳調節シグナルが発現ベクターに含まれるようにすべきである。外来性の翻訳要素及び開始コドンは、様々な天然物及び合成物を起源とし得る。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である。（Scharf, D. 他（1994）Results Probl. Cell Differ. 20: 125-162. 等を参照）。

40

50

【0162】

当業者に周知の方法を用いて、TRICHをコードする配列、好適な転写及び翻訳調節エレメントを含む発現ベクターを作製することが可能である。これらの方法には、in vitro組換えDNA技術、合成技術、及びin vivo遺伝子組換え技術が含まれる（例えば、Sambrook, J. 他. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, 4章及び8章, および16-17章; およびAusubel, F. M. 他. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9章、13章および16章を参照）。 10

【0163】

種々の発現ベクター／宿主系を利用して、TRICHをコードする配列の保持及び発現が可能である。限定するものではないがこのような発現ベクター／宿主系には、組換えバクテリオファージ、プラスミドまたはコスミドDNA発現ベクターで形質転換させた細菌や、酵母菌発現ベクターで形質転換させた酵母菌や、ウイルス発現ベクター（例えばバキュロウイルス）に感染した昆虫細胞系や、ウイルス発現ベクター（例えばカリフラワーモザイクウイルス、CaMVまたはタバコモザイクウイルスTMV）または細菌発現ベクター（例えばTiまたはpBR322プラスミド）で形質転換させた植物細胞系、動物細胞系などの微生物等がある。（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) EMBOJ. 6:307-311、; 『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, 191-196ページ、Logan, J. および T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659、Harrington, J. J. 他 (1997) Nat. Genet. 15:345-355等を参照）レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ輸送することができる（Di Nicola, M. 他 (1998) Cancer Gen. Ther. 5(6):350-356、Yu, M. 他 (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344、Buller, R. M. 他 (1985) Nature 317(6040):813-815; McGregor, D. P. 他 (1994) Mol. Immunol. 31(3):219-226、Verma, I. M. および N. Somia (1997) Nature 389:239-242等を参照）。本発明は使用される宿主細胞によって限定されるものではない。 20 30 40

【0164】

細菌系では、多数のクローニングベクター及び発現ベクターが、TRICHをコードするポリヌクレオチド配列の使用目的に応じて選択可能である。例えば、TRICHをコードするポリヌクレオチド配列の日常的なクローニング、サブクローニング、増殖には、PBUESCRIPT (Stratagene, La Jolla CA) またはpSPORT1プラスミド (GIBCO BRL) などの多機能の大腸菌ベクターを用いることができる。ベクターの多数のクローニング部位にTRICH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 40 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. 264:55035509を参照）。例えば、抗体の産生のためなどに多量のTRICHが必要な場合は、TRICH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。

【0165】

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

【0166】

植物系もTRICHの発現に使用可能である。TRICH 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る用いられるようなCaMV由来の35S及び19Sプロモーターによって促進される。或いは、RUBISCOの小サブユニット等の植物プロモーターまたは熱ショックプロモーターを用いてもよい（例えば、Coruzzi, G. 他 (1984) EMBO J. 3:1671-1680; Broglie, R. 他 (1984) Science 224:838-843; および Winter, J. 他 (1991) Results Probl. Cell Differ. 17:85-105を参照）。これらの構成物は、直接DNA形質転換または病原体を媒介とする形質移入によって、植物細胞内に導入可能である。（『マグローヒル科学技術年鑑』（The McGraw Hill Yearbook of Science and Technology）(1992) McGraw Hill New York NY, 191-196ページ等を参照）。

【0167】

哺乳動物細胞においては、多数のウイルスベースの発現系を利用し得る。アデノウイルスが発現ベクターとして用いられる場合、後発プロモーター及び3連リーダー配列からなるアデノウイルス転写物／翻訳複合体にTRICHをコードする配列を結合し得る。アデノウイルスゲノムの非必須のE1またはE3領域へ挿入することにより、宿主細胞でTRICH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 81:36553659を参照）。更に、ラウス肉腫ウイルス(RSV)エンハンサー等の転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させ得る。SV40またはEBVをベースにしたベクターを用いてタンパク質を高レベルで発現させることもできる。

【0168】

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

【0169】

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

る配列を株化細胞に形質転換することが可能である。このような発現ベクターは、ウイルス起源の複製及び／または内在性の発現要素や、同じ或いは別のベクターの上の選択マーカー遺伝子を含む。ベクターの導入後、選択培地に移す前に強化培地で約1～2日間細胞を増殖させることができる。選択可能マーカーの目的は選択培地への抵抗性を与えることであり、選択可能マーカーが存在することにより、導入された配列をうまく発現するような細胞の成長及び回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に適した組織培養技術を用いて増殖可能である。

【0170】

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チミジンキナーゼ遺伝子と、 apr^{-} 細胞のために用いられる単純ヘルペスウイルスのアデニンホスホリボシルトランスフェラーゼ遺伝子がある（例えば、Wigler, M. 他 (1977) Cell 11:223-232; 及びLowy, I. 他 (1980) Cell 22:817-823を参照）。また、選択の基礎として代謝拮抗物質、抗生物質或いは除草剤への耐性を用いることができる。例えば $d h f r$ はメトトレキセートに対する耐性を与え、 neo はアミノグリコシッドネオマイシン及びG-418に対する耐性を与え、 als はクロルスルフロンに対する耐性を、 pat はホスフィノトリシンアセチルトランスフェラーゼに対する耐性を各々与える（Wigler, M. 他 (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbergarapin, F. 他 (1981) J. Mol. Biol. 150:114等を参照）。この他の選択可能な遺伝子、例えば、代謝のための細胞の必要条件を変える $trpB$ 及び $hisD$ は、文献に記載されている（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、例えばアントシアニン、緑色蛍光タンパク質（GFP; Clontech）、 β グルクロニダーゼ及びその基質 β グルクロニド、またはルシフェラーゼ及びその基質ルシフェリン等を用いてもよい。これらのマーカーを用いて、トランスフォーマントを特定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することが可能である（Rhodes, C. A. (1995) Methods Mol. Biol. 55:121-131等を参照）。

【0171】

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

【0172】

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

【0173】

特異的なポリクローナル抗体またはモノクローナル抗体のどちらかを用いるTRICHの発現の検出及び計測のための免疫学的方法は、当分野で周知である。このような技法には、酵素に結合したイムノソルベントアッセイ（ELISA）、ラジオイムノアッセイ（RIA）、蛍光標示式細胞分取器（FACS）などがある。TRICH上の2つの非干

渉エピトープに反応するモノクローナル抗体を用いた、2部位のモノクローナルベース免疫アッセイ (two-site, monoclonal-based immunoassay) が好ましいが、競合の結合アッセイも用いることもできる。これらのアッセイ及びこれ以外のアッセイは、当分野で公知である (Hampton, R. 他 (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Sect. IV, Coligan, J. E. 他 (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY, Pound, J. D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ等を参照)。

【0174】

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

【0175】

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

【0176】

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

【0177】

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

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

【0178】

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

【0179】

本発明のTRICHまたはその断片を用いて、TRICH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、TRICH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。

【0180】

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TRICHの天然のリガンド、または天然の基質、構造的または機能的な擬態性または自然結合パートナーに密接に関連している (Coligan, J. E. 他 (1991) Current Protocols in Immunology 1 (2) の5章等を参照)。

同様に、化合物は、TRICHが結合する天然受容体、或いは例えばリガンド結合部位などの少なくとも受容体のある断片に密接に関連し得る。何れの場合も、既知の技術を用いてこの化合物を合理的に設計することができる。一実施例では、このような化合物に対するスクリーニングには、分泌タンパク質或いは細胞膜上のタンパク質の何れか一方としてTRICH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。TRICHを発現する細胞またはTRICHを含有する細胞膜断片を試験化合物と接触させて、TRICHまたは化合物の何れかの結合、刺激または阻害を分析する。

【0181】

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

【0182】

本発明のTRICHまたはその断片を用いて、TRICHの活性を調整する化合物をスクリーニングすることが可能である。このような化合物には、アゴニスト、アンタゴニスト

、或るいは部分的または逆アゴニスト等が含まれる。一実施例では、T R I C Hの活性が許容される条件下でアッセイを実施し、そのアッセイでは少なくとも1つの試験化合物をT R I C Hと混合し、試験化合物の存在下でT R I C Hの活性を試験化合物不在下でのT R I C Hの活性と比較する。試験化合物の存在下でのT R I C Hの活性の変化は、T R I C Hの活性を調整する化合物の存在を示唆する。別法では、試験化合物をT R I C Hの活性に適した条件下でT R I C Hを含む *i n v i t r o* または細胞遊離系と結合させてアッセイを実施する。これらアッセイの何れかにおいて、T R I C Hの活性を調整する試験化合物は間接的に結合することが可能であり、試験化合物と直接接触する必要がない。少なくとも1つから複数の試験化合物をスクリーニングすることができる。

【0183】

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

【0184】

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

【0185】

T R I C Hをコードするポリヌクレオチドを用いて、ヒト疾患をモデルとした「ノックイン」ヒト化動物（ブタ）または遺伝子組換え動物（マウスまたはラット）を作製することが可能である。ノックイン技術を用いて、T R I C Hをコードするポリヌクレオチドの或る領域を動物E S細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について研究し、可能性のある医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばT R I C Hを乳汁内に分泌するなどT R I C H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（*J a n n e*, J. 他 (1998) *B i o t e c h n o l . A n n u . R e v .* 4:55-74）。

【0186】

（治療）

T R I C Hのある領域と輸送体及びイオンチャネルのある領域との間に、例えば配列及びモチーフの文脈における化学的及び構造的類似性が存在する。更に、T R I C Hの発現は、脾臓腫瘍組織、食道腫瘍組織、脳腫瘍組織、心房組織の粘液腫等の腫瘍組織および腎臓、肝臓、鼻ポリープ、前立腺、甲状腺、臍帯血、神経、消化、子宮内膜組織、正常脳組織（線条体、淡蒼球、被殻組織等）等の正常組織に密接に関連する。従って、T R I C Hは

10

20

30

40

50

、輸送障害、神経の疾患、筋疾患、及び免疫疾患や細胞異常増殖においてある役割を果たすと考えられる。T R I C Hの発現若しくは活性の増大に関連する疾患の治療においては、T R I C Hの発現または活性を低下させることが望ましい。また、T R I C H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 R I C Hの発現または活性を増大させることが望ましい。

【0187】

従って、一実施例において、T R I C H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 R I C H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、このような疾患のうち、輸送障害には運動不能症、筋萎縮性側索硬化症、毛細血管拡張性運動失調、ベッカー筋ジストロフィー、顔面麻痺、シャルコー-マリー-トゥス病、糖尿病、尿崩症、糖尿病性ニューロパシー、デュシェンヌ型筋ジストロフィー、高カリウム血性周期性四肢麻痺、正常カリウム血性周期性四肢麻痺、パーキンソン病、悪性高熱、多剤耐性、重症筋無力症、筋緊張性異栄養症、緊張病、錐体外路性終末欠陥症候群、ジストニー、末梢神経疾患、脳性腫瘍、前立腺癌と、口峡炎、徐脈型不整脈、頻拍性型不整脈、高血圧症、遺伝性QT延長症候群、心筋炎、心筋症、ネマリニンミオパシーラネマリン筋障害、中心核ミオパシー、脂質ミオパシー、ミトコンドリアミオパシー、甲状腺中毒性ミオパシー、エタノールミオパシー、皮膚筋炎、封入体筋炎、感染性筋炎、及び多発性筋炎などの輸送に関連した心臓病と、アルツハイマー病、健忘症、双極性障害、痴呆、うつ病、てんかん、トゥーレット病、妄想性精神病、及び分裂病などの輸送に関連した神経障害と、神経線維腫症、帯状疱疹後神経痛、3又神経ニューロパシー、サルコイドーシス、鎌状赤血球性貧血、ウィルソン病、白内障、不妊症、肺動脈狭窄症、常染色体性感音性難聴、高/低血糖症、グレーブス病、甲状腺腫、クッシング病と、副腎機能不全、グルコース-ガラクトース吸収不全症候群、グリコーゲン貯蔵疾患、高コレステロール血症、副腎性白質ジストロフィー、ツェルヴェーガー症候群、メンケス病、後角症候群、フォンギルケ症候群、偽性アルドステロン減少症（タイプ1）、リドル症候群、シスチン尿症、イミノグリシン尿症、Hartup病、ファンコニ病およびBartter病が含まれ、神経の疾患の中には、癲癇、虚血性脳血管障害、脳卒中、大脳新生物、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病及びその他の錐体外路障害、筋萎縮性側索硬化及びその他の運動ニューロン障害、進行性神経性筋萎縮症、色素性網膜炎、遺伝性運動失調、多発性硬化症及び他の脱髄疾患、細菌性及びウイルス性髄膜炎、脳膿瘍、硬膜下蓄膿症、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎及び神経根炎、ウイルス性中枢神経系疾患と、クールー及びクロイツフェルト-ヤコブ病、ゲルストマン症候群、Gerstmann-Straussler-Scheinker症候群を含むプリオン病（prion disease）と、致死性家族性不眠症、神経系性栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜血管芽腫（cerebellar retinal hemangioblastomatosis）、脳3又神経血管症候群、ダウン症を含む中枢神経系性精神薄弱及び他の発生障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィー及び他の神経筋障害、末梢神経疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性、及び中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分性及び不安性精神障害、及び妄想性精神病と、季節性の感情の障害（SAD）、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、片麻痺性片頭痛、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、帯状疱疹後神経痛、及びトゥーレット病が含まれ、筋疾患の中には、心筋症、心筋炎、デュシェンヌ型筋ジストロフィー、ベッカー型偽肥大性筋ジストロフィー、筋緊張性ジストロフィー、中心コア病、ネマリニンミオパシーラネマリン筋障害、中心核ミオパシー、脂質ミオパシー、ミトコンドリアミオパシー、感染性筋炎、多発性筋炎、皮膚筋炎、封入体筋炎、甲状腺中毒性ミオパシー、エタノールミオパシー（ethanol myopathy）、口峡炎、アナフィラキシー、不整脈、喘息、心血管ショック、クッシング病、高血圧症、低血糖症、心筋梗塞、片頭痛、クロム親和細胞腫、脳症、てんかん、カーンズ-セイヤ症候群、乳酸アシドーシス、ミオクロノス疾患、眼筋麻痺、酸性マルターゼ欠損症（AM 50

D、ポンペ病としても知られる)、全身性筋緊張症および遺伝性筋緊張症を含む筋障害が含まれ、免疫疾患の中には、後天性免疫不全症候群(AIDS)及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、腓炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷、および細胞異常増殖には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病(MCTD)、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌が含まれる。

10

【0188】

20

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TRICH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、TRICH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。

【0189】

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TRICH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TRICH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。

【0190】

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TRICH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、TRICH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。

30

【0191】

更なる実施例では、TRICH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TRICH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、上記した輸送障害、神経疾患、筋疾患、免疫疾患および細胞の増殖異常が含まれる。一実施態様では、TRICH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TRICHを発現する細胞または組織に薬剤を運ぶターゲティング或いは運搬機構として間接的に用いられ得る。

【0192】

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TRICH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、TRICHをコードするポリヌクレオチドの相補配列を発現するベクターを患者に投与することも可能である。

40

【0193】

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

【0194】

50

TRICHのアンタゴニストは、当分野で一般的な方法を用いて製造することが可能である。詳しくは、精製されたTRICHを用いて抗体を作ったり、治療薬のライブラリをスクリーニングしてTRICHと特異的に結合するものを同定が可能である。TRICHの抗体も、当分野で一般的な方法を用いて製造することが可能である。このような抗体には、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、一本鎖、Fabフラグメント、及びFab発現ライブラリによって作られたフラグメントが含まれる。但し、これらに限定されるものではない。中和抗体（即ち二量体の形成を阻害する抗体）は通常、治療用に好適である。

【0195】

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

【0196】

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

【0197】

TRICHに対するモノクローナル抗体は、培地内の連続した細胞株によって、抗体分子を産生する任意の技術を用いて作製することが可能である。限定するものではないがこのような技術には、ハイブリドーマ技術、ヒトB細胞ハイブリドーマ技術及び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等を参照）。

【0198】

更に、「キメラ抗体」作製のために発達したヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの技術が、好適な抗原特異性及び生物学的活性を備える分子を得るために用いられる（例えば、Morrison, S. L. 他（1984）Proc. Natl. Acad. Sci. 81: 6851-6855; Neuberger, M. S. 他（1984）Nature 312: 604-608; Takeda, S. 他（1985）Nature 314: 452, 454を参照。）別法では、当分野で周知の方法を用いて、一本鎖抗体の産生のための記載された技術を適用して、TRICH特異性一本鎖抗体を生成する。関連特異性を有するがイディオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリからチェーンシャッフリングによって産生することもできる（Burton D. R.（1991）Proc. Natl. Acad. Sci. USA 88: 10134-10137等を参照）。

【0199】

抗体の産生は、リンパ球集団における*in vivo*産生の誘導によって、或いは免疫グロブリンライブラリのスクリーニングまたは文献に開示されているような高特異結合試薬のパネルのスクリーニングによっても行い得る（Orlandi, R. 他（1989）

10

20

30

40

50

) Proc. Natl. Acad. Sci. USA 86: 3833-3837、Winter, G. 他 (1991) Nature 349: 293-299等を参照)。

【0200】

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

【0201】

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

【0202】

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、TRICHに対する抗体の親和性を評価する。親和性を結合定数K_aで表すが、このK_aは、平衡状態の下でTRICH抗体複合体のモル濃度を遊離抗体と遊離抗原のモル濃度で除して得られる値である。多数のTRICHエピトープに対して親和性が不均一なポリクローナル抗体医薬のK_aは、TRICHに対する抗体の平均親和性または結合活性を表す。特定のTRICHエピトープに単一特異的なモノクローナル抗体医薬のK_aは、親和性の真の測定値を表す。K_a値が10⁹~10¹² L/molの高親和性抗体医薬は、TRICH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10⁶~10⁷ L/molの低親和性抗体医薬は、TRICHが抗体から最終的に活性化状態で解離する必要がある免疫精製 (immunopurification) 及び類似の処理に用いるのが好ましい。 (Catty, D. (1988) Antibodies, Volume I: A Practical Approach. IRL Press, Washington, DC; Liddell, J. E. およびCryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY)。

【0203】

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

【0204】

本発明の別の実施例では、TRICHをコードするポリヌクレオチド、またはその任意の断片や相補配列が、治療目的で 사용할 ことができる。ある実施態様では、TRICHをコードする遺伝子のコーディング領域や調節領域に相補的な配列やアンチセンス分子 (DNA及びRNA、修飾ヌクレオチド) を設計して遺伝子発現を変更することができる。こ

のような技術は当分野では周知であり、センスまたはアンチセンスオリゴヌクレオチドまたは大きな断片が、TRICHをコードする配列の制御領域から、またはコード領域に沿ったさまざまな位置から設計可能である。(Agrawal, S., 編集 (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJを参照。)

【0205】

治療に用いる場合、アンチセンス配列を好適な標的細胞に導入するのに好適な任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を作製する発現プラスミドの形で細胞内に輸送することが可能である(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関連ウイルスベクター等のウイルスベクターを用いて細胞内に導入することもできる(Miller, A. D. (1990) *Blood* 76: 271、前出の Ausubel、Uckert, W. および W. Walther (1994) *Pharmacol. Ther.* 63 (3): 323-347等を参照)。その他の遺伝送達機構には、リポソーム系、人工的なウイルスエンベロップ及び当分野で公知のその他の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. 等を参照)。

【0206】

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

【0207】

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

10

【0208】

限定するものではないがTRICHの発現に影響を及ぼし得る発現ベクターには、PCDNA 3.1、EPITAG、PRCCMV2、PREP、PVAX、PCR2-TOPOTAベクター(Invitrogen, Carlsbad CA)、PCMV-SCRIPT、PCMV-TAG、PEGSH/PERV(Stratagene, La Jolla CA) 及び PTET-OFF、PTET-ON、PTRE2、PTRE2-LUC、PTK-HYG(Clontech, Palo Alto CA) がある。ECMCADを発現させるために、(i) 恒常的に活性なプロモーター(例えば、サイトメガロウイルス(CMV)、ラウス肉腫ウイルス(RSV)、SV40ウイルス、チミジンキナーゼ(TK)、若しくは β -アクチン遺伝子等)、(ii) 誘導性プロモーター(例えば、市販されているT-REXプラスミド(Invitrogen)に含まれている、テトラサイクリン調節性プロモーター(Gossen, M. および H. Bujard (1992) *Proc. Natl. Acad. Sci. U. S. A.* 89:5547-5551; Gossen, M. 他 (1995) *Science* 268:1766-1769; Rossi, F. M. V. および H. M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456))、エクジソン誘導性プロモーター(市販されているプラスミドPVGRXR及びPINDに含まれている: Invitrogen)、FK506/ラパマイシン誘導性プロモーター、またはRU486/ミフェプリストーン誘導性プロモーター(Rossi, F. M. V. および H. M. Blau, 前出)、または(iii) 正常な個体に由来するECMCADをコードする内在性遺伝子の天然のプロモーター若しくは組織特異的プロモーターを用いることが可能である。

20

30

【0209】

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

40

【0210】

本発明の別の実施例では、TRICHの発現に関連する遺伝子欠損によって起こる疾患や異常症は、(i) レトロウイルス末端反復配列(LTR)プロモーター若しくは独立したプロモーターのコントロール下でTRICHをコードするポリヌクレオチドと、(ii) 好適なRNAパッケージングシグナルと、(iii) 追加のレトロウイルス・シス作用性RNA配列及び効率的なベクターの増殖に必要なコーディング配列を伴うRev応答性エレメント(RRE)とからなるレトロウイルスベクターを作製して治療することができる。レトロウイルスベクター(例えばPFB及びPFBNEO)はStratagene社から市販されており、刊行データ(Riviere, I. 他 (1995) *Proc.*

50

Natl. Acad. Sci. U. S. A. 92:6733-6737)に基づいている。上記データを引用することをもって本明細書の一部とする。ベクターは、好適なベクター産生細胞株(VPCL)において増殖され、VPCLは、標的細胞上の受容体に対する親和性を有するエンベロップ遺伝子またはVSVg等の汎親和性エンベロップタンパク質を発現する(Armentano, D. 他(1987) J. Virol. 61:1647-1650、Bender, M. A. 他(1987) J. Virol. 61:1639-1646、Adam, M. A. および A. D. Miller(1988) J. Virol. 62:3802-3806、Dull, T. 他(1998) J. Virol. 72:8463-8471、Zufferey, R. 他(1998) J. Virol. 72:9873-9880)。RIGGに付与された米 10
国特許第5,910,434号(「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」)において、レトロウイルスパッケージング細胞系を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、細胞集団(例えばCD4⁺ T細胞)の形質導入、及び形質導入した細胞の患者への戻しは、遺伝子治療の分野では当業者に公知の方法であり、多数の文献に記載されている(Ranga, U. 他(1997) J. Virol. 71:7020-7029、Bauer, G. 他(1997) Blood 89:2259-2267、Bonyhadi, M. L. (1997) J. Virol. 71:4707-4716、Ra 20
nga, U. 他(1998) Proc. Natl. Acad. Sci. U. S. A. 95:1201-1206、Su, L. (1997) Blood 89:2283-2290)。

【0211】

別法では、アデノウイルス系遺伝子治療の送達系を用いて、TRICHの発現に関連する1或いは複数の遺伝子異常を有する細胞にTRICHをコードするポリヌクレオチドを送達する。アデノウイルス系ベクターの作製及びパッケージングについては、当業者に公知である。複製欠損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を脾臓の無損傷の脾島内に導入するために可変性であることが証明された(Csete, M. E. 他(1995) Transplantation 27:263-268) 30
。使用できる可能性のあるアデノウイルスベクターは、Armentanoに付与された米国特許第5,707,618号(「Adenovirus vectors for gene therapy」)に記載されており、引用することをもって本明細書の一部とする。アデノウイルスベクターについては、Antinuzzi, P. A. 他(1999) Annu. Rev. Nutr. 19:511-544 及び Verma, I. M. 及び N. Somia (1997) Nature 18:389:239-242も参照されたい。両文献は、引用することをもって本明細書の一部とする。

【0212】

別法では、ヘルペス系遺伝子治療の送達系を用いて、TRICHの発現に関連する1或いは複数の遺伝子異常を有する標的細胞にTRICHをコードするポリヌクレオチドを送達 40
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TRICHを導入する際に特に有要である。ヘルペス系ベクターの作製及びパッケージングは、当業者に公知である。複製適格性単純ヘルペスウイルス(HSV) I型系のベクターは、レポーター遺伝子を霊長類の眼に送達するために用いられてきた(Liu, X. 他(1999) Exp. Eye Res. 169:385-395)。HSV-1ウイルスベクターの作製についても、DeLucaに付与された米国特許第5,804,413号(「Herpes simplex virus swains for gene transfer」)が開示されており、該特許の引用をもって本明細書の一部とする。米 50
国特許第5,804,413号には、ヒト遺伝子治療を含む目的のために好適なプロモーターの制御下において細胞に導入される少なくとも1つの外在性遺伝子を有するゲノムを

含む組換えHSV d92についての記載がある。上記特許はまた、ICP4、ICP27及びICP22のために除去される組換えHSV系統の作製及び使用について開示している。HSVベクターについては、Goins, W. F. 他 (1999) J. Virol. 73:519-532 及び Xu, H. 他 (1994) Dev. Biol. 163:152-161も参照されたい。両文献は、引用をもって本明細書の一部とする。クローン化ヘルペスウイルス配列の操作、巨大ヘルペスウイルスのゲノムの異なった部分を含む多数のプラスミドを形質移入した後の組換えウイルスの産生、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は、当業者に公知の技術である。

【0213】

別法では、 α ウイルス（正の一本鎖RNAウイルス）ベクターを用いてTRICHをコードするポリヌクレオチドを標的細胞に送達する。プロトタイプの α ウイルスであるセムリキ森林熱ウイルス（Semliki Forest Virus, SFV）の生物学的研究が広範に行われており、遺伝子導入ベクターがSFVゲノムに基づいていることが分かった（Garoff, H. 及び K. -J. Li (1998) Curr. Opin. Biotech. 9:464-469）。 α ウイルスRNAの複製中に、通常はウイルスのキャプシッドタンパク質をコードするサブゲノムRNAが作り出される。このサブゲノムRNAは、完全長のゲノムRNAより高いレベルに複製されるため、酵素活性（例えばプロテアーゼ及びポリメラーゼ）を有するウイルスタンパク質に比べてキャプシッドタンパク質が過剰産生される。同様に、TRICHをコードする配列を α ウイルスゲノムのカ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TRICHをコードするRNAが産生され、高いレベルでTRICHが合成される。通常は α ウイルスの感染が数日以内での細胞溶解に関係する一方で、シンドビスウイルス（SIN）の変異体を有するハムスター正常腎臓細胞（BHK-21）の持続的な感染を確立する能力は、 α ウイルスの溶解複製を遺伝子治療に適用できるように好適に変更可能であることを示唆している（Dryga, S. A. 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TRICHを導入することできる。或る集団におけるサブセットの細胞の特定形質導入は、形質導入前に細胞の選別を必要とし得る。 α ウイルスの感染性cDNAクローンの処置方法、 α ウイルスのcDNA及びRNAの形質移入方法及び α ウイルスの感染方法は、当業者に公知である。

【0214】

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

【0215】

リボザイムは酵素性RNA分子であり、RNAの特異的切断を触媒するためにリボザイムを用いることもできる。リボザイム作用のメカニズムは、相補的標的RNAへのリボザイム分子の配列特異性ハイブリダイゼーションとその後に起こる内ヌクレオチド鎖切断に関与している。例えば、遺伝子操作で作られたハンマーヘッド型リボザイム分子が、TRICH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。

10

20

30

40

50

【0216】

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

【0217】

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

10

【0218】

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

20

【0219】

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

30

【0220】

特異ポリヌクレオチドの変異発現における有効性に対して、少なくとも1個から複数個の試験化合物をスクリーニングし得る。試験化合物は、当分野で通常知られている任意の方法により得られる。このような方法には、ポリヌクレオチドの発現を変異させる場合と、既存の、市販のまたは専売の、天然または非天然の化合物ライブラリから選択する場合と、標的ポリヌクレオチドの化学的及び／または構造的特性に基づく化合物を合理的にデザインする場合と、組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合に有効であることが知られているような化合物の化学修飾がある。TRICHをコードするポリヌクレオチドを含むサンプルは、このようにして得られた試験化合物の少なくとも1つに曝露する。サンプルには例えば、無傷細胞、透過化処理した細胞、無細胞再構成系または再構成生化学系があり得る。TRICHをコードするポリヌクレオチドの発現に

40

50

おける変化は、当分野で周知の任意の方法でアッセイする。通常、T R I C H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し、それによって1つ若しくは複数の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較に対する基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現における変化の検出は、ポリヌクレオチドの発現を変異する際に試験化合物が有効であることを示している。特異ポリヌクレオチドの変異発現に有効な化合物に対して、例えば分裂酵母 (Schizosaccharomyces pombe) 遺伝子発現系 (Atkins, D. 他 (1999) 米国特許第5, 932, 435号、Arndt, G. M. 他 (2000) *Nucleic Acids Res.* 28: E15) またはHeLa細胞等のヒト細胞系 (Clarke, M. L. 他 (2000) *Biochem. Biophys. Res. Commun.* 268: 8-13) を用いてスクリーニングを実行する。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性のためのオリゴヌクレオチド (デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾オリゴヌクレオチド) の組み合わせライブラリをスクリーニングすることに関与している (Bruce, T. W. 他 (1997) 米国特許第5, 686, 242号、Bruce, T. W. 他 (2000) 米国特許第6, 022, 691号)。

10

【0221】

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

20

【0222】

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

【0223】

本発明の追加実施例は、通常薬剤として許容できる賦形剤で処方される活性成分を有する成分の投与に関連する。賦形剤には例えば、糖、でんぷん、セルロース、ゴム及びタンパク質がある。様々な処方通常知られており、詳細はRemington's Pharmaceutical Sciences (Maack Publishing, Easton PA) の最新版に記載されている。このような組成物は、T R I C H、T R I C Hの抗体、擬態、アゴニスト、アンタゴニスト、またはT R I C Hのインヒビターなどかなる。

30

【0224】

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

40

【0225】

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

50

【0226】

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

【0227】

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

【0228】

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

【0229】

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

【0230】

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

【0231】

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

【0232】

(診断)

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

TRICHの診断アッセイには、抗体及び標識を用いてヒトの体液或いは細胞や組織から採取されたものからTRICHを検出する方法が含まれる。この抗体は修飾されたものもされていないものも可能であり、レポーター分子との共有結合または非共有結合で標識化できる。レポーター分子としては広くさまざまな種類が本分野で知られており、また使用可能であるが、そのうちのいくつかは上記で説明されている。

【0233】

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

10

【0234】

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

20

【0235】

一実施形態では、TRICHまたは近縁の分子をコードする遺伝子配列を含むポリヌクレオチド配列を検出可能なPCRプローブを用いたハイブリダイゼーションによって、TRICHをコードする核酸配列を同定することが可能である。例えば5'調節領域である高度に特異的な領域か、例えば保存されたモチーフであるやや特異性の低い領域から作られているかのプローブの特異性と、ハイブリダイゼーション或いは増幅のストリンジェントは、プローブがTRICHをコードする自然界の配列のみを同定するかどうか、或いは対立遺伝子や関連配列コードする自然界の配列のみを同定するかどうかによって決まるであろう。

30

【0236】

プローブはまた、関連する配列の検出に利用され、TRICHをコードする任意の配列と少なくとも50%の配列同一性を有し得る。目的の本発明のハイブリダイゼーションプローブには、DNAあるいはRNAが可能であり、SEQ ID NO: 21-40の配列、或いはTRICH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。

【0237】

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

40

【0238】

TRICHをコードするポリヌクレオチド配列を用いて、TRICHの発現に関連する疾患を診断することが可能である。限定するものではないが、このような疾患のうち、輸送障害には運動不能症、筋萎縮性側索硬化症、毛細血管拡張性運動失調、ベッカー筋ジスト

50

ロフィー、顔面麻痺、シャルコー-マリー-トゥース病、糖尿病、尿崩症、糖尿病性ニューロパシー、デュシェンヌ型筋ジストロフィー、高カリウム血性周期性四肢麻痺、正常カリウム血性周期性四肢麻痺、パーキンソン病、悪性高熱、多剤耐性、重症筋無力症、筋緊張性異栄養症、緊張病、錐体外路性終末欠陥症候群、ジストニー、末梢神経疾患、脳性腫瘍、前立腺癌と、口峡炎、徐脈型不整脈、頻拍性型不整脈、高血圧症、遺伝性QT延長症候群、心筋炎、心筋症、ネマリンミオパシーラネマリン筋障害、中心核ミオパシー、脂質ミオパシー、ミトコンドリアミオパシー、甲状腺中毒性ミオパシー、エタノールミオパシー、皮膚筋炎、封入体筋炎、感染性節炎、及び多発性筋炎などの輸送に関連した心臓病と、アルツハイマー病、健忘症、双極性障害、痴呆、うつ病、てんかん、トゥーレット病、妄想性精神病、及び分裂病などの輸送に関連した神経障害と、神経線維腫症、帯状疱疹後神経痛、3叉神経ニューロパシー、サルコイドーシス、鎌状赤血球性貧血、ウィルソン病、白内障、不妊症、肺動脈狭窄症、常染色体性感音性難聴、高/低血糖症、グレーブス病、甲状腺腫、クッシング病と、副腎機能不全、グルコース-ガラクトース吸収不全症候群、グリコーゲン貯蔵疾患、高コレステロール血症、副腎性白質ジストロフィー、ツェルヴェーガー症候群、メンケス病、後角症候群、フォンギルケ症候群、偽性アルドステロン減少症（タイプ1）、リドル症候群、シスチン尿症、イミノグリシン尿症、Hartup病、ファンコニ病およびBartter病が含まれ、神経の疾患の中には、癲癇、虚血性脳血管障害、脳卒中、大脳新生物、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病及びその他の錐体外路障害、筋萎縮性側索硬化及びその他の運動ニューロン障害、進行性神経性筋萎縮症、色素性網膜炎、遺伝性運動失調、多発性硬化症及び他の脱髄疾患、細菌性及びウイルス性髄膜炎、脳膿瘍、硬膜下蓄膿症、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎及び神経根炎、ウイルス性中枢神経系疾患と、クールー及びクロイツフェルト-ヤコブ病、ゲルストマン症候群、Gerstmann-Strausler-Scheinker症候群を含むプリオン病（prion disease）と、致死性家族性不眠症、神経系性栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜血管芽腫（cerebelloretinal hemangioblastomatosis）、脳3叉神経血管症候群、ダウン症を含む中枢神経系性精神薄弱及び他の発生障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィー及び他の神経筋障害、末梢神経疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性、及び中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分性及び不安性精神障害、及び妄想性精神病と、季節性の感情の障害（SAD）、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、片麻痺性片頭痛、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、帯状疱疹後神経痛、及びトゥーレット病が含まれ、筋疾患の中には、心筋症、心筋炎、デュシェンヌ型筋ジストロフィー、ベッカー型偽肥大性筋ジストロフィー、筋緊張性ジストロフィー、中心コア病、ネマリンミオパシーラネマリン筋障害、中心核ミオパシー、脂質ミオパシー、ミトコンドリアミオパシー、感染性節炎、多発性筋炎、皮膚筋炎、封入体筋炎、甲状腺中毒性ミオパシー、エタノールミオパシー（ethanol myopathy）、口峡炎、アナフィラキシー、不整脈、喘息、心血管ショック、クッシング病、高血圧症、低血糖症、心筋梗塞、片頭痛、クロム親和細胞腫、脳症、てんかん、カーンズ-セイヤ症候群、乳酸アシドーシス、ミオクロヌス疾患、眼筋麻痺、酸性マルターゼ欠損症（AMD、ポンペ病としても知られる）、全身性筋緊張症および遺伝性筋緊張症を含む筋障害が含まれ、免疫疾患の中には、後天性免疫不全症候群（AIDS）及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー（APCED）、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時的リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、肺炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身

性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷、および細胞異常増殖には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病（MCTD）、骨髓線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髓腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髓、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌が含まれる。TRICHをコードするポリヌクレオチド配列は、サザン法やノーザン法、ドットブロット法、或いはその他の膜系の技術、PCR法、ディップスティック（dipstick）、ピン（pin）、ELISA式アッセイ、及び変異TRICHの発現を検出するために患者から採取した体液或いは組織を利用するマイクロアレイに使用することが可能である。このような定性方法または定量方法は、当分野で公知である。 10

【0239】

ある実施態様では、TRICHをコードするヌクレオチド配列は、関連する疾患、特に上記した疾患を検出するアッセイにおいて有用であろう。TRICHをコードするヌクレオチド配列は、標準的な方法で標識化され、ハイブリダイゼーション複合体の形成に好適な条件の下、患者から採取した体液或いは組織のサンプルに加えることができるであろう。好適なインキュベーション期間が経過したらサンプルを洗浄し、シグナルを定量して標準値と比較する。患者のサンプルのシグナルの量が、制御サンプルと較べて著しく変わっている場合は、サンプル内のTRICHをコードするヌクレオチド配列の変異レベルにより、関連する疾患の存在が明らかになる。このようなアッセイは、動物実験、臨床試験における特定の治療効果を推定するため、或いは個々の患者の治療をモニターするために用いることもできる。 20

【0240】

TRICHの発現に関連する疾患の診断の基準となるものを提供するために、発現の正常すなわち標準的なプロファイルが確立される。これは、ハイブリダイゼーション或いは増幅に好適な条件の下、動物或いはヒトの何れかの正常な被験者から抽出された体液或いは細胞と、TRICHをコードする配列或いはその断片とを混合させることにより達成され得る。実質的に精製されたポリヌクレオチドを既知量用いて行った実験から得た値を正常な対象から得た値と比較することにより、標準ハイブリダイゼーションを定量することができる。このようにして得た標準値は、疾患の徴候を示す患者から得たサンプルから得た値と比較することができる。標準値からの偏差を用いて疾患の存在を確定する。 30

【0241】

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

【0242】

癌に関しては、個体からの生体組織における異常な量の転写物（過少発現または過剰発現）の存在は、疾患の発生素質を示したり、実際に臨床的症状が現れる前に疾患を検出する方法を提供したりし得る。この種のより明確な診断により、医療の専門家が予防方法または積極的な治療法を早くから利用し、それによって癌の発生または更なる進行を防止することが可能となる。 40

【0243】

TRICH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、或いは *in vitro* で産出し得る。オリゴマーは、好ましくはTRICHをコードするポリヌクレオチドの断片、或いはTRICHをコードするポリヌクレ 50

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

【0244】

或る実施態様において、TRICH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、挿入及び欠失である。限定するものではないがSNPの検出方法には、SSCP (single-stranded conformation polymorphism) 及び蛍光SSCP (fSSCP) がある。SSCPでは、TRICHをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いたポリメラーゼ連鎖反応 (PCR) でDNAを増幅する。DNAは例えば、病変組織または正常組織、生検サンプル、体液その他に由来し得る。DNA内のSNPは、一本鎖形状のPCR生成物の2次及び3次構造に差異を生じさせる。差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。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の実験室での調整及び統計モデル及びDNA配列クロマトグラムの自動分析を用いたシーケンシングのエラーに起因する配列の変異をフィルタリングして除去する。別の態様では、例えば高処理MASSARRAYシステム (Sequenom, Inc., San Diego CA) を用いた質量分析によりSNPを検出し、特徴付ける。

【0245】

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

【0246】

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

【0247】

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

10

20

30

40

50

【0248】

或る実施例は、或る組織または細胞タイプの転写イメージを生成するような本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞タイプにより遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所定の条件下で所定の時間に発現した遺伝子の数及び相対存在量を定量することにより分析される（Seilhamer 他、米国特許第5,840,484号の「Comparative Gene Transcript Analysis」を参照。この特許に言及することを以って本明細書の一部とする）。従って、特定の組織または細胞タイプの転写または逆転写全体に本発明のポリヌクレオチドまたはその補体をハイブリダイズすることにより、転写イメージを生成し得る。或る実施例では、本発明のポリヌクレオチドまたはその補体がマイクロ

10

【0249】

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

【0250】

本発明のポリヌクレオチドの発現のプロフィールを作製する転写イメージはまた、工業的または天然の環境化合物の毒性試験のみならず、in vitro モデル系及び薬剤の前臨床評価に関連して使用し得る。全ての化合物は、作用及び毒性のメカニズムを暗示する、しばしば分子フィンガープリントまたは毒性シグネチャと称されるような特徴的な遺伝子発現パターンを惹起する（Nuwaysir, E. F. 他（1999）Mol. Carcinog. 24:153-159、Steiner, S. 及び N. L. Anderson（2000）Toxicol. Lett. 112-113:467-471、該文献は特に引用することを以って本明細書の一部となす）。試験化合物が、既知の毒性を有する化合物のシグネチャと同一のシグネチャを有する場合には、毒性特性を共有している可能性がある。フィンガープリントまたはシグネチャは、多数の遺伝子及び遺伝子ファミリーからの発現情報を含んでいる場合には、最も有用且つ正確である。理想的には、発現のゲノム全域にわたる測定が最高品質のシグネチャを提供する。たとえ、発現が任意の試験された化合物によって変化しない遺伝子があったとしても、それらの発現レベルを残りの発現データを標準化するために使用できるため、それらの遺伝子は重要である。標準化手順は、異なる化合物で処理した後の発現データの比較に有用である。毒性シグネチャの要素に遺伝子機能を割り当てることが毒性メカニズムの解釈に役立つが、毒性の予測につながるシグネチャの統計的に一致させるのに遺伝子機能の知識は必要とされない（例えば2000年2月29日に米国国立環境健康科学研究所（National Institute of Environmental Health Sciences）より発行されたPress Release 00-02を参照されたい。これについては <http://www.niehs.nih.gov/oc/news/toxchip.htm> で入手可能である）。従って、毒性シグネチャを用いる中毒学的スクリーニングの際に、全ての発現した遺伝子配列を含めることは重要且つ望ましいことである。

20

30

40

【0251】

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

【0252】

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

50

を分析することに関連する。プロテオームの語は、特定の組織または細胞タイプでのタンパク質発現の包括的パターンを指す。プロテオームの各タンパク質成分は、個々に更に分析の対象とすることができる。プロテオーム発現パターン即ちプロファイルは、所与の条件下で所与の時間に発現したタンパク質の数及び相対存在量を定量することにより分析し得る。従って細胞のプロテオームのプロファイルは、特定の組織または細胞タイプのポリペプチドを分離及び分析することにより作成し得る。或る実施例では、1次元等電点電気泳動によりサンプルからタンパク質を分離し、2次元ドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に応じて分離するような2次元ゲル電気泳動により分離が達成される（前出のSteiner および Anderson）。タンパク質は、通常クーマシーブルーまたは銀染色液または蛍光染色液などの物質を用いてゲルで染色することにより、分散した、独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得られるタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を同定する。スポット内のタンパク質は、例えば化学的または酵素的切断とそれに続く質量分析を用いる標準的な方法を用いて部分的にシークエンシングする。スポット内のタンパク質の同一性は、その部分配列を、好適には少なくとも5個の連続するアミノ酸残基を、本発明のポリペプチド配列と比較することにより決定し得る。場合によっては、決定的なタンパク質同定のための更なる配列が得られる。

10

【0253】

プロテオームのプロファイルは、TRICHに特異的な抗体を用いてTRICH発現レベルを定量することによっても作成可能である。或る実施例では、マイクロアレイ上でエレメントとして抗体を用い、マイクロアレイをサンプルに曝して各アレイ要素へのタンパク質結合レベルを検出することによりタンパク質発現レベルを定量する（Lueking, A. 他（1999）Anal. Biochem. 270:103-111、Mendoe, 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。

20

【0254】

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。或る組織の或るタンパク質に対しては、転写とタンパク質の存在量の相関が乏しいこともあるので（Anderson, N. L. 及び J. Seilhamer（1997）Electrophoresis 18:533-537）、転写イメージにはそれ程影響しないがプロテオームのプロファイルを変化させるような化合物の分析においてプロテオーム毒性シグネチャは有用たり得る。更に、体液中での転写の分析はmRNA急速な分解により困難であるので、プロテオームのプロファイル作成はこのような場合により信頼でき、情報価値がある。

30

【0255】

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

40

【0256】

別の実施例では、タンパク質を含有する生物学的サンプルを試験化合物で処理することにより試験化合物の毒性を算定する。生物学的サンプルから得たタンパク質は、本発明のポリペプチドに特異的な抗体を用いてインキュベートする。抗体により認識されたタンパク

50

質の量を定量する。処理された生物学的サンプル中のタンパク質の量を、未処理生物学的サンプル中のタンパク質の量と比較する。両サンプルのタンパク質量の差は、処理サンプル中の試験化合物に対する反応を示す。

【0257】

マイクロアレイは、本技術分野でよく知られている方法を用いて調製し、使用し、そして分析する (Brennan, T. M. 他 (1995) の米国特許第 5, 474, 796 号、Schena, M. 他 (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619、Baldeschweiler 他 (1995) PCT 出願第 WO95/251116 号、Shalon, D. 他 (1995) PCT 出願第 WO95/35505 号、Heller, R. A. 他 (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155、Heller, M. J. 他 (1997) 米国特許第 5, 605, 662 号等を参照)。様々なタイプのマイクロアレイが公知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, 編集。(1999) Oxford University Press, London に記載されている。該文献は、特に引用することを以って本明細書の一部となす。

【0258】

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

蛍光 in situ ハイブリッド形成法 (FISH) は、他の物理的及び遺伝地図データと関連し得る (前出の Heinz-Ulrich, 他 (1995) in Meyers, 965-968 ページ。等を参照) 遺伝地図データの例は、種々の科学雑誌あるいは Online Mendelian Inheritance in Man (OMIM) のウェブサイトに見ることができる。物理的な染色体地図上の TRICH 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。

【0259】

確定した染色体マーカーを用いた連鎖分析等の物理的マッピング技術及び染色体標本原位ハイブリッド形成法を用いて、遺伝地図を拡張することができる。例えばマウスなど別の哺乳動物の染色体上に遺伝子を配置することにより、正確な染色体上の遺伝子座が未知でも、関連するマーカー類をしばしば明らかにし得る。この情報は、位置クローニングその他の遺伝子発見技術を用いて遺伝的疾患を探す研究者にとって価値がある。疾患または症候群に関連する遺伝子が、血管拡張性失調症の 11q22-23 領域等、特定の遺伝子領域への遺伝的結合によって大まかに位置決めがなされると、該領域に対するいかなる配列マッピングも、更なる調査のための関連遺伝子或いは調節遺伝子を表すことができる (

G a t t i, R. A. 他 (1988) N a t u r e 336:577-580等を参照) 転座、反転等に起因する、健常者、保有者、感染者の三者間における染色体位置の相違を発見するために本発明のヌクレオチド配列を用いてもよい。

【0260】

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

10

【0261】

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 (G e y s e n, 他 (1984) P C T a p p l i c a t i o n W O 84/03564等を参照。) この方法においては、多数の異なる小さな試験用化合物を固体基板上で合成する。試験用化合物は、T R I C H、或いはその断片と反応してから洗浄される。次に、結合されたT R I C Hが、当分野で周知の方法で検出される。精製されたT R I C Hはまた、前記した薬剤をスクリーニングする技術に用いられるプレート上で直接被覆することもできる。別法では、非中和抗体を用いてペプチドを捕捉し、ペプチドを固体支持物に固定することもできる。

20

【0262】

別の実施例では、T R I C Hと特異結合可能な中和抗体がT R I C Hとの結合について試験用化合物競合する、競合的薬剤スクリーニングアッセイを用いることができる。このようにして、T R I C Hと1つ以上の抗原決定因子を共有するどのペプチドの存在をも、抗体を使って検出できる。

【0263】

別の実施例では、将来に開発される分子生物学技術で、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) に依存しているならば、T R I C Hをコードするヌクレオチド配列にその新技術を用い得る。

30

【0264】

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

【0265】

本明細書において開示した全ての特許、特許出願及び刊行物、特に米国特許第60/243,989号、第60/245,904号、第60/249,661号、第60/247,673号、第60/252,232号、および第60/250,790号は、言及することをもって本明細書の一部となす。

【0266】

(実施例)

1 cDNAライブラリの作製

I n c y t e c D N A は、L I F E S E Q G O L D データベース (I n c y t e G e n o m i c s, P a l o A l t o C A) に記載されたcDNAライブラリに由来するものであり、表4の列5に列記した。ホモジナイズしてグアニジニウムイソチオシアネート溶液に溶解した組織もあり、また、ホモジナイズしてフェノールまたは好適な変性剤の混合液に溶解した組織もある。変性剤の混合液は、例えばフェノールとグアニジニウムイソチオシアネートの単相溶液であるT R I Z O L (L i f e T e c h n o l o g i e s) 等である。結果として得られた溶解物は、塩化セシウムにおいて遠心分離するかクロロホルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方

40

50

、或いは別の方法を用いて、溶解物からRNAを沈殿させた。

【0267】

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

【0268】

場合によってはStratagene社へのRNA提供を行い、対応するcDNAライブラリをStratagene社が作製することもあった。そうでない場合は、UNIZAPベクターシステム(Stratagene)またはSUPERSCRIPTプラスミドシステム(Life Technologies)を用いて本技術分野で公知の推奨方法または類似の方法でcDNAを合成し、cDNAライブラリを作製した(前出のAusubel, 1997, 5.1-6.6ユニット等を参照)。逆転写は、オリゴd(T)またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖cDNAに連結反応させ、好適な制限酵素でcDNAを消化した。殆どのライブラリに対して、cDNAのサイズ(300~1000bp)選択は、SEPHACRYL S1000、SEPHAROSE CL2BまたはSEPHAROSE CL4Bカラムクロマトグラフィー(Amersham Pharmacia Biotech)、或いは調製用アガロースゲル電気泳動法を用いて行った。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でcDNAを消化した。好適なプラスミドは、例えばPBLUESCRIPTプラスミド(Stratagene)、PSPORT1プラスミド(Life Technologies)、PCDNA2.1プラスミド(Invitrogen, Carlsbad CA)、PBK-CMVプラスミド(Stratagene)、PCR2-TOPOTAプラスミド(Invitrogen)、PCMV-ICISプラスミド(Stratagene)、pIGEN(Incyte Genomics, Palo Alto CA)、pINCY(Incyte Genomics)等およびその誘導体である。組換えプラスミドは、Stratagene社のXL1-Blue、XL1-BlueMRFまたはSOLR、或いはLife Technologies社のDH5α、DH10BまたはELECTROMAX DH10Bを含む大腸菌細胞に形質転換した。

【0269】

2 cDNAクローンの単離

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

【0270】

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

10

20

30

40

50

ene OR) 及び FLUOROSKANI I 蛍光スキャナ (Labsystems Oy, Helsinki, Finland) を用いて蛍光分析的に定量した。

【0271】

3 シークエンシング及び分析

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

10

20

【0272】

Incyte cDNA 配列に由来するポリヌクレオチド配列は、ベクター、リンカー及びポリ (A) 配列を除去し、あいまいな塩基対をマスクすることによって有効性を確認した。その際、BLAST、動的プログラミング及び隣接ジヌクレオチド頻度分析に基づくアルゴリズム及びプログラムを用いた。Incyte の cDNA 配列、またはその翻訳を公共のデータベース (例えば GenBank の霊長類及びげっ歯類、哺乳動物、脊椎動物、真核生物のデータベースと、BLOCKS、PRINTS、DOMO、PRODOM、PROTEOME データベース (Incyte Genomics, Palo Alto CA) (ヒト、ラット、マウス、線虫、酵母、分裂酵母、及び鵝口瘡カンジダ (*Candida albicans*) の配列を含む)、および PFAM 等隠れマルコフモデル (HMM) ベースのタンパク質ファミリーデータベース) に対して問い合わせた (HMM は、遺伝子ファミリーのコンセンサス 1 次構造を分析する確率的アプローチである。Eddy, S. R. (1996) Cuff. Opin. Struct. Biol. 6: 361-365 等を参照)。問合せは、BLAST、FASTA、BLIMPS 及び HMMER に基づくプログラムを用いて行った。Incyte cDNA 配列は、完全長のポリヌクレオチド配列を産出するように構築された。或いは、GenBank cDNA、GenBank EST、ステッチされた配列、ストレッチされた配列または Genscan 予測コード配列 (実施例 4 及び 5 を参照) を用いて Incyte cDNA の集団を完全長まで伸長させた。Phred、Phrap 及び Consed に基づくプログラムを用いて構築し、GenMark、BLAST 及び FASTA に基づくプログラムを用いて cDNA の集団をオープンリーディングフレームに対してスクリーニングした。対応する完全長ポリペプチド配列を誘導するべく完全長ポリヌクレオチド配列を翻訳した。或いは、本発明のポリペプチドは完全長翻訳ポリペプチドの任意のメチオニン残基で開始し得る。引き続き、GenBank タンパク質データベース (genpept)、SwissProt、PROTEOME データベース、BLOCKS、PRINTS、DOMO、PRODOM 及び Prosite 等のデータベース、PFAM 等の隠れマルコフモデル (HMM) に基づいた、タンパク質ファミリーデータベースに対する問合せによって完全長ポリペプチド配列を分析した。完全長ポリヌクレオチド配列はまた、MACDNASIS

30

40

50

PROソフトウェア（日立ソフトウェアエンジニアリング，South San Francisco CA）及びLASERGENEソフトウェア（DNASTAR）を用いて分析した。ポリヌクレオチド及びポリペプチド配列アラインメントは、アラインメントした配列と配列の一致率も計算するMEGALIGNマルチシークエンスアラインメントプログラム（DNASTAR）に組み込まれているようなCLUSTALアルゴリズムによって特定されるデフォルトパラメータを用いて作製する。

【0273】

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

10

【0274】

完全長ポリヌクレオチド配列およびポリペプチド配列の組み立て及び分析に用いる上記のプログラムは、SEQ ID NO: 21-40のポリヌクレオチド配列断片の同定にも利用できる。ハイブリダイゼーション及び増幅技術に有用である約20～約4000ヌクレオチドの断片を表4の列4に示した。

【0275】

20

4 ゲノムDNAからのコード配列の同定及び編集

推定輸送体及びイオンチャネルは、公共のゲノム配列データベース（例えば、gbpriやgbhtg）においてGenscan遺伝子同定プログラムを実行して初めに同定された。Genscanは、様々な生物からのゲノムDNA配列を分析する汎用遺伝子同定プログラムである（Burge, C. 及び S. Karlin (1997) J. Mol. Biol. 268: 78-94、Burge, C. 及び S. Karlin (1998) Curr. Opin. Struct. Biol. 8: 346-354を参照）。プログラムは予測エキソンを連結し、メチオニンから終止コドンに及ぶ構築されたcDNA配列を形成する。Genscanの出力は、ポリヌクレオチド及びポリペプチド配列のFASTAデータベースである。Genscanが一度に分析する配列の最大範囲は、30kbに設定した。これらのGenscan推定cDNA配列の内、どの配列が輸送体及びイオンチャネルをコードするかを決定するために、コードされたポリペプチドをPFAMモデルにおいて輸送体及びイオンチャネルについて問合せて分析した。潜在的な輸送体及びイオンチャネルが、輸送体及びイオンチャネルとして注釈が付けられたインサイトcDNA配列に対する相同性を基に同定された。こうして選択されたGenscan予測配列は、次にBLAST分析により公共データベースgenpept及びgbpriと比較した。必要であれば、genpeptからのトップのBLASTヒットと比較することによりGenscan予測配列を編集し、余分なまたは取り除かれたエキソンなどのGenscanにより予測された配列のエラーを修正する。BLAST分析はまた、任意のIncyte cDNAまたはGenscan予測配列の公共cDNA適用範囲の発見に用いられるので、転写の証拠を提供する。Incyte cDNA適用範囲が利用できる場合には、この情報を用いてGenscan予測配列を修正または確認した。完全長ポリヌクレオチド配列は、実施例3に記載された構築プロセスを用いて、Incyte cDNA配列及び／または公共のcDNA配列でGenscan予測コード配列を構築することにより得た。或いは、完全長ポリヌクレオチド配列は編集または未編集のGenscan予測コード配列に完全に由来する。

30

40

【0276】

5 cDNA配列データを使ったゲノム配列データの構築

ステッチ配列 (Stitched Sequence)

部分cDNA配列は、実施例4に記載のGenscan遺伝子同定プログラムにより予測

50

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

【0277】

ストレッチ配列 (Stretched Sequence)

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

【0278】

6 TRICH 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: 21-40 を構築するために用いた配列を、BLAST 及び Smith-Waterman アルゴリズムを用いて、Incyte LIFESEQ データベース及び公共のドメインデータベースの配列と比較した。SEQ ID NO: 21-40 と一致するこれらのデータベースの配列を、Phrap (表 7) などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。スタンフォード・ヒトゲノムセンター (SHGC)、ホワイトヘッド・ゲノム研究所 (WIGR)、Genethon 等の公的な情報源から入手可能な放射線ハイブリッド及び遺伝地図データを用いて、クラスタ化された配列が前もってマッピングされたかを決定した。マッピングされた配列がクラスタに含まれている場合は、個々の配列番号を含めてそのクラスタの全配列が地図上の位置に割り当てられた。

【0279】

地図上の位置は、ヒト染色体の範囲または間隔として表される。センチモルガン間隔の地図上の位置は、染色体の p アームの末端に関連して測定する。(センチモルガン (cM) は、染色体マーカー間の組換え頻度に基づく計測単位である。平均して、1 cM は、ヒト中の DNA の 1 メガベース (Mb) にほぼ等しい。尤も、この値は、組換えのホットスポ

ット及びコールドスポットに起因して広範囲に変化する。) cM距離は、配列が各クラスタ内に含まれるような放射線ハイブリッドマーカーに対して境界を提供するような Genethon によってマッピングされた遺伝マーカーに基づく。NCBI「GeneMap 99」(<http://www.ncbi.nlm.nih.gov/genemap/>)などの一般個人が入手可能なヒト遺伝子マップおよびその他の情報源を用いて、上記した区間が既に同定されている疾患遺伝子マップ内若しくは近傍に位置するかを決定できる。

【0280】

7 ポリヌクレオチド発現の分析

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

【0281】

BLASTを適用した類似のコンピュータ技術を用いて、GenBankやLifeSeq (Incyte Genomics) 等の cDNA データベースにおいて同一または関連分子を検索した。ノーザン分析は、多数の膜系ハイブリダイゼーションよりも非常に速い。更に、特定の同一を厳密な或いは相同的なものとして分類するか否かを決定するため、コンピュータ検索の感度を変更することができる。検索の基準は積スコアであり、次式で定義される。

【0282】

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

【0283】

或いは、TRICHをコードするポリヌクレオチド配列は、由来する組織に対して分析する。例えば或る完全長配列は、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cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各ヒト組織は、以下の生物/組織カテゴリー即ち心血管系、結合組織、消化器系、胎芽構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液及び免疫系、肝、筋骨格系、神経系、脾臓、呼吸器系、感覚器、皮膚、顎口腔系、非分類性/混合性または尿路の1つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。同様に、各ヒト組織は、以下の疾患/病状カテゴリー即ち癌、細胞株、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、TRICHをコードするcDNAの疾患特異的な発現を反映する。cDNA配列およびcDNAライブラリ／組織の情報は、LIFESEQ GOLD データベース (Incyte Genomics, Palo Alto CA) から得ることができる。

【0284】

8 TRICHをコードするポリヌクレオチドの伸長

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

10

【0285】

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

【0286】

高忠実度の増幅が、当業者によく知られている方法を利用したPCR法によって得られた。PCRは、PTC-200 サーマルサイクラー (MJ Research, Inc.) を用いて96穴プレート内で行った。反応混合液は、鋳型DNA及び200 nmolの各プライマー、 Mg^{2+} と $(NH_4)_2 SO_4$ と β -メルカプトエタノールを含むバッファー、Taq DNAポリメラーゼ (Amersham Pharmacia Biotech)、ELONGASE酵素 (Life Technologies)、Pfu DNAポリメラーゼ (Stratagene) を含む。プライマーの組、PCI AとPCI Bに対して以下のパラメータで増幅を行った。ステップ1: 94℃で3分間 ステップ2: 94℃で15秒 ステップ3: 60℃で1分間 ステップ4: 68℃で2分間 ステップ5: ステップ2、3、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。プライマー対T7、SK+に対しては、上記パラメータに代えて以下のパラメータを用いた。ステップ1: 94℃で3分間 ステップ2: 94℃で15秒 ステップ3: 57℃で1分間 ステップ4: 68℃で2分間 ステップ5: ステップ2、3、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。

20

30

【0287】

各ウェルのDNA濃度は、1X TE及び0.5 μ lの希釈していないPCR産物に溶解した100 μ 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プレート (Coming Costar, Acton MA) の各ウェルに分配してDNAが試薬と結合できるようにして測定する。サンプルの蛍光を計測してDNAの濃度を定量するべくプレートをFluoroskan II (Lab Systems Oy, Helsinki, Finland) でスキャンした。反応混合物のアリコート5～10 μ lを1%アガロースミニゲル上で電気泳動法によって解析し、どの反応が配列の伸長に成功したかを決定した。

40

【0288】

伸長させたヌクレオチドは、脱塩及び濃縮して384穴プレートに移し、CviJIコラウイルスエンドヌクレアーゼ (Molecular Biology Research, Madison WI) を用いて消化し、pUC 18ベクター (Amersham Pharmacia Biotech) への再連結反応前に音波処理またはせん断した。ショットガン・シーケンシングのために、消化したヌクレオチドを低濃度 (0.6～0.

50

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、大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りとって LB/2X カルベニシリン培養液の 384 ウェルプレートに 37℃ で一晩培養した。

【0289】

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

【0290】

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、或いはそのような伸長のために設計されたオリゴヌクレオチド及び適切なゲノムライブラリを用いて 5' 調節配列を得る。

【0291】

9 個々のハイブリダイゼーションプローブの標識及び使用

SEQ ID NO: 21-40 から導き出されたハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノム DNA をスクリーニングする。約 20 塩基対からなるオリゴヌクレオチドの標識について特に記載するが、より大きなヌクレオチド断片に対しても事実上同一の手順が用いられる。オリゴヌクレオチドは、OLIGO 4.06 ソフトウェア (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 ビードカラム (Amersham Pharmacia Biotech) を用いて十分に精製する。Ase I、Bgl II、Eco RI、Pst I、Xba I または Pvu II (DuPont NEN) のいずれか 1 つのエンドヌクレアーゼで消化されたヒトゲノム DNA の典型的な膜ベースのハイブリダイゼーション解析において、毎分 10^7 カウントの標識されたプローブを含むアリコットを用いる。

【0292】

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

【0293】

10 マイクロアレイ

マイクロアレイの表面上でアレイエレメントの結合または合成は、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等を参照）、機械的マイクロスポッティング技術及びこれらから派生したものを用いて達成することが可能である。上記各技術において基板は、均一且つ非多孔性の固体とするべきである（Sचना (1999) 前出）。推奨する基板には、シリコン、シリカ、スライドガラス、ガラスチップ及びシリコンウエハがある。或いは、ドットブロット法またはスロットブロット法に類似のアレイを利用して、熱的、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及び結合させてもよい。通常のアレイは、手作業で、または利用可能な方法や機械を用いて作製でき、任意の適正数のエレメントを有し得る（Sचना, M. 他 (1995) Science 270: 467-470、Shallon, D. 他 (1996) Genome Res. 6: 639-645、Marshall, A. および J. Hodgson (1998) Nat. Biotechnol. 16: 27-31. を参照）。

10

【0294】

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

20

【0295】

組織または細胞サンプルの準備

グアニジウムチオシアネート法を用いて組織サンプルから全RNAを単離し、オリゴ（dT）セルロース法を用いてポリ（A）⁺ RNAを精製する。各ポリ（A）⁺ RNAサンプルを、MMLV逆転写酵素、0.05 pg/μlのオリゴ（dT）プライマー（21mer）、1×第一鎖合成バッファー、0.03 unit/μlのRNA分解酵素阻害因子、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）を用いてポリ（A）⁺ RNA含有の25体積ml内で行う。特異的対照ポリ（A）⁺ RNAは、非コード酵母ゲノムDNAから*in vitro*転写により合成する。37℃で2時間インキュベートされた後、各反応サンプル（1つはCy3、もう1つはCy5標識）は、2.5 mlの0.5 M水酸化ナトリウムで処理し、85℃で20分間インキュベートし、反応を停止させてRNAを分解させる。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピンカラム（CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA）を用いて精製する。混合後、2つの反応サンプルは、1 mlのグリコーゲン（1 mg/ml）、60 mlの酢酸ナトリウム及び300 mlの100%エタノールを用いてエタノール析出させる。サンプルは次に、SpeedVAC（Savant Instruments Inc., Holbrook NY）を用いて乾燥して仕上げ、14 μlの5×SSC/0.2% SDS中で再懸濁する。

30

40

【0296】

50

マイクロアレイの準備

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

【0297】

精製したアレイエレメントは、ポリマーコートされたスライドガラス上に固定する。顕微鏡スライドガラス (Corning) は、処理中及び処理後に0.1%のSDS及びアセトン中で超音波処理をかけ、蒸留水で非常に良く洗浄する。スライドガラスは、4%フッ化水素酸 (VWR Scientific Products Corporation (VWR), West Chester PA) 中でエッチングし、蒸留水中で非常に良く洗浄し、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℃の天火で硬化させる。

10

【0298】

米国特許第5,807,522号で説明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。該特許は、引用を以って本明細書の一部となす。平均濃度が100 ng/µlのアレイエレメントDNA 1 µlを高速機械装置により開放型キャピラリープリンティングエレメント (open capillary printing element) に充填する。装置はここで、スライド毎に約5 nlのアレイエレメントサンプルを加える。

20

【0299】

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

30

ハイブリダイゼーション

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

40

【0300】

検出

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

50

【0301】

2つの異なるスキャンで、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。アレイと光電子増倍管間に設置された好適なフィルターを用いて、シグナルをフィルターする。用いる蛍光色素の最大発光の波長は、Cy3では565nm、Cy5では650nmである。装置は両方の蛍光色素からのスペクトルを同時に記録し得るが、レーザ源において好適なフィルターを用いて、蛍光色素1つにつき1度スキャンし、各アレイを通常2度スキャンする。

【0302】

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

【0303】

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

【0304】

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（Incycyte）である。

【0305】

1.1 相補的ポリヌクレオチド

TRICHをコードする配列或いはその任意の一部に対して相補的な配列は、天然のTRICHの発現を低下させるため即ち阻害するために用いられる。約15～30塩基対を含むオリゴヌクレオチドの使用について記すが、これより小さな或いは大きな配列の断片の場合でも本質的に同じ方法を用いることができる。Oligo4.06ソフトウェア（National Biosciences）及びTRICHのコーディング配列を用いて、適切なオリゴヌクレオチドを設計する。転写を阻害するためには、最も独特な5'配列から相補的オリゴヌクレオチドを設計し、これを用いてプロモーターがコーディング配列に結合するのを阻害する。翻訳を阻害するためには、相補的なオリゴヌクレオチドを設計して、リボソームがTRICHをコードする転写物に結合するのを阻害する。

【0306】

1.2 TRICHの発現

TRICHの発現及び精製は、細菌若しくはウイルスを基にした発現系を用いて行うことができる。細菌でTRICHが発現するために、抗生物質耐性及びcDNAの転写レベルを高める誘導性のプロモーターを含む好適なベクターにcDNAをサブクロニングする。このようなプロモーターには、lacオペレーター調節エレメントに関連するT5またはT7バクテリオファージプロモーター及びtrp-lac（tac）ハイブリッドプロモーターが含まれるが、これらに限定するものではない。組換えベクターを、BL21（

D E 3) 等の好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピル β -Dチオガラクトピラノシド (IPTG) で誘発されるとTRICHを発現する。真核細胞でのTRICHの発現は、昆虫細胞株または哺乳動物細胞株に一般にバキュロウイルスとして知られている Autographica californica 核多面性ウイルス (AcMNPV) を感染させて行う。バキュロウイルスの非必須ポリヘドリン遺伝子を、相同組換え或いは転移プラスミドの媒介を伴う細菌の媒介による遺伝子転移のどちらかによって、TRICHをコードするcDNAと置換する。ウイルスの感染力は維持され、強い多角体プロモーターによって高いレベルのcDNAの転写が行われる。組換えバキュロウイルスは、多くの場合は Spodoptera frugiperda (Sf9) 昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる (Engelhard, E. K. 他 (1994) Proc. Natl. Acad. Sci. USA 91: 3224-3227, Sandig, V. 他 (1996) Hum. Gene Ther. 7: 1937-1945. 等を参照)。

10

20

30

40

50

【0307】

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

【0308】

1.3 機能的アッセイ

TRICHの機能は、哺乳動物細胞培養系において生理学的に高められたレベルでのTRICHをコードする配列の発現によって評価する。cDNAを、cDNAを高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにサブクローニングする。選択されるベクターには、pCMV SPORTプラスミド (Life Technologies) 及びpCR3.1プラスミド (Invitrogen) が含まれ、どちらもサイトメガロウイルスプロモーターを有する。リポソーム製剤或いは電気穿孔法を用いて、5~10 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、形質移入された細胞と形質移入されていない細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、cDNAの組換えベクターからの発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質 (GFP; Clontech)、CD64またはCD64-GFP融合タンパク質から選択できる。自動化された、レーザー光学に基づく技術であるフローサイトメトリー (FCM) を用いて、GFPまたはCD64-GFPを発現する形質移入された細胞を同定し、その細胞のアポトーシス状態や他の細胞特性を評価する。FCMは、細胞死に先行するか或いは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、プロピジウムヨウ化物によるDNA染色によって計測される核DNA内容物の変化、前方散乱光と90°側方散乱光によって計測される細胞サイズと顆粒状性の変化、プロモデオキシウリジンの取込量の低下によって計測されるDNA合成の下方調節、特異抗体との反応性によって計測される細胞表面及び細胞

内におけるタンパク質の発現の変化、及び蛍光複合アネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変化とがある。フローサイトメトリー法については、Ormerod, M. G. (1994) Flow Cytometry Oxford, New York, NY. に記述がある。

【0309】

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

10

【0310】

14 TRICHに特異的な抗体の作製

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

20

【0311】

別法では、TRICH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)。

【0312】

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

30

【0313】

15 特異的な抗体を用いる天然TRICHの精製

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

40

【0314】

TRICHを含む培養液をイムノアフィニティークラムに通し、TRICH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TRICHとの結合を切るような条件で (例えば、pH2~3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ピックイオンで) 溶出させ、TRICHを回収する。

50

【0315】

16 TRICHと相互作用する分子の同定

TRICHと相互作用する分子には、輸送体基質、アゴニスト若しくはアンタゴニスト、 $G\beta\gamma$ タンパク質（Reimann既出）のような調節性タンパク質、又はMAGUKs（Craven既出）の様なTRICHの局在化若しくはクラスター形成に關与するタンパク質が含まれる。TRICHまたは生物学的に活性であるTRICH断片を、¹²⁵I ボルトンハンター試薬で標識する。（例えば Bolton A. E. および W. M. Hunter (1973) Biochem. J. 133:529-539を参照。）
マルチウェルプレートに予め配列しておいた候補の分子を、標識したTRICHと共にインキュベートし、洗浄して、標識したTRICH複合体を有する全てのウェルをアッセイする。様々なTRICH濃度で得られたデータを用いて、候補分子と結合したTRICHの数量及び親和性、会合についての値を計算する。 10

【0316】

別法では、TRICHと相互作用する分子がFields, S. 及びO. Song (1989, Nature 340:245-246) に記載の酵母2-ハイブリッドシステム（yeast two-hybrid system）で単離された。TRICHもしくはその断片は、Gal4若しくはlexAのDNA結合ドメインを有する融合タンパク質として発現し、また潜在的な相互作用タンパク質が活性ドメインを有する融合タンパク質として発現する。TRICH融合タンパク質とTRICH相互作用タンパク質（活性化ドメインを有する融合タンパク質）再構成物との間の相互作用、すなわちトランス活性化機能は、レポーター遺伝子の発現によって観察される。酵母2-ハイブリッドシステムは一般に利用可能であり、イオンチャネルタンパク質を有する酵母2-ハイブリッドシステムの使用法は、Niethammer, M. およびM. Shengらの著書（1998, Meth. Enzymol. 293:104-122）に記載されている。 20

【0317】

TRICHはまた、ハイスループット型の酵母2-ハイブリッドシステムを使用するPATHCALLINGプロセス（CuraGen Corp., New Haven CT）に用いて、遺伝子の2つの大きなライブラリによってコードされるタンパク質間の全ての相互作用を決定することができる（Nandabalan, K. 他（2000）米国特許第6,057,101号）。 30

【0318】

潜在的なTRICHアゴニスト若しくはアンタゴニストは、実施例18に記載のアッセイを用いてTRICHイオンチャネル活性の活性化若しくは阻害のために試されてもよい。

【0319】

17 TRICH活性の実証

TRICHのイオンチャネル活性は、イオンコンダクタンスのための電気生理学的アッセイを用いて実証される。TRICHは、TRICHをコードする真核生物発現ベクターを備えたCOS7、HeLa、若しくはCHOのようなほ乳類株化細胞を形質転換することで発現されうる。真核生物発現ベクターは市販されており、それらを細胞内に導入する技術は当業者には周知である。 β -ガラクトシダーゼのような複数の標識遺伝子いずれか一つを発現させる第二のプラスミドは、細胞へと同時形質転換され、異質なDNAを吸収し発現させるそれら細胞の迅速な同定を可能とする。細胞は、形質転換の後、株化細胞がTRICH及び 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、48-72時間に渡ってインキュベートされる。 40

【0320】

β -ガラクトシダーゼを発現させる形質転換細胞は、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。染色された細胞は、本技術分野で既知の電気生理学技術を用いて膜電気伝導度の違いを試験される。TRICHを発現する細胞は、対照細胞に対して高いアニオンコンダクタンス、若しくはカチオンコンダクタンスを有する。TRICHの伝導度への寄与は、TRICHに特異的な抗体を用いて細胞を 50

インキュベーションすることによって確認できる。抗体は、TRICHの細胞外側面に結合し、それによってイオンチャネル内の孔とそれに伴うコンダクタンスをブロッキングする。

【0321】

別法では、TRICHのイオンチャネル活性は、二電極電位クランプ技術（Ishii 他既出、Jegla, T. 及び L. Salukoff (1997) J. Neurosci. 17 : 32-44) を用いて、TRICH含有アフリカツメガエル卵母細胞膜を通る電流として測定される。TRICHは、pBFのような適切なアフリカツメガエル卵母細胞発現ベクターへとサブクローニングされ、また0.5~5 ngのmRNAが成熟段階4の卵母細胞へ注入される。注入された卵母細胞は摂氏18度で1-5日間インキュベートされる。インサイド-アウトマクロパッチが、切り取られて、116 mMのK-グルコン酸、4 mMのKCl、および10 mMのHepes (pH 7.2) を含む細胞内溶液へと移される。細胞内溶液には、適切な、cAMP、cGMP、または Ca^{+2} ($CaCl_2$ の形態) のような様々な濃度のTRICH媒介物質が補足される。電極の抵抗は2-5 MΩでセットされ、電極は媒介物質のない細胞内溶液で満たされる。実験は、室温で0 mVの保持電位で実行される。-100 mVから100 mVの電圧ランプが、サンプリング周波数500 Hzで得られる。電流測定はアッセイに於けるTRICHの活性に比例する。

10

【0322】

特に、TRICH-2の活性は電圧依存性の Ca^{+2} または Na^{+} の伝導度として測定され、TRICH-15の活性は Ca^{+2} の伝導度として測定され、TRICH-16の活性は K^{+} の伝導度として測定される。

20

【0323】

TRICHの変換活性は、アフリカツメガエル卵母細胞への標識された基質の取り込みによって検査される。ステージ5及び6における卵母細胞は、TRICH mRNA (卵母細胞あたり10 ng) で注入され、OR2培地 (82.5 mMのNaCl、2.5 mMのKCl、1 mMの $CaCl_2$ 、1 mMの $MgCl_2$ 、1 mM Na_2HPO_4 、5 mMのHepes、3.8 mMのNaOH、50 μg/ml ゲンタマイシン、pH 7.8) 内で摂氏18度で3日間インキュベートされ、TRICHの発現を可能とする。卵母細胞は、次に標準取り込み培地 (100 mMのNaCl、2 mMのKCl、1 mMの $CaCl_2$ 、1 mMの $MgCl_2$ 、10 mMのHepes/Tris、pH 7.5) へと移される。様々な基質 (例えばアミノ酸、糖、薬剤、イオン、及び神経伝達物質) の取り込みは、標識された基質 (例えば ^3H で放射標識されていたり、ローダミンで蛍光標識されている) を添加することで開始される。30分間のインキュベートの後、取り込みは、卵母細胞を3回、 Na^{+} のない培養液で洗浄することで終了し、取り込まれた標識を測定して、対照と比較する。TRICH活性は、取り込まれた標識基質のレベルに比例する。具体的な試験基質としては、TRICH-1に対してトリカルボン酸、TRICH-4に対して硫酸、TRICH-5に対して Na^{+} 、TRICH-6に対してアニオン性代謝産物、TRICH-8に対してグルコース6リン酸、TRICH-10に対してアミノ酸が挙げられる。

30

40

【0324】

TRICHに関連するATPase活性は、放射性ラベルされたATP- [γ - ^{32}P] の加水分解、クロマトグラフィー法による加水分解生成物の分離、及びシンチレーションカウンターを用いる再生 ^{32}P の定量化によって測定される。反応混合物は、ATP- [γ - ^{32}P] と、好適な時間をかけて摂氏37度でインキュベートされた好適なバッファ中の様々な量のTRICHとを有する。反応は、トリクロロ酢酸の沈殿によって終了し、塩基で中和され、反応混合物のアリコットは、反応混合物を分離するべく、膜若しくはペーパークロマトグラフィーにかけられる。遊離された ^{32}P の量は、シンチレーション計数器で測定される。回収した放射活性の量はアッセイ中のTRICHのATPase活性に比例する。

50

【0325】

TRICHのlipocalin活性はリガンド蛍光エンハンスメント分光蛍光分析によって測定される(Lin 他 (1997) Molecular Vision 3:17)。リガンドの例としては、レチノール(Sigma, St. Louis MO) および16アンスリロキシパルミチン酸(16-AP)(Molecular Probes Inc., Eugene OR)等が挙げられる。リガンドは100%エタノールに溶解され、濃度は既知の吸光係数(レチノール: $46,000 \text{ A/M/cm @ } 325 \text{ nm}$; 16-AP: $8,200 \text{ A/M/cm @ } 361 \text{ nm}$)を使って推測される。10 mM Tris (pH 7.5)、2 mM EDTAおよび500 mM NaCl 溶液中に溶かした1 μM TRICH の700 μl アリコットを光路長1 cm の石英キュベットに入れ、1 μl アリコットのリガンド溶液を加えた。蛍光は読取値が安定化するまで、各追加時の100秒後に測定された。リガンド濃度の単位変化あたりの蛍光の変化がTRICHの活性に比例する。

【0326】

18 TRICHアゴニスト及びアンタゴニストの同定

TRICHは、CHO(チャイニーズハムスターの卵巣)またはHEK(ヒト胎児腎臓)293のような真核細胞株内で発現する。形質転換細胞のイオンチャネル活性は、候補アゴニスト又はアンタゴニストの存在若しくは不在条件で測定される。イオンチャネル活性は実施例17に記載の本技術分野では既知のパッチクランプ法を用いて検定される。別法では、イオンチャネル活性が、細胞膜を通るイオンの流れを測定する蛍光技術を用いて検定される(Velicelebi, G.ら(1999) Meth. Enzymol. 294: 20-47; West, M. R. 及びC. R. Mollroy (1996) Anal. Biochem. 241: 51-58)。これらアッセイは、ミクロプレートを用いる高処理のスクリーニングに適合しうる。内部イオン濃度の変化は、 Ca^{2+} 指示薬Fluo-4 AM、SBFI及びナトリウムグリーンのようなナトリウム感受性色素、Cl指示薬MQAE(すべてMolecular Probes社より入手可能)のような蛍光色素を用い、FLIPR蛍光定量プレートリーディングシステム(Molecular Device社)と併せて測定される。アッセイのさらに一般的な別法では、原形質膜を通るイオンの流れによって生じる膜電位の変化は、DiBAC₄(Molecular Probes社)のようなoxonyl色素を用いて測定される。DiBAC₄は、細胞膜電位に従って細胞外溶液と細胞内との間で平衡となる。色素の蛍光強度は、疎水性の細胞内部に結合すると20倍以上となり、細胞内へのDiBAC₄の流入が検出可能となる(Gonzalez, J. E. 及びP. A. Negulescu (1998) Curr. Opin. Biotechnol. 9: 624-631)。候補アゴニスト若しくはアンタゴニストは、既知のイオンチャネルアゴニスト、若しくはアンタゴニスト、ペプチドライブラリ、若しくは組み合わせの化学ライブラリより選択されてもよい。

【0327】

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

【0328】

(表の簡単な説明)

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

【0329】

表2は、GenBank識別番号及び本発明のポリペプチドに最も近いGenBank相

同体の注釈を示す。各ポリペプチドとその相同体が一致する確率スコアも併せて示す。

【0330】

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

【0331】

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

【0332】

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。

10

【0333】

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

【0334】

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

【表1】

表 1

IncYTE クト ID	プロジェクト NO:	IncYTE ポリペプチド ID	ポリヌクレオチド SEQ ID NO:	IncYTE ポリヌクレオチド ID
1626101	1	1626101CD1	21	1626101CB1
2907828	2	2907828CD1	22	2907828CB1
3968527	3	3968527CD1	23	3968527CB1
7472732	4	7472732CD1	24	7472732CB1
7476938	5	7476938CD1	25	7476938CB1
8128531	6	8128531CD1	26	8128531CB1
7476757	7	7476757CD1	27	7476757CB1
266243	8	266243CD1	28	266243CB1
6585710	9	6585710CD1	29	6585710CB1
7483599	10	7483599CD1	30	7483599CB1
2507246	11	2507246CD1	31	2507246CB1
3033505	12	3033505CD1	32	3033505CB1
4027693	13	4027693CD1	33	4027693CB1
7472030	14	7472030CD1	34	7472030CB1
7476089	15	7476089CD1	35	7476089CB1
6428177	16	6428177CD1	36	6428177CB1
7477243	17	7477243CD1	37	7477243CB1
7473042	18	7473042CD1	38	7473042CB1
7482060	19	7482060CD1	39	7482060CB1
1578772	20	1578772CD1	40	1578772CB1

【表 2】

10

20

30

40

表2-1

ポリパプチドSEQ ID NO:	Incyte ポリパプチ ドID	GenBank ID NO:	確率スコア	GenBank 相合体
1	1626101CD1	g13785618	1.00E-105	[f1][マウス]sideroflexin 4 Fleming, M. D. 他 (2001) A mutation in a mitochondrial transmembrane protein is responsible for the pleiotropic hematological and skeletal phenotype of flexed-tail (f/f) mice. Genes Dev. 15:652-657
		g545998	1.00E-15	[ラット] トリカルボン酸キャリア Azzi, A. 他 (1993) The mitochondrial tricarboxylate carrier. J. Bioenerg. Biomembr. 25:515-524
2	2907828CD1	g12004581	0.0	[f1][マウス]カルシウムチャネル
3	3968527CD1	g6434968	0.0	[マウス] 推定上の E1-E2 ATP 分解酵素 Halleck, M.S. 他 (1999) Differential expression of putative transbilayer amphipath transporters. Physiol. Genomics (Online) 1:139-150
4	7472732CD1	g15341552	0.0	[f1] [ヒト] 推定上のアニオン輸送体
		g575895	1.40E-82	[マウス] 硫酸輸送体 Kobayashi, T. 他 (1997) Cloning of mouse diastrophic dysplasia sulfate transporter gene induced during osteoblast differentiation by bone morphogenetic protein-2. Gene 198:341-349
5	7476938CD1	g4894991	6.00E-139	[ショウジョウバエ] ナトリウム水素交換体 NHE1
6	8128531CD1	g5052319	6.00E-51	[ヒト] シトリン; 成人発症型タイプIIシトルリン血タンパク質 Kobayashi, T. 他 (1999) The gene mutated in adult-onset type II citrullinaemia encodes a putative mitochondrial carrier protein. Nat. Genet. 22:159-163
7	7476757CD1	g1653342	1.80E-16	[ラン藻] メリオースキヤリアタンパク質 Kaneko, T. 他 (1996) DNA Res. 3:109-136
8	266243CD1	g7229675	6.90E-39	[シロイヌナズナ] グルコース 6 リン酸/リン酸トランスロケータ

10

20

30

40

表2-2

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	確率スコア	GenBank 相同体
9	6585710CD1	g15146444	0.0	[ヒト]sterolin-2 Lu, K. 他 (2001) Two genes that map to the stsl locus cause sitosterolemia: genomic structure and spectrum of mutations involving sterolin-1 and sterolin-2, encoded by ABCG5 and ABCG8, respectively. Am. J. Hum. Genet. 69:278-290
11	2507246CD1	g472900	3.10E-66	[線虫] キヤリアタンパク質 (c2) Runswick, M.J. 他 (1994) Extension of the mitochondrial transport superfamily: sequences of five members from the nematode worm <i>Caenorhabditis elegans</i> . DNA Seq. 4:281-291
12	3033505CD1	g6978016	4.40E-239	[ラット] 神経グルタミン輸送体 Varoqui, H. 他 (2000) Cloning and functional identification of a neuronal glutamine transporter. J. Biol. Chem. 275:4049-4054
13	4027693CD1	g2198807	1.10E-53	[ニワトリ] モノカルボン酸輸送体 3 Yoon H. 他 (1997) Biochem. Biophys. Res. Commun. 234:90-94; Yoon H. および Philp N. (1998) J. Exp. Eye Res. 67:417-424; Yoon H. 他 (1999) Genomics 60:366-370
14	7472030CD1	g387426	0.0	[マウス] 多剤耐性タンパク質 Gros, P. 他 (1986) Cell 47:371-380
15	7476089CD1	g2826759	2.50E-11	[線虫] ナトリウムカルシウム交換体
16	6428177CD1	g3880445	1.70E-14	[線虫] Pfam ドメインに類似した部分あり: PF02214 (K ⁺ チャネル四量体化ドメイン)
17	7477243CD1	g6457274	0.0	[マウス] 推定上の E1-E2 ATP 分解酵素 Halleck, M.S. 他 (1999) Physiol. Genomics (Online) 1:139-150

【表 4】

表2-3

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	確率スコア	GenBank 相同体
18	7473042CD1	g15211933	0.0	[F1] [ヒト] プロトンマイオイノシトール輸送体 Uldry, M. 他 (2001) Identification of a mammalian H(+)-myo-inositol symporter expressed predominantly in the brain. The EMBO Journal 20:4467-4477
19	7482060CD1	g6006493	8.80E-83	[ヒト] 心臓カリウムチャネルサブユニット (Kv6.2)
20	1578772CD1	g11933425	2.00E-05	[F1] [シロイヌナズナ] 硫酸輸送体
		g11907976	3.00E-05	[F1] [ジャガイモ] 高濃和性硫酸輸送体タイプ1

表3-1

SEQ ID NO:	incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
1	1626101CD1	337	S221 S317 S324 T158 T34 T71	N154	タンパク質 膜貫通 染色体 推定 輸送体 C17G6.15C 輸送 XV リーディングフレーム PD006986:A20-P264	BLAST_PRODOM
2	2907828CD1	816	S123 S264 S351 S359 S375 S395 S4 S54 S697 S703 S716 S745 S769 T14 T322 T382 T559 T602 T618 T639 T764 Y624	N599 N611 N616 N695	膜貫通ドメイン: L107-L124, I235-F254, C297-F320, G506-L523, M560-F577, I666-I686 イオン輸送タンパク質ドメイン: L437-I686	HMMER HMMER_PFAM
3	3968527CD1	1047	S1038 S179 S346 S366 S417 S453 S491 S498 S499 S548 S559 S605 S624 S629 S835 S920 T143 T147 T207 T212 T240 T276 T377 T390 T397 T445 Y528 T634 T649 I665 T687 T707 I763 T776 T981 Y611	N182 N285 N535	ナトリウムチャネルシグネチャ PR00170: Q227-F254, S296-D325 膜貫通ドメイン: V300-V319, I953-M980, V1004-S1023 E1-E2 ATP 分解酵素ドメイン: G146-I174, N256-E279 E1-E2 ATP 分解酵素リン酸化部位 タンパク質シグネチャ BL00154: G158-F175, I385-F403, D653-L693 E1-E2 ATP 分解酵素リン酸化部位 T371-D419 P 型陽イオン輸送 ATP 分解酵素スーパーファミリーシグネチャ PR00119: L389-F403, A669-D679	HMMER HMMER_PFAM BLIMPS_BLOCKS PROFILES SCAN BLIMPS_PRINTS

【表 6】

10

20

30

40

表3-2

SEQ ID NO:	Incyte 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
					ATP 分解酵素 ヒドラーゼ 膜貫通 リン酸化 ATP 結 合 タンパク質 推定 カルシウム輸送 カルシウム 輸送 PD004657:S817-T1009 PD149930:C757-F816 PD006317:K149-D245	BLAST_PRODOM
					do ATP 分解酵素; カルシウム; 輸送; DM02405 Q10309 127-865:Y495-S881, S152-A484, S91-V270 DM02405 P40527 208-977:R201-S881, S91-A482 DM02405 Q09891 206-1107:E715-L851, L444-F703, N154-R327, L336-G414 DM02405 P39524 236-1049:V148-V869, Q92-D365, I926-S1038	BLAST_DOMO
4	7472732CD1	671	S10 S138 S225 S311 S345 S352 S494 S556 S640 S658 T507 T595	N125 N131 N661	E1-E2 ATP 分解酵素モチーフ: D391-T397 膜貫通ドメイン: L228-N248, I399-Y417, V470-P488 硫酸輸送体ファミリードメイン: M162-F487 STAS (硫酸輸送体およびアンチンシグマファクターア ンタゴニスト) ドメイン: E508-A652 硫酸輸送体タンパク質シグネチャ BL01130: A150-M201, S53-L106	MOTIFS HMMER HMMER_PFAM HMMER_PFAM BLIMPS_BLOCKS

10

20

30

40

表3-3

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					タンパク質輸送 硫酸輸送体 膜貫通 ペルミアゼ 遺伝子間領域 親和性 糖タンパク質 PD001255:V164-R486 PD001121:P33-G168	BLAST_PRODOM
					硫酸輸送体 DM01229 P40879 5-462:P33-W446 DM01229 P50443 49-505:L32-N447 DM01229 P45380 10-468:R6-W446 DM01229 P53393 11-447:P33-M201, H303-W446	BLAST_DOMO
5	7476938CD1	671	S15 S288 S582 S661 S663 S68 S85 T104 T120 T334 T440 T503 T584 T600 T630 T634 T649	N103 N110 N276 N337 N47 N580	膜貫通ドメイン: M150-R171, Y172-I194, I241-I263, G390-G418, V452-L480, L537-G555 ナトリウム/水素交換体ファミリー: I152-D568 Na+/H+ 交換体シグネチャ PR01084: M215-F226, G229-S243, I244-T252, G284-A294 + 輸送交換体 NA PD01672: M215-I263, A297-L333 NA+/H+ タンパク質 膜貫通 輸送 アンチポータ シンポータ ナトリウム交換体 糖タンパク質 ナトリウム/水素 PD000631:G149-E567 do ベータ: 交換体 NA: DM02572 P48764 10-703: DM02572 P48761 17-738:S155-N580 DM02572 P26434 14-716:L156-R632 DM02572 P48764 10-734:L156-E636	HMNER HMNER_PFAM BLIMPS_PRINTS BLIMPS_PRODOM BLAST_PRODOM BLAST_DOMO

【表 8】

10

20

30

40

表3-4

SEQ ID NO.	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	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
6	8128531CD1	315	S165 T149 T160 T2 T240 T251 T55 T9 Y261	N71	ミトコンドリアキヤリアタンパク質ドメイン: S7-Q99, N101-G217 ミトコンドリアエネルギー伝達タンパク質シグネチャ BL00215: I13-Q37, L173-G185 ミトコンドリアエネルギー伝達タンパク質シグネチャ: H3-T55, L102-S150, F221-L271 タンパク質輸送 膜貫通 リポート ミトコンドリア キヤリア 膜内 ミトコンドリア ADP/ATP PD00117: S7-Y302	HMMER_Pfam BLIMPS_BLOCKS PROFILES SCAN BLAST_PRODOM
					ミトコンドリアエネルギー伝達タンパク質 DM00026 S54495 534-620:G15-G97 DM00026 S54495 622-719:E105-Q131, A170-L207 DM00026 S60949 16-113:I13-L92	BLAST_DOMO
					ミトコンドリアキヤリアタンパク質モチーフ: P28-L36 P241-I249	MOTIFS
7	7476757CD1	445	S247 S26 S389 S393 S440 T181 T355 T380 T405		ナトリウム: ガラクタクトシドシンボータファミリー DM01084 P30868 1-456:L178-K382 (p=6.1e-07) 輸送 タンパク質 膜貫通 シンボータ 糖 シンボータ ペルミアアゼ 内膜 キヤリア PD003362: R183-L375 (p=5.7e-09)	BLAST_DOMO BLAST_PRODOM
8	266243CD1	410	S162 S208 S356 S363 T140 T180	N295 N333	シグナルペプチド signal_peptide: M1-G29 膜内在性タンパク質 DUF6 DUF6: A39-V181	HMMER HMMER_Pfam

【表 9】

10

20

30

40

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
			T297 T335 T351		膜間スペースドメイン DM02684 P52178 1-401: V112-K336, P15-L63 DM02684 S37550 1-407: R110-K345, L36-V88 DM02684 S37497 1-409: P93-K345, H10-G90 DM02684 P52177 1-406: R110-K345, V6-S89 トランスロケータ 前駆体 膜貫通 CTPP リン酸/ホスホエノールピルビン酸 トリオース リン酸/リン酸 ノングリシン PLASTID PD150555: L184-K336 transmembrane_domain: L169-F186, S316-Y332	BLAST_DOMO
9	6585710CD1	374	S145 S174 S310 S332 S362 T26 T313 T73 T80	N280 N320	do ホワイット; FRUIT; FLY; スカーレット; DM05200 P45844 289-650: M1-V348 DM05200 P10090 317-666: V2-L346 do ペルミアーズ; 依存; ATP; PDR10; DM01528 P51533 406-797: Y4-P340 DM01528 S55517 406-797: Y4-P340 transmembrane_domain: L118-F137, Y339-V361	BLAST_DOMO BLAST_DOMO
10	7483599CD1	443	S412 S59 T26 T369 T388	N259	膜貫通アミノ酸輸送体 Aa_trans: A104-F438 酸アミノタンパク質輸送体 透過酵素 膜貫通 遺伝子間領域 推定 プロリン PD001875: G79-L367 transmembrane_domain: V115-C134, V177-F195, Y231-F254, F292-L310, L326-G344 signal_cleavage: M1-G25	HMMER_PPFAM BLAST_PRODOM
11	2507246CD1	321	S209 S307 S41 S80 T226 T274 Y268	N224 N229	ミトコンドリア キヤリアタンパク質 N10-P125, S127-A220, S232-X322	HMMER SPSCAN HMMER_PPFAM

【表10】

10

20

30

40

表3-6

SEQ ID NO.	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					ミトコンドリアエネルギー伝達タンパク質 BL00215:F16-Q40, V177-G189	BLIMPS_BLOCKS
					ミトコンドリアエネルギー伝達タンパク質シグネチャ: Q6-V101, Q6-G100, Q6-V99, Q6-L98, Q6-N97, Q6-P96, Q6-G95, Q6-L94, Q6-K84, F233-G283	PROFILES SCAN
					輸送 膜貫通 ミトコンドリア キヤリア ミトコンドリア内 ADP/ATP PD000117:L171-E317, H14-E216, S127-F292, S127-E264, L12-S232, R184-T320, N10-E117, T231-E310	BLAST_PRODOM
					ミトコンドリアエネルギー伝達タンパク質 DM000026 S44092 201-284:P125-L214, G73-I122, M237-E310, N10-Q40 DM000026 S44092 302-380:M237-L316, S134-E208, T68-Y112, L15-L82 DM000026 P38127 56-163:A17-F119, E263-Y309 DM000026 P38127 291-375:A239-L316, S134-L214, G73-Y112, A17-S42	BLAST_DOMO
					ミトコンドリアエネルギー伝達タンパク質シグネチャ: P31-L39 P253-L261	MOTIFS
12	3033505CD1	487	S303 S347 S378 S413 S45 S481 S482 S49 S56 S6 T145 T17 T259 T265 T32 T332 T355 T374 T442	N15 N23 N251 N257 N26 N312 N79	膜貫通ドメイン: M85-L107, V198-T221, F224-I243, F316-N336, L352-P372, I399-V421, L457-W477 膜貫通アミノ酸輸送体: A95-S469	HMVER HMVER_PFAM

【表 1 1】

10

20

30

40

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
			T450		酸 アミノ 輸送体 透過酵素 膜貫通 遺伝子間 プロリン PD001875:S76-V370 輸送体タンパク質 PD138374:H343-H487	BLAST_PRODOR BLAST_PRODOR
13	4027693CD1	509	S233 T267 T419 Y277	N411	膜貫通ドメイン: P344-F366, L400-V418, F464-L483 モノカルボン酸輸送体: S19-S467 輸送体; 連結 DM05037 Q03064 1-475:P7-N244, V298-L480 DM05037 P53988 1-465:P7-S194, A299-L480 DM05037 P36021 155-612:P7-Y211, L300-L483	HMMER HMMER_PFAM BLAST_DOMO
14	7472030CD1	1232	S1081 S1111 S1159 S1165 S253 S369 S401 S426 S554 S651 S654 S673 S887 T1178 T1186 T1211 T137 T20 T233 T491 T586 T595 T640 T728 T886 Y505 Y875	N1079 N1163 N189 N300 N372 N391 N424 N703 N764 N794 N86 N885 N92	膜貫通ドメイン: V112-I131, I735-Y751, M812-M834 ABC 輸送体膜貫通領域: M49-I340, V693-I943 ABC 輸送体: G1018-G1204, G415-G599 ABC 輸送体ファミリーシグネチャ: M526-L540, I1131-L1145 ATP/GTP 結合部位モチーフ A (P-loop): G422-S429, G1025-S1032 ABC 輸送体ファミリーシグネチャ: I508-D557, I1113-D1162	HMMER HMMER-PFAM HMMER-PFAM MOTIFS MOTIFS ProfileScan

【表 1 2】

10

20

30

40

表3-8

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					ABC 輸送体ファミリー BL00211: L420-V431, L1131-D1162	BLIMES-BLOCKS
					ATP 結合輸送膜貫通タンパク質 PD00131: G162-D171, C1029-T1082, G1179-L1216	BLIMES-PRODOM
					MALK タンパク質 DM00130 P21448 53-386: G52-G385, F748-G988, G696-N723	BLAST-DOMO
					MALK タンパク質 DM00130 P21448 53-399: L54-G385, F748-G988, T697-N723	BLAST-DOMO
					MALK タンパク質 DM00130 P21448 53-391: L54-G385, F748-G988	BLAST-DOMO
					MALK タンパク質 DM00130 P21448 53-391: L54-G385, F748-G988, T697-N724	BLAST-DOMO
					ATP 結合 膜貫通輸送体、多剤耐性、ABC 輸送体 PD00130: L48-I334, F748-F933, V693-I718	BLAST-PRODOM
					P-糖タンパク質、多剤耐性、ATP 結合輸送体 PD167072: I465-A524	BLAST-PRODOM
					ATP 結合膜貫通輸送体 PD000101: E1058-G1128	BLAST-PRODOM
					膜貫通ドメイン: I251-V267 ナトリウム/カルシウム 交換体: DM05297 P48765 6-969: V117-I346 (P 値 = 6.0e-10)	HMME BLAST-DOMO
15	7476089CD1	759	S128 S161 S236 S315 S349 S405 S47 S505 S644 S646 S663 S710 S77 S84 T12 T130 T271 T435 T628 T735	N173 N184 N218 N250 N306 N334 N393 N461 N521 N545 N626 N682 N72 N729 N739		
16	6428177CD1	283	S163 S96 T119 Y112	N26 N87	K+ チャネル四量体化ドメイン: V58-Q155	HMME-PFAM

【表 13】

10

20

30

40

表3-9

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
					カリウムチャネルシグネチャ: H100-T119 カリウムチャネル CDRK、SHAW: DM00490 P17972 1-102: V58-L143 (P 値 = 2.8e-7)	BLIMPS-PRINTS BLAST-DOMO
17	7477243CD1	1129	S30 S47 S138 S278 S282 S442 S494 S495 S548 S708 S733 S736 S762 S813 S924 S982 S1094 S1100 S1105 S1109 S1113 T204 T250 T254 T264 T308 T328 T334 T408 T413 T449 T646 T680 T693 T701 T704 T1008 T1121 Y258 Y747	N121 N392 N761 N992 N1098	膜貫通ドメイン: F995-A1012, I1070-K1088 E1-E2 ATP 分解酵素リン酸化部位 タンパク質 BL00154: G144-L161, V403-F421, K563-V573, D650-L690, T811-K834 E1-E2 ATP 分解酵素リン酸化部位: A389-V438 P 型陽イオン輸送 ATP 分解酵素スーパファミリシグネチャ PR00119: F407-F421, A666-D676, I814-I833 ATP 分解酵素 ヒドロラーゼ 膜貫通 リン酸化 ATP 結合 カルシウム 輸送 PD004657: S848-K1088 PD149930: C787-Y847 PD006317: R135-I225 推定上のカルシウム輸送 ATP 分解酵素 ヒドロラー ゼ カルシウム 輸送 膜貫通 リン酸化 マグネシウム ATP 結合 PD101227: R458-V583, R16-L77	HMMER BLIMPS_BLOCKS PROFILES SCAN BLIMPS_PRINTS BLAST_PROD OM BLAST_PROD OM

10

20

30

40

表3-10

SEQ ID NO:	Incyte ペリペチド 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
					ATP 分解酵素; カルシウム; 輸送 DM02405 P39524 236-1049: Q79-L760, S708-N912, V989-F1024 DM02405 P32660 318-1225: E460-N912, E460-N912, I141-G446 DM02405 Q09891 206-1107: E460-N912, I141-G446, Y975-F1024 DM02405 S51243 356-1267: E460-Y911, E139-G446, V989-F1039, K453-G504	BLAST_DOMO
					ATP/GTP 結合部位モチーフA(Pループ) A271-S278	MOTIFS
					E1-E2 ATP 分解酵素リン酸化部位 D409-T415	MOTIFS
18	7473042CD1	648	S2 S6 S224 S262 S294 S355 S419 S474 S607 S619 S640 S645 T12 T272 T287 T322 T394 T423 T461 Y11	N285 N433 N458 N485	膜貫通ドメイン: A79-F95, L366-V385, F395-S415 糖(およびその他の)輸送体: V84-F609 糖輸送タンパク質 BL00216: G92-S103, L174-A223 糖輸送タンパク質シグネチャ: L366-R421, S162-V225 糖輸送体シグネチャ PR00171: G92-V102, L175-V194, Q336-Y346, L510-V531, S533-N545	HMMER HMMER_PFAM BLIMPS_BLOCKS PROFILES SCAN BLIMPS_PRINTS
					グルコース輸送体シグネチャ PR00172: L326-Y347, I364-V385, L90-K110, L510-S533, T543-L561, G574-L594 糖輸送体タンパク質 PD000537: K296-R391	BLIMPS_PRINTS BLAST_PRODOM

【表 15】

10

20

30

40

表3-11

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
					糖輸送タンパク質 DM00135 Q01440 101-433: R178-S419, L513-G599 DM00135 P54723 120-454: R178-L414, T503-K600 DM00135 S25009 121-478: G161-Q417, L510-K598 DM00135 S43230 170-502: R178-Q417, S506-K600	BLAST_DOMO
					糖輸送タンパク質シグネチャ1: G381-G396	MOTIFS
					糖輸送タンパク質シグネチャ2: V180-R205	MOTIFS
19	7482060CD1	545	S36 S43 S128 S198 S335 S405 T185 T269 T297 T441 Y95	N196	膜貫通ドメイン: I165-C182, V268-A285, V306-L325 イオン輸送タンパク質: I175-I390 K+ チャネル四量体化ドメイン: A9-L116 カリウムチャネルシグネチャ PR00169: E60-G79, A157-T185, I205-K228, F231-V251, M275-C301, E304-E327, F339-M361, G368-F394 チャネル イオン性 タンパク質 カリウム サブユニット 電圧依 存性 膜貫通 カルシウム 輸送 イオン PD000141: F231-Y398	HMMER HMMER_PFAM HMMER_PFAM BLIMPS_PRINTS BLAST_PRODUM

【表 16】

10

20

30

40

表3-12

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
					チャネル; カリウム; CDRK; SHAW; DM00490 JH0595 26-142: V11-R115 DM00490 P15387 18-134: R5-R115 DM00436 JH0595 144-307: A163-I278 DM00490 P17970 268-384: V11-R115	BLAST_DOMO
20	1578772CD1	262	S87 S154 S232 S249 T153 T178		signal cleavage: M1-G45 シグナルペプチド: M1-A19 STAS ドメイン (硫酸輸送体およびアンチシグマファクターアンタゴニスト) : H110-A236 硫酸輸送体タンパク質輸送膜貫通親和性糖タンパク質硫酸高疾患 PD001755:H110-A236 E値: 3.0e-08	SPSCAN HMME HMME_PPFAM BLAST_PRODOR

【表 17】

10

20

30

40

表4-1

ポリヌクレオチドSEQ ID NO:	Incyte ポリヌクレオチド ID	配列長	選択された断片	配列断片	5'位置	3'位置
21	1626101CB1	1373	577-657	6122382H1 (BRAHNON05)	220	880
				2822668H1 (ADRETUT06)	1	305
				72008374V1	322	1036
				2378333T6 (ISLTNOT01)	822	1373
22	2907828CB1	3231	1-94, 2624-2681, 807-866	7276683H1 (LIVRDIS04)	2591	3231
				55084582J1	1325	1931
				55124156J1	574	1337
				7602868J1 (ESOGTME01)	468	1025
				7697278J1 (KIDPTDE01)	1914	2573
				7348609H1 (COLNNON05)	2022	2587
				7765470J1 (URETTUE01)	2378	3108
				6830849H1 (SINTNOR01)	1281	1833
				7376421H1 (ESOGTUE01)	1	582
				GBI.g10277937_edit	1	429
23	3968527CB1	3160	2860-3160, 1-434, 1496-1790	7068888H1 (BRATNOR01)	2443	3108
				7755687H1 (SPLNTUE01)	1389	2085
				7069701H1 (BRAUTDR02)	2679	3160
				7039903H1 (UTRSTMR02)	155	778
				7755687J1 (SPLNTUE01)	2047	2527
				8104892H1 (MIXDDIE02)	43	363
				55052339H1	380	1204
				7097441H1 (BRACDIR02)	2098	2643
				7032041H1 (BRAXTDR12)	1092	1580
				GBI.g8748893_000007. edit	1919	2722
				GNN.g6598919_006.edit	677	1384
				GBI.g8748893_000006_00	1271	1558
				0003.edit		
				g3179340	1705	2116
24	7472732CB1	2848	2653-2848, 2455-2571, 817-1609, 1-72, 150- 195			

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
				GRI.g8748893_000003_00 0004.regenscan.edit	1463	1871
				55061545H1	1	427
				55061546H1	116	728
				2467913T6 (THYRN0T08)	2098	2755
				2467913F6 (THYRN0T08)	1405	1945
				6489280R6 (MIXDUNB01)	2234	2848
25	7476938CB1	3727	1-1490	4107326H1 (BRSTTUT17)	3169	3445
				5958480H1 (BRAINTOT05)	2598	3198
				8242492H1 (BONEUNR01)	1879	2518
				71063602V1	1741	2405
				6932813H1 (SINTTMR02)	667	1257
				7171144H1 (BRSTTMC01)	196	613
				7226459H1 (LUNGTMCO1)	2496	3097
				8144835J1 (MIXDTME01)	424	840
				6799132H1 (COLENOT03)	3188	3727
				8190281H1 (BMARTXN03)	1329	1773
				7963983H1 (SPLNFEA02)	1201	1697
26	8128531CB1	2571	1-925	GNN.g9187761_004.edit	1	559
				3187659H1 (THYMNON04)	2225	2571
				70030270D1	909	1312
				4860138F6 (BRSTTUT22)	953	1509
				2232088T6 (PROSNOT16)	1644	2204
				4341662H1 (BRAUNOT02)	1	266
				6883871J1 (BRAHTDR03)	477	945
				2232088F6 (PROSNOT16)	1991	2478
				4001257T6 (HNT2AZS07)	1372	1963
				8128531H1 (SCOMDICO1)	150	850
27	7476757CB1	1660	1107-1660, 490-803	GNN:g7712233_000033_00 2	667	1660

【表19】

10

20

30

40

表4-3

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
28	266243CB1	2743	1-155, 2720-2743, 2490-2605	55136433H1 7765596J1 (URETUE01) 7629030H1 (GBLADIE01) 71153607V1 7279379H1 (BMARTXE01) 6618283J1 (BRAUTDR03) 7629030J1 (GBLADIE01) GNN:g8575919_008	1 2067 655 1700 1160 563 6 1 1862 1332 2433 580 1 2584 968 2740 1261 826 243 334 421 679 521 925 1015	737 2743 1237 2171 1778 1096 660 1233 2503 2068 3230 1190 669 3231 1611 3239 1615 1096 420 517 606 825 663 1332 1419
29	6585710CB1	3239	1-899, 2163-2217, 2516-2653	72460988D1 72458459D1 71978812V1 72463146D1 72461256D1 71977010V1 72462439D1 71875053V1	1862 1332 2433 580 1 2584 968 2740 1261 826 243 334 421 679 521 925 1015	737 2743 1237 2171 1778 1096 660 1233 2503 2068 3230 1190 669 3231 1611 3239 1615 1096 420 517 606 825 663 1332 1419
30	7483599CB1	1615	749-823, 114-353	g2077361 ENST00000023927 FL7483599_g7708819_000 010_g7293314_1_2-3 FL7483599_g7708819_000 010_g7293314_1_3-4 FL7483599_g7708819_000 010_g7293314_1_4-5 FL7483599_g7708819_000 010_g7293314_1_7 FL7483599_g7708819_000 010_g7293314_1_5-6 FL7483599_g7708819_000 010_g7293314_1_9 g2077387	1261 826 243 334 421 679 521 925 1015	1615 1096 420 517 606 825 663 1332 1419

【表 2 0】

10

20

30

40

表 4-4

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
				GNN.g7417485_000010_00 2	1	1332
31	2507246CB1	1245	922-1245	FL7483599_g7708819_000 010_g7293314_1_1-2 71424096V1 504936R6 (TMLR3DT02) 354532F1 (RATRN0T01) 72229434D1	115 300 739 619 1	333 1021 1245 1237 481
32	3033505CB1	4169	793-1236, 2297-2327, 4074-4169, 3031-3429	6045025J1 (BRABDIR02) 7359286H1 (BRAIFEE05) 6547566H2 (PROSUNT01) 4104913F6 (BRSTTUT17) 4529404H1 (LYMBTXT01) 71059135V1 487605R6 (HNT2AGT01) 6859847H1 (BRAIFEN08) 4324588H1 (TLYMUNT01) 7103133H1 (BRAWTDR02) 7288253H1 (BRAIFER06) 6913384J1 (PITUDIR01) 3033505F6 (TLYMNOT05) 71246947V1	3278 656 55 3125 1 2442 1194 1945 557 877 3584 2564 1354 1864	3674 1046 758 3616 256 3042 1700 2572 810 1301 4169 3174 1913 2523
33	4027693CB1	3440	2635-2665, 558-1546	2866257F6 (KIDNNOT20) 7364929H1 (OVARDIC01) 7617003H1 (KIDNTUE01) 5501287H1 (BRABDIR01) 91485147 70503458V1 7648768J1 (STOMTDE01) 70618525V1	2069 2923 2423 450 1 1705 1193 1489	2700 3440 3031 614 451 2332 1833 2083

10

20

30

40

表4-5

ポリスクレオチド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
				8107676H1 (MIXDDIE02)	582	1038
				7363817H1 (OVARDIC01)	733	1273
				2866257H1 (KIDNNOT20)	2068	2402
				g3147430	211	561
34	7472030CB1	3699	2778-3235, 2347-2550, 1-982, 3286-3699, 1712-2259	1242602R6 (LUNGNOT03) FL7472030_g10445386_g3 07181_1_7-8 FL7472030_g10445386_g3 07181_1_8-9 56004293H1 (FLP600128) 8180328H1 (EYERNON01) GBI_g4508130_g10445386 _edit	1359 2347 2551 1815 941 1	1710 2651 2792 2263 1562 3699
35	7476089CB1	2428	1-861, 1225-2428	FL7476089_g8656012_g59 02966 58016676J1 56003593J1 58007776J1 7289568R6 (BRAIFER06) 1991187F6 (CORPNOT02) 7924964H1 (COLNTUS02) 8059028J1 (LIVRTUE01) 2132191H1 (OVARNOT03) 6437511H1 (BRAENOT02) 7393426R8 (BRABDIE02) 3234007H2 (COLNUCT03) 8219858J2 (SINTFER02) 55120512J1 55120612J1 56000471J1 8228596J1 (BRAUTDR02)	1300 1702 1 897 512 1264 1 616 1995 1412 351 1862 996 252 1325 1 2417	2346 2428 583 1407 967 1808 592 1137 2243 1968 986 2112 1698 1056 2135 498 3044
36	6428177CB1	2243	1052-1089, 1737-2243			
37	7477243CB1	3711	1-673, 2513-3058, 3676-3711			

【表 2 2】

10

20

30

40

表4-6

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
38	7473042CB1		2717	1-607, 1886-2006	6983392F7 (BRAIFER05)	1911	2665
					2863115T6 (KIDNNOT20)	3051	3711
					55155912J1	1730	2139
					5885787F8 (LIVENON08)	2646	3327
					6258661F6 (BMARIXT06)	856	1444
					56008775J1	966	1762
39	7482060CB1		2235	827-859, 1642-2235, 1525-1551	72622070V1	532	1238
					GNN.G7008856_000017_00 2	24	776
					7179755V1	1628	2287
					8195488H2 (BRAINOR03)	336	1087
					72457143D1	2069	2717
					55141001H1	1	253
40	1578772CB1		2563	710-845, 1- 138, 2543- 2563	55061745J1	282	393
					GNN.G9454649_000007_00 0012	428	2065
					6770140R8 (BRAUNOR01)	1	871
					5923423H1 (BRAIFET02)	1932	2235
					6770140F8 (BRAUNOR01)	1416	1648
					5402544H1 (BRAHNOT01)	549	807
					1619843T6 (BRAITUT13)	1940	2563
					70880785V1	769	1319
					6449038H1 (BRAINOC01)	1275	1715
					2083536H1 (UTRSNOT08)	1122	1380
					6536573H1 (OVARIN02)	1	475
					2755946R6 (THPLAZS08)	1445	1833
					6553201H1 (BRAFNON02)	1531	2121
					7239234H1 (BRAINOY02)	238	689

【表 2 3】

10

20

30

40

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
21	1626101CB1	NOSEDIN01
22	2907828CB1	UCMCL5T01
23	3968527CB1	SPLNTUE01
24	7472732CB1	THYRNOT08
25	7476938CB1	LIVRNOT03
26	8128531CB1	THYMNOT08
27	7476757CB1	ESOGTUE01
28	266243CB1	BMARTXE01
29	6585710CB1	SINTNOT21
31	2507246CB1	LATRTUT02
32	3033505CB1	TLYMNOT05
33	4027693CB1	KIDNNOT20
34	7472030CB1	LUNGNOT03
35	7476089CB1	BRAIFER06
36	6428177CB1	BRABDIE02
37	7477243CB1	BMARTXT06
38	7473042CB1	UTRENOT10
39	7482060CB1	BRAUNOR01
40	1578772CB1	BRAITUT12

【表 2 4】

10

20

30

40

表6-1

ライブラリ	ベクター	ライブラリの説明
BMARTXE01	pINCY	この5'パイアスのランダムプライムされたライブラリは、4才の白人女子から除去した転移性骨髄神経芽腫由来の処理された SH-SY5Y 細胞から単離した RNA を用いて作製した(Schering AG)。培地は 10%ウシ胎仔血清を含む MEM/HAM'S F12 であった。約80%コンフルエントに達した後、細胞を 100 マイクロモルの6ヒドロキシドパミン(6-OHDA)で8時間処理した。
BMARTXT06	pINCY	ライブラリは 4 才の白人女子から除去した骨髄神経芽腫腫瘍細胞由来の未処理の SH-SY5Y 細胞系から単離した RNA を用いて作製した。
BRAEDIE02	pINCY	この5'パイアスのランダムプライムされたライブラリは脳血管事故死亡の 57 才の白人男性の脳から除去した腫瘍小脳組織から単離した RNA を用いて作製した。血清検査は陰性だった。患者の病歴は、ハンチントン病、気腫およびタバコの乱用 (40 年間 1 日 3~4 箱)などであった。
BRAIFER06	PCDNA2.1	このランダムプライムされたライブラリは、妊娠 23 週間で死産の左心低形成症胎児(白人男子)から除去された脳組織から単離した RNA を用いて作製した。血清検査は陰性だった。
BRAITUT12	pINCY	ライブラリは、40 才の白人女性の脳腫瘍病変の切除時に、左前頭葉から採取した脳腫瘍組織から単離した RNA を用いて作製した。病理はグレード 4 の大円形細胞性星状細胞腫を示した。
BRAUNOR01	pINCY	このランダムプライムされたライブラリは、アテローム性動脈硬化症により胸部大動脈破裂による出血で死亡の 81 才白人女性から得た線条体、淡蒼球および被殻組織から単離した RNA を用いて作製した。病理学検査では、双方の内部頸動脈に中度のアテローム性動脈硬化、前頭皮質および海馬に微視的な梗塞、および散在性びまん性のアミロイド斑、および神経原線維濃縮体が見られたが、これは年齢と矛盾しない。肉眼で見て、軟膜は上矢状静脈洞に沿った軽度の硝子質化と厚化が見られるのみであった。残りの軟膜は薄く、充血した血管がいくらか見られた。主に前頭極および前頭葉、及び側頭葉(双方)に軽度の萎縮が見られた。微視的には、アルツハイマータイプ II 星状細胞のペリアが新皮質の深層にあった。中央前頭皮質の灰白質の奥深くのニューロンの周りでサテライトシス(satellitosis)が増大していた。扁桃はまれなびまん性斑と神経原線維濃縮体を含んでいた。後部海馬は反応性グリオシスに囲まれたヘモジリン含有マクロファージを伴う微視的な囊胞性の空洞化領域を含んでいた。患者の病歴は、敗血症、胆管炎、術後無気肺、肺炎冠動脈疾患(CAD)、左心室肥大に起因する心肥大、巨脾症、細動脈性腎硬化、結節性コロイド甲状腺腫、気腫、鬱血心不全(CHF)、甲状腺機能低下症および抹消血管疾患を含んでいた。
ESOGTUE01	pINCY	この5'パイアスのランダムプライムされたライブラリは、61 才の白人男性の食道部分切除、近位胃切除、幽門筋層切開術、および限局的リンパ節切除時に取り除いた食道腫瘍組織から単離した RNA を用いて作製した。病理検査は、食道での侵襲性グラーデ3の腺癌を示し、伸長して末端が胃食道境界部に至っていた。腫瘍は筋層を通過して食道周囲と胃周辺の軟組織にまで及んでいた。胃周囲の1つのリンパ節と食道周囲の2つのリンパ節が腫瘍に陽性であった。胃周囲と食道周囲に複数の腫瘍転移が見つけられた。患者は欠乏性貧血と脊髄形成異常を示した。患者の病歴は高脂血症と寛解期にあるタバコ中毒、アル

10

20

30

40

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、アデノイド扁桃摘出術、造鼻術、精管切除、および待切除がある。以前の骨髓吸引では、その年齢としては骨髓が細胞過多であり、実質/脂肪比が95:5であった。骨髓は限局的に高密度の線維性であった。顆粒球前駆体は、正常な成熟と共にわずかに増加していた。芽細胞の推定は5%以上であった。巨核球は増加しており、群の中で異常に見えた。肝臓細胞と肉芽腫はなかった。患者の服用薬剤には Etoposide, Danocrine, Berocin Plus 錠剤、Selenium、ビタミン B6 リン酸、ビタミン E とビタミン C、β カロチンが含まれる。家族の病歴としては、父親にアルコール中毒、アテローム性冠動脈疾患、タイプ II 糖尿病、慢性肝臓病、および一次性心筋症があり、母親に良性高血圧症と脳血管病がある。
KIDN0T20	pINCY	ライブラリは、43 才白人男性の腎尿管切除術、局所リンパ節切除、左片側副腎摘出時に切除された左腎組織から単離された RNA を用いて作製された。関連する腫瘍組織の病理学検査では、グレード 2 の腎細胞癌が見られた。家族の病歴には、アテローム性冠動脈疾患が見られた。
LATRT0T02	pINCY	ライブラリは 43 才白人男性の弁輪形成時に左心房から採取した粘液腫から単離した RNA を用いて作成した。病理は心房性粘液腫を示した。患者の病歴には肺不全、急性心筋梗塞、アテローム性冠動脈疾患、高脂血症および喫煙がある。家族歴には良性高血圧症、急性心筋梗塞、アテローム性冠動脈疾患および II 型糖尿病がある。
LIVR0T03	pINCY	ライブラリは、20 週間の妊娠中にパトシンドローム (13 トリソミー) で死亡した白人男子胎児から採取した肝臓組織の分離 RNA を使用して作製された。
LUNG0T03	PSORT1	ライブラリは 79 才の白人男性の肺組織から単離した RNA を用いて作製した。関連する腫瘍組織の病理学検査では、グレード 4 の癌が見られた。患者の病歴には良性の前立腺腫瘍とアテローム性冠動脈硬化症が含まれている。
NCSEDI01	pINCY	この標準化したライブラリは、プールされた鼻ポリープ組織ライブラリからの 108 万個の独立性クローンから作製した。開始 RNA は 2 人の提供者からのプールされた cDNA から作製された。cDNA は 78 歳の白人男性 (提供者 A) の鼻ポリープ摘出術中に除去された鼻ポリープから単離された mRNA および 1 人の提供者 (提供者 B) からの鼻ポリープから単離された mRNA を使って作製された。病理検査 (A) では鼻ポリープのほかに著しい好酸球増加症が特に上皮組織の深部に見られた。多くの場合好酸球増加症は明らかな壊死になりつつあり、Charcot-Leyden 結晶の著しい堆積が見られた。複数の小さな細胞の島 (islands) で好酸球増加症が侵入した病巣が一部の場所に見られ、その外見上の表面に近いところでは輪郭を失って明らかに壊死が進んでいた。呼吸器の上皮の検査では、多くの場所で表面上皮が失われていて、上皮の周囲に細胞が凝集する傾向が見られた。鼻ポリープは、アレルギー疾患に伴うポリープの変化に典型的な組織学上の所見を示した。患者の病歴は喘息、アレルギー性鼻炎 (ヒスタミンは陽性で、一般の物質は陰性)、肺機能テスト (PFT: 強制肺活量 (FEV) が減少していて、気管支拡張剤を使用した後に増加した) および鼻ポリープなどであった。患者 (A) の病歴には、喘息がある。以前の手術 (A) としては鼻ポリープ摘出術があった。患者は喘息の治療にグルココルチコイドを使っていなかった。このライブラリは、極めて長時間 (48 時間/回の再アニーリングによるハイブリダイゼーションを用いたことを除いては、Soares 他, PNAS (1994) 91:9228-9232

10

20

30

40

表6-3

ライブラリ	ベクター	ライブラリの説明
SINNOT21	pINCY	および Bonaldo 他, Genome Research (1996) 6:791 を応じた条件を用いて1回で標準化した。 ライブラリは無酸素症で死亡した8才の黒人男子から採取した小腸組織から単離した RNA を用いて作製した。血清学は陰生であった。
SPLNTUE01	PCDNA2.1	この5'パイアスのランダムプライムされたライブラリは28歳白人男性の脾臓全摘出術時に切除された脾臓腫瘍組織から単離した RNA を用いて作製した。病理は悪性リンパ腫、びまん性大細胞タイプ、B細胞表現型を示し、ほぼ45個の小結節形成がある脾臓、肝臓および多数のリンパ節にわたって多くの反応性T細胞および著しい肉芽腫反応が示された。
THYMNOT08	pINCY	ライブラリは4ヶ月の白人男子の低温開心術による房室管欠陥の修復および全胸腺摘出術中に取り除かれた胸腺組織から単離した RNA を用いて作製された。病理検査では肉眼で見る限り胸腺が正常であった。患者は先天性の心臓異常、鬱血性心不全およびダウン症を示していた。患者の病歴としては、甲状腺機能検査の異常および早産があった。以前の処置としては右および左の心臓の血管心臓造影法があった。患者の使用薬剤としてはジギキシン、Synthroid およびラジックス等があった。
THYMNOT08	pINCY	ライブラリは13才白人女性の甲状腺全摘出術時に取り除かれた病変左甲状腺から単離した RNA を用いて作成した。病理検査ではリンパ球性甲状腺炎が示された。一致する腫瘍組織の病理学検査では、グレード1の乳頭状癌が見られた。首の右、左および中央線部分の複数のリンパ節は腫瘍について陰性であった。胸腺の断片は良性であった。線維脂肪組織が右下および上方の副甲状腺領域で見つかった。首の右側の複数のリンパ節(6つのうち2つ)が転移性乳頭状癌の微視的な病巣を含んでいた。患者の病歴は、多動性傷害を伴う注意欠陥障害を含んでいた。以前の手術としては、外耳の手術処置などがあった。患者の使用薬剤には Prozac があった。家族の病歴としては、母親に慢性閉塞性喘息があり、祖父母にアルコール濫用、良性の高血圧および抑鬱性障害、兄弟姉妹に多動性傷害を伴う注意欠陥障害があった。
TLYMNOT05	pINCY	ライブラリは非活性化 Th2 細胞から単離された RNA を用いて作製した。これらの細胞は、IL-12 抗免疫体がある IL-4 を持った脾臓の緒の CD4 T 細胞と B7-核酸 COS 細胞から区別した。
UCMCL5T01	PBLUESCRIPT	ライブラリは、12人の被験者の臍帯の血液から得られた単核細胞から単離された RNA を使って作製された。細胞を IL-5 と共に12日培養してから、RNA をプールされた溶解物から得た。
UTRENOT10	pINCY	ライブラリは3人の成人女性の子宮内膜生検時に取り除かれたプールされた子宮内膜組織から単離した RNA を用いて作製した。病理検査は正常な子宮内膜を示していた。3人の患者は全てベータ3インテグリンについて陽性であった。

10

20

30

40

表7-1

プログラム	説明	引用文献	パラメータ関連
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キングするプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	アミノ酸または核酸配列の比較、注釈に有用なアア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ペーシングク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tblastx。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) (1997) Nucleic Acids Res. 25:3389-3402.	ESTs確率値=1.0E-8以下 完全長配列確率値=1.0E-10以下
FASTA	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、tfastxおよびfssearch。	Pearson, W.R.及びJ.D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; およびSmith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs fasta E 値=1.06E-6 構築された ESTs fasta 同一性=95%以上 一致した長さ=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グering領域を検索するBLocks IMPROVED Searcher	Henikoff, S.及びJ.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G.及びS. Henikoff (1996) Methods Enzymol. 266:88-105. 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; Sonhammer, E.L.L. 他 (1988) Nucleic Acids Res. 26:320-322; Durbin, R. 他. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, 1-350ページ。	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

40

表7-2

プログラム	説明	引用文献	パラメータ閾値
ProfileScan	Prositeで定義された配列パターンと一致するタンパク質配列の構造的モチーフおよび配列モチーフを検索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4:61-66; Gribskov, M. 他 (1989) Methods Enzymol.183:146-159; Bairoch, A. 他. (1997) Nucleic Acids Res. 25:217-221.	標準化された質スコアと特定の 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クインサートレースを調べる塩基コールングアルゴリズム。	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Phils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. 及び S. Audic (1997) CABIOS 12:431-438.	スコア=3.5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.	
TMHMMER	隠れMarkov 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。	Somhammer, 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。	Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual. パージョン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
23 May 2002 (23.05.2002)

PCT

(10) International Publication Number
WO 02/40541 A2

(51) International Patent Classification: C07K 14/705

(21) International Application Number: PCT/US91/66055

(22) International Filing Date: 25 October 2001 (25.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/213,989 27 October 2000 (27.10.2000) US
60/245,904 3 November 2000 (03.11.2000) US
60/247,673 9 November 2000 (09.11.2000) US
60/249,661 17 November 2000 (17.11.2000) US
60/252,232 20 November 2000 (20.11.2000) US
60/250,790 1 December 2000 (01.12.2000) US

95056 (US); AZIMZAI, Yalda (US/US); 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US); KHAN, Farrah, A. (IN/US); 9445 Harrison Street, Des Plaines, IL 60016 (US); THANGAVELU, Kavitha (IN/US); 1950 Montecito Avenue, #23, Mountain View, CA 94043 (US); THORNTON, Michael (US/US); 9 Midway Road, Woodside, CA 94062 (US); LU, Dwyang, Aina, M. (US/US); 233 Coy Drive, San Jose, CA 95123 (US); TRIBOULEY, Catherine, M. (FR/US); 1121 Tennessee Street, #5, San Francisco, CA 94107 (US); WARREN, Bridget, A. (US/US); 10130 Parkwood Drive, #2, Cupertino, CA 95014 (US); ISON, Craig, II (US/US); 1242 Weathersfield Way, San Jose, CA 95118 (US); DAS, Debopriya (IN/US); 1267 Parkington Avenue, Sunnyvale, CA 94087 (US); RAUMANN, Brigitte, E. (US/US); 5801 South Dorchester Avenue #3B, Chicago, IL 60637 (US); POLICKY, Jennifer, L. (US/US); 1511 Court Jarvis Court, San Jose, CA 95118 (US); KEARNEY, Liam (US/US); 50 Woodside Avenue, San Francisco, CA 94127 (US).

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

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

(72) Inventors: and

(75) Inventors/Applicants (for US only): TANG, Yi, Tom (US/US); 4250 Ranwick Court, San Jose, CA 95118 (US); YUE, Henry (US/US); 826 Lois Avenue, Sunnyvale, CA 94087 (US); NGUYEN, Daniel, B. (IN/US); 1403 Ridgewood Drive, San Jose, CA 95118 (US); HAFALIA, April, J., A. (US/US); 2227 Calle de Primavera, Santa Clara, CA 95054 (US); ELLIOTT, Vicki, S. (US/US); 5770 Pullina Plaza Way, San Jose, CA 95121 (US); LU, Yan (CN/US); 3885 Corrina Way, Palo Alto, CA 94303 (US); WALLA, Narinder, K. (US/US); 890 Davis Street, #305, San Leandro, CA 94577 (US); YAO, Monique, G. (US/US); 1189 Woodgate Drive, Carmel, IN 46033 (US); BAUGHN, Marilyn, R. (US/US); 14244 Santiago Road, San Leandro, CA 94577 (US); GANDHI, Anurag, R. (US/US); 705 5th Avenue, San Francisco, CA 94118 (US); DING, Li (CN/US); 3353 Alton Street, 4146, Palo Alto, CA 94306 (US); SANJANWALA, Madhusudan (US/US); 210 Silvia Court, Los Altos, CA 94024 (US); RAMKUMAR, Jayalaxmi (IN/US); 34359 Maybird Circle, Fremont, CA 94555 (US); ARVIZO, Chandra (US/US); 490 Sherwood Way, #1, Menlo Park, CA 94025 (US); GIETZEN, Kimberly, J. (US/US); 691 Los Hozos Drive, San Jose, CA 95123 (US); LAL, Preeti, G. (IN/US); P.O. Box 5142, Santa Clara, CA

(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, DO, DZ, EC, EE, EG, FI, GB, GR, GU, HK, 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, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TH, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW); Eurasian patent (AM, AZ, BY, EG, KG, KZ, MD, RU, TJ, TM); European patent (AL, 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, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

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

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

(54) Title: TRANSPORTERS AND ION CHANNELS

(57) Abstract: The invention provides human transporters and ion channels (TRICH) and polynucleotides which identify and encode TRICH. 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 TRICH.

WO 02/40541 A2

WO 02/40541

PCT/US01/46055

TRANSPORTERS AND ION CHANNELS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of transporters and ion channels and to the use of these sequences in the diagnosis, treatment, and prevention of transport, neurological, muscle, immunological and cell proliferative disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of transporters and ion channels.

BACKGROUND OF THE INVENTION

Eukaryotic cells are surrounded and subdivided into functionally distinct organelles by hydrophobic lipid bilayer membranes which are highly impermeable to most polar molecules. Cells and organelles require transport proteins to import and export essential nutrients and metal ions including K^+ , NH_4^+ , P_i , SO_4^{2-} , sugars, and vitamins, as well as various metabolic waste products. Transport proteins also play roles in antibiotic resistance, toxin secretion, ion balance, synaptic neurotransmission, kidney function, intestinal absorption, tumor growth, and other diverse cell functions (Griffith, J. and C. Sansom (1998) The Transporter Facts Book, Academic Press, San Diego CA, pp. 3-29). Transport can occur by a passive concentration-dependent mechanism, or can be linked to an energy source such as ATP hydrolysis or an ion gradient. Proteins that function in transport include carrier proteins, which bind to a specific solute and undergo a conformational change that translocates the bound solute across the membrane, and channel proteins, which form hydrophilic pores that allow specific solutes to diffuse through the membrane down an electrochemical solute gradient.

Carrier proteins which transport a single solute from one side of the membrane to the other are called uniporters. In contrast, coupled transporters link the transfer of one solute with simultaneous or sequential transfer of a second solute, either in the same direction (symport) or in the opposite direction (antiport). For example, intestinal and kidney epithelium contains a variety of symporter systems driven by the sodium gradient that exists across the plasma membrane. Sodium moves into the cell down its electrochemical gradient and brings the solute into the cell with it. The sodium gradient that provides the driving force for solute uptake is maintained by the ubiquitous Na^+/K^+ ATPase system. Sodium-coupled transporters include the mammalian glucose transporter (SGLT1), iodide transporter (NIS), and multivitamin transporter (SMVT). All three transporters have twelve putative transmembrane segments, extracellular glycosylation sites, and cytoplasmically-oriented N- and C-termini. NIS plays a crucial role in the evaluation, diagnosis, and treatment of various thyroid pathologies because it is the molecular basis for radioiodide thyroid-imaging

WO 02/40541

PCT/US01/46055

techniques and for specific targeting of radioisotopes to the thyroid gland (Levy, O. et al. (1997) Proc. Natl. Acad. Sci. USA 94:5568-5573). SMVT is expressed in the intestinal mucosa, kidney, and placenta, and is implicated in the transport of the water-soluble vitamins, e.g., biotin and pantothenate (Prasad, P.D. et al. (1998) J. Biol. Chem. 273:7501-7506).

5 One of the largest families of transporters is the major facilitator superfamily (MFS), also called the uniporter-symporter-antiporter family. MFS transporters are single polypeptide carriers that transport small solutes in response to ion gradients. Members of the MFS are found in all classes of living organisms, and include transporters for sugars, oligosaccharides, phosphates, nitrates, nucleosides, monocarboxylates, and drugs. MFS transporters found in eukaryotes all have a structure
10 comprising 12 transmembrane segments (Pao, S.S. et al. (1998) Microbiol. Molec. Biol. Rev. 62:1-34). The largest family of MFS transporters is the sugar transporter family, which includes the seven glucose transporters (GLUT1-GLUT7) found in humans that are required for the transport of glucose and other hexose sugars. These glucose transport proteins have unique tissue distributions and physiological functions. GLUT1 provides many cell types with their basal glucose requirements and
15 transports glucose across epithelial and endothelial barrier tissues; GLUT2 facilitates glucose uptake or efflux from the liver; GLUT3 regulates glucose supply to neurons; GLUT4 is responsible for insulin-regulated glucose disposal; and GLUT5 regulates fructose uptake into skeletal muscle. Defects in glucose transporters are involved in a recently identified neurological syndrome causing infantile seizures and developmental delay, as well as glycogen storage disease, Fanconi-Bickel
20 syndrome, and non-insulin-dependent diabetes mellitus (Mueckler, M. (1994) Eur. J. Biochem. 219:713-725; Longo, N. and L.J. Elsas (1998) Adv. Pediatr. 45:293-313).

Monocarboxylate anion transporters are proton-coupled symporters with a broad substrate specificity that includes L-lactate, pyruvate, and the ketone bodies acetate, acetoacetate, and beta-hydroxybutyrate. At least seven isoforms have been identified to date. The isoforms are
25 predicted to have twelve transmembrane (TM) helical domains with a large intracellular loop between TM6 and TM7, and play a critical role in maintaining intracellular pH by removing the protons that are produced stoichiometrically with lactate during glycolysis. The best characterized H⁺-monocarboxylate transporter is that of the erythrocyte membrane, which transports L-lactate and a wide range of other aliphatic monocarboxylates. Other cells possess H⁺-linked monocarboxylate
30 transporters with differing substrate and inhibitor selectivities. In particular, cardiac muscle and tumor cells have transporters that differ in their K_m values for certain substrates, including stereoselectivity for L- over D-lactate, and in their sensitivity to inhibitors. There are Na⁺-monocarboxylate cotransporters on the luminal surface of intestinal and kidney epithelia, which allow the uptake of lactate, pyruvate, and ketone bodies in these tissues. In addition, there are
35 specific and selective transporters for organic cations and organic anions in organs including the

WO 02/40541

PCT/US01/46055

kidney, intestine and liver. Organic anion transporters are selective for hydrophobic, charged molecules with electron-attracting side groups. Organic cation transporters, such as the ammonium transporter, mediate the secretion of a variety of drugs and endogenous metabolites, and contribute to the maintenance of intercellular pH (Poole, R.C. and A.P. Halestrap (1993) *Am. J. Physiol.*

5 264:C761-C782; Price, N.T. et al. (1998) *Biochem. J.* 329:321-328; and Martinelle, K. and I. Haggstrom (1993) *J. Biotechnol.* 30:339-350).

ATP-binding cassette (ABC) transporters are members of a superfamily of membrane proteins that transport substances ranging from small molecules such as ions, sugars, amino acids, peptides, and phospholipids, to lipopeptides, large proteins, and complex hydrophobic drugs. ABC
10 transporters consist of four modules: two nucleotide-binding domains (NBD), which hydrolyze ATP to supply the energy required for transport, and two membrane-spanning domains (MSD), each containing six putative transmembrane segments. These four modules may be encoded by a single gene, as is the case for the cystic fibrosis transmembrane regulator (CFTR), or by separate genes. When encoded by separate genes, each gene product contains a single NBD and MSD. These "half-
15 molecules" form homo- and heterodimers, such as Tap1 and Tap2, the endoplasmic reticulum-based major histocompatibility (MHC) peptide transport system. Several genetic diseases are attributed to defects in ABC transporters, such as the following diseases and their corresponding proteins: cystic fibrosis (CFTR, an ion channel), adrenoleukodystrophy (adrenoleukodystrophy protein, ALDP), Zellweger syndrome (peroxisomal membrane protein-70, PMP70), and hyperinsulinemic
20 hypoglycemia (sulfonylurea receptor, SUR). Overexpression of the multidrug resistance (MDR) protein, another ABC transporter, in human cancer cells makes the cells resistant to a variety of cytotoxic drugs used in chemotherapy (Taglicht, D. and S. Michaelis (1998) *Meth. Enzymol.* 292:130-162).

A number of metal ions such as iron, zinc, copper, cobalt, manganese, molybdenum,
25 selenium, nickel, and chromium are important as cofactors for a number of enzymes. For example, copper is involved in hemoglobin synthesis, connective tissue metabolism, and bone development, by acting as a cofactor in oxidoreductases such as superoxide dismutase, ferroxidase (ceruloplasmin), and lysyl oxidase. Copper and other metal ions must be provided in the diet, and are absorbed by transporters in the gastrointestinal tract. Plasma proteins transport the metal ions to the liver and
30 other target organs, where specific transporters move the ions into cells and cellular organelles as needed. Imbalances in metal ion metabolism have been associated with a number of disease states (Danks, D.M. (1986) *J. Med. Genet.* 23:99-106).

Transport of fatty acids across the plasma membrane can occur by diffusion, a high capacity, low affinity process. However, under normal physiological conditions a significant fraction of fatty
35 acid transport appears to occur via a high affinity, low capacity protein-mediated transport process.

WO 02/40541

PCT/US01/46055

Fatty acid transport protein (FATP), an integral membrane protein with four transmembrane segments, is expressed in tissues exhibiting high levels of plasma membrane fatty acid flux, such as muscle, heart, and adipose. Expression of FATP is upregulated in 3T3-L1 cells during adipose conversion, and expression in COS7 fibroblasts elevates uptake of long-chain fatty acids (Hui, T.Y. et al. (1998) *J. Biol. Chem.* 273:27420-27429).

The lipocalin superfamily constitutes a phylogenetically conserved group of more than forty proteins that function as extracellular ligand-binding proteins which bind and transport small hydrophobic molecules. Members of this family function as carriers of retinoids, odorants, chromophores, pheromones, allergens, and sterols, and in a variety of processes including nutrient transport, cell growth regulation, immune response, and prostaglandin synthesis. A subset of these proteins may be multifunctional, serving as either a biosynthetic enzyme or as a specific enzyme inhibitor. (Tanaka, T. et al. (1997) *J. Biol. Chem.* 272:15789-15795; and van't Hof, W. et al. (1997) *J. Biol. Chem.* 272:1837-1841.)

Members of the lipocalin family display unusually low levels of overall sequence conservation. Pairwise sequence identity often falls below 20%. Sequence similarity between family members is limited to conserved cysteines which form disulfide bonds and three motifs which form a juxtaposed cluster that functions as a target cell recognition site. The lipocalins share an eight stranded, anti-parallel beta-sheet which folds back on itself to form a continuously hydrogen-bonded beta-barrel. The pocket formed by the barrel functions as an internal ligand binding site. Seven loops (L1 to L7) form short beta-hairpins, except loop L1 which is a large omega loop that forms a lid to partially close the internal ligand-binding site (Flower (1996) *Biochem. J.* 318:1-14).

Lipocalins are important transport molecules. Each lipocalin associates with a particular ligand and delivers that ligand to appropriate target sites within the organism. Retinol-binding protein (RBP), one of the best characterized lipocalins, transports retinol from stores within the liver to target tissues. Apolipoprotein D (apo D), a component of high density lipoproteins (HDLs) and low density lipoproteins (LDLs), functions in the targeted collection and delivery of cholesterol throughout the body. Lipocalins are also involved in cell regulatory processes. Apo D, which is identical to gross-cystic-disease-fluid protein (GCDFP)-24, is a progesterone/pregnenolone-binding protein expressed at high levels in breast cyst fluid. Secretion of apo D in certain human breast cancer cell lines is accompanied by reduced cell proliferation and progression of cells to a more differentiated phenotype. Similarly, apo D and another lipocalin, α_1 -acid glycoprotein (AGP), are involved in nerve cell regeneration. AGP is also involved in anti-inflammatory and immunosuppressive activities. AGP is one of the positive acute-phase proteins (APP); circulating levels of AGP increase in response to stress and inflammatory stimulation. AGP accumulates at sites of inflammation where it inhibits platelet and neutrophil activation and inhibits phagocytosis. The

WO 02/40541

PCT/US01/46055

immunomodulatory properties of AGP are due to glycosylation. AGP is 40% carbohydrate, making it unusually acidic and soluble. The glycosylation pattern of AGP changes during acute-phase response, and deglycosylated AGP has no immunosuppressive activity (Flower (1994) FEBS Lett. 354:7-11; Flower (1996) supra).

- 5 The lipocalin superfamily also includes several animal allergens, including the mouse major urinary protein (mMUP), the rat α -2-microglobulin (rA2U), the bovine β -lactoglobulin (B β), the cockroach allergen (Bla g4), bovine dander allergen (Bos d2), and the major horse allergen, designated *Equus caballus* allergen 1 (Equ c1). Equ c1 is a powerful allergen responsible for about 80% of anti-horse IgE antibody response in patients who are chronically exposed to horse allergens.
- 10 It appears that lipocalins may contain a common structure that is able to induce the IgE response (Gregoire, C. et al., (1996) J. Biol. Chem. 271:32951-32959).

Lipocalins are used as diagnostic and prognostic markers in a variety of disease states. The plasma level of AGP is monitored during pregnancy and in diagnosis and prognosis of conditions including cancer chemotherapy, renal dysfunction, myocardial infarction, arthritis, and multiple
15 sclerosis. RBP is used clinically as a marker of tubular reabsorption in the kidney, and apo D is a marker in gross cystic breast disease (Flower (1996) supra). Additionally, the use of lipocalin animal allergens may help in the diagnosis of allergic reactions to horses (Gregoire supra), pigs, cockroaches, mice and rats.

Mitochondrial carrier proteins are transmembrane-spanning proteins which transport ions and
20 charged metabolites between the cytosol and the mitochondrial matrix. Examples include the ADP, ATP carrier protein; the 2-oxoglutarate/malate carrier; the phosphate carrier protein; the pyruvate carrier; the dicarboxylate carrier which transports malate, succinate, fumarate, and phosphate; the tricarboxylate carrier which transports citrate and malate; and the Grave's disease carrier protein, a protein recognized by IgG in patients with active Grave's disease, an autoimmune disorder resulting
25 in hyperthyroidism. Proteins in this family consist of three tandem repeats of an approximately 100 amino acid domain, each of which contains two transmembrane regions (Stryer, L. (1995) Biochemistry, W.H. Freeman and Company, New York NY, p. 551; PROSITE PDOC00189 Mitochondrial energy transfer proteins signature; Online Mendelian Inheritance in Man (OMIM) *275000 Graves Disease).

- 30 This class of transporters also includes the mitochondrial uncoupling proteins, which create proton leaks across the inner mitochondrial membrane, thus uncoupling oxidative phosphorylation from ATP synthesis. The result is energy dissipation in the form of heat. Mitochondrial uncoupling proteins have been implicated as modulators of thermoregulation and metabolic rate, and have been proposed as potential targets for drugs against metabolic diseases such as obesity (Ricquier, D. et al.
35 (1999) J. Int. Med. 245:637-642).

WO 02/40541

PCT/US01/46055

Ion Channels

The electrical potential of a cell is generated and maintained by controlling the movement of ions across the plasma membrane. The movement of ions requires ion channels, which form ion-selective pores within the membrane. There are two basic types of ion channels, ion transporters and gated ion channels. Ion transporters utilize the energy obtained from ATP hydrolysis to actively transport an ion against the ion's concentration gradient. Gated ion channels allow passive flow of an ion down the ion's electrochemical gradient under restricted conditions. Together, these types of ion channels generate, maintain, and utilize an electrochemical gradient that is used in 1) electrical impulse conduction down the axon of a nerve cell, 2) transport of molecules into cells against concentration gradients, 3) initiation of muscle contraction, and 4) endocrine cell secretion.

Ion Transporters

Ion transporters generate and maintain the resting electrical potential of a cell. Utilizing the energy derived from ATP hydrolysis, they transport ions against the ion's concentration gradient. These transmembrane ATPases are divided into three families. The phosphorylated (P) class ion transporters, including Na⁺-K⁺ ATPase, Ca²⁺-ATPase, and H⁺-ATPase, are activated by a phosphorylation event. P-class ion transporters are responsible for maintaining resting potential distributions such that cytosolic concentrations of Na⁺ and Ca²⁺ are low and cytosolic concentration of K⁺ is high. The vacuolar (V) class of ion transporters includes H⁺ pumps on intracellular organelles, such as lysosomes and Golgi. V-class ion transporters are responsible for generating the low pH within the lumen of these organelles that is required for function. The coupling factor (F) class consists of H⁺ pumps in the mitochondria. F-class ion transporters utilize a proton gradient to generate ATP from ADP and inorganic phosphate (P_i).

The P-ATPases are hexamers of a 100 kD subunit with ten transmembrane domains and several large cytoplasmic regions that may play a role in ion binding (Scarborough, G.A. (1999) *Curr. Opin. Cell Biol.* 11:517-522). The V-ATPases are composed of two functional domains: the V₁ domain, a peripheral complex responsible for ATP hydrolysis; and the V₀ domain, an integral complex responsible for proton translocation across the membrane. The F-ATPases are structurally and evolutionarily related to the V-ATPases. The F-ATPase F₀ domain contains 12 copies of the c subunit, a highly hydrophobic protein composed of two transmembrane domains and containing a single buried carboxyl group in TM2 that is essential for proton transport. The V-ATPase V₀ domain contains three types of homologous c subunits with four or five transmembrane domains and the essential carboxyl group in TM4 or TM3. Both types of complex also contain a single a subunit that may be involved in regulating the pH dependence of activity (Forgac, M. (1999) *J. Biol. Chem.* 274:12951-12954).

The resting potential of the cell is utilized in many processes involving carrier proteins and

WO 02/40541

PCT/US01/46055

gated ion channels. Carrier proteins utilize the resting potential to transport molecules into and out of the cell. Amino acid and glucose transport into many cells is linked to sodium ion co-transport (symport) so that the movement of Na^+ down an electrochemical gradient drives transport of the other molecule up a concentration gradient. Similarly, cardiac muscle links transfer of Ca^{2+} out of the cell with transport of Na^+ into the cell (antiport).

Gated Ion Channels

Gated ion channels control ion flow by regulating the opening and closing of pores. The ability to control ion flux through various gating mechanisms allows ion channels to mediate such diverse signaling and homeostatic functions as neuronal and endocrine signaling, muscle contraction, fertilization, and regulation of ion and pH balance. Gated ion channels are categorized according to the manner of regulating the gating function. Mechanically-gated channels open their pores in response to mechanical stress; voltage-gated channels (e.g., Na^+ , K^+ , Ca^{2+} , and Cl^- channels) open their pores in response to changes in membrane potential; and ligand-gated channels (e.g., acetylcholine-, serotonin-, and glutamate-gated cation channels, and GABA- and glycine-gated chloride channels) open their pores in the presence of a specific ion, nucleotide, or neurotransmitter. The gating properties of a particular ion channel (i.e., its threshold for and duration of opening and closing) are sometimes modulated by association with auxiliary channel proteins and/or post translational modifications, such as phosphorylation.

Mechanically-gated or mechanosensitive ion channels act as transducers for the senses of touch, hearing, and balance, and also play important roles in cell volume regulation, smooth muscle contraction, and cardiac rhythm generation. A stretch-inactivated channel (SIC) was recently cloned from rat kidney. The SIC channel belongs to a group of channels which are activated by pressure or stress on the cell membrane and conduct both Ca^{2+} and Na^+ (Suzuki, M. et al. (1999) *J. Biol. Chem.* 274:6330-6335).

The pore-forming subunits of the voltage-gated cation channels form a superfamily of ion channel proteins. The characteristic domain of these channel proteins comprises six transmembrane domains (S1-S6), a pore-forming region (P) located between S5 and S6, and intracellular amino and carboxy termini. In the Na^+ and Ca^{2+} subfamilies, this domain is repeated four times, while in the K^+ channel subfamily, each channel is formed from a tetramer of either identical or dissimilar subunits.

The P region contains information specifying the ion selectivity for the channel. In the case of K^+ channels, a GYG tripeptide is involved in this selectivity (Ishii, T.M. et al. (1997) *Proc. Natl. Acad. Sci. USA* 94:11651-11656).

Voltage-gated Na^+ and K^+ channels are necessary for the function of electrically excitable cells, such as nerve and muscle cells. Action potentials, which lead to neurotransmitter release and muscle contraction, arise from large, transient changes in the permeability of the membrane to Na^+

WO 02/40541

PCT/US01/46055

and K⁺ ions. Depolarization of the membrane beyond the threshold level opens voltage-gated Na⁺ channels. Sodium ions flow into the cell, further depolarizing the membrane and opening more voltage-gated Na⁺ channels, which propagates the depolarization down the length of the cell. Depolarization also opens voltage-gated potassium channels. Consequently, potassium ions flow outward, which leads to repolarization of the membrane. Voltage-gated channels utilize charged residues in the fourth transmembrane segment (S4) to sense voltage change. The open state lasts only about 1 millisecond, at which time the channel spontaneously converts into an inactive state that cannot be opened irrespective of the membrane potential. Inactivation is mediated by the channel's N-terminus, which acts as a plug that closes the pore. The transition from an inactive to a closed state requires a return to resting potential.

Voltage-gated Na⁺ channels are heterotrimeric complexes composed of a 260 kDa pore-forming α subunit that associates with two smaller auxiliary subunits, β 1 and β 2. The β 2 subunit is an integral membrane glycoprotein that contains an extracellular Ig domain, and its association with α and β 1 subunits correlates with increased functional expression of the channel, a change in its gating properties, as well as an increase in whole cell capacitance due to an increase in membrane surface area (Isom, L.L., et al. (1995) *Cell* 83:433-442).

Non voltage-gated Na⁺ channels include the members of the amiloride-sensitive Na⁺ channel/degenerin (NaC/DEG) family. Channel subunits of this family are thought to consist of two transmembrane domains flanking a long extracellular loop, with the amino and carboxyl termini located within the cell. The NaC/DEG family includes the epithelial Na⁺ channel (ENaC) involved in Na⁺ reabsorption in epithelia including the airway, distal colon, cortical collecting duct of the kidney, and exocrine duct glands. Mutations in ENaC result in pseudohypoaldosteronism type 1 and Liddle's syndrome (pseudohyperaldosteronism). The NaC/DEG family also includes the recently characterized H⁺-gated cation channels or acid-sensing ion channels (ASIC). ASIC subunits are expressed in the brain and form heteromultimeric Na⁺-permeable channels. These channels require acid pH fluctuations for activation. ASIC subunits show homology to the degenerins, a family of mechanically-gated channels originally isolated from *C. elegans*. Mutations in the degenerins cause neurodegeneration. ASIC subunits may also have a role in neuronal function, or in pain perception, since tissue acidosis causes pain (Waldmann, R. and M. Lazdunski (1998) *Curr. Opin. Neurobiol.* 8:418-424; Eglon, R.M. et al. (1999) *Trends Pharmacol. Sci.* 20:337-342).

K⁺ channels are located in all cell types, and may be regulated by voltage, ATP concentration, or second messengers such as Ca²⁺ and cAMP. In non-excitabile tissue, K⁺ channels are involved in protein synthesis, control of endocrine secretions, and the maintenance of osmotic equilibrium across membranes. In neurons and other excitable cells, in addition to regulating action potentials and repolarizing membranes, K⁺ channels are responsible for setting the resting membrane

WO 02/40541

PCT/US01/46055

potential. The cytosol contains non-diffusible anions and, to balance this net negative charge, the cell contains a Na⁺-K⁺ pump and ion channels that provide the redistribution of Na⁺, K⁺, and Cl⁻. The pump actively transports Na⁺ out of the cell and K⁺ into the cell in a 3:2 ratio. Ion channels in the plasma membrane allow K⁺ and Cl⁻ to flow by passive diffusion. Because of the high negative charge within the cytosol, Cl⁻ flows out of the cell. The flow of K⁺ is balanced by an electromotive force pulling K⁺ into the cell, and a K⁺ concentration gradient pushing K⁺ out of the cell. Thus, the resting membrane potential is primarily regulated by K⁺ flow (Salkoff, L. and T. Jørgensen (1995) *Neuron* 15:489-492).

Potassium channel subunits of the *Shaker*-like superfamily all have the characteristic six transmembrane/1 pore domain structure. Four subunits combine as homo- or heterotetramers to form functional K channels. These pore-forming subunits also associate with various cytoplasmic β subunits that alter channel inactivation kinetics. The *Shaker*-like channel family includes the voltage-gated K⁺ channels as well as the delayed rectifier type channels such as the human ether-a-go-go related gene (HERG) associated with long QT, a cardiac dysrhythmia syndrome (Curran, M.E. (1998) *Curr. Opin. Biotechnol.* 9:565-572; Kaczorowski, G.J. and M.L. Garcia (1999) *Curr. Opin. Chem. Biol.* 3:448-458).

A second superfamily of K⁺ channels is composed of the inward rectifying channels (Kir). Kir channels have the property of preferentially conducting K⁺ currents in the inward direction. These proteins consist of a single potassium selective pore domain and two transmembrane domains, which correspond to the fifth and sixth transmembrane domains of voltage-gated K⁺ channels. Kir subunits also associate as tetramers. The Kir family includes ROMK1, mutations in which lead to Bartter syndrome, a renal tubular disorder. Kir channels are also involved in regulation of cardiac pacemaker activity, seizures and epilepsy, and insulin regulation (Doupnik, C.A. et al. (1995) *Curr. Opin. Neurobiol.* 5:268-277; Curran, *supra*).

The recently recognized TWIK K⁺ channel family includes the mammalian TWIK-1, TREK-1 and TASK proteins. Members of this family possess an overall structure with four transmembrane domains and two P domains. These proteins are probably involved in controlling the resting potential in a large set of cell types (Duprat, F. et al. (1997) *EMBO J* 16:5464-5471).

The voltage-gated Ca²⁺ channels have been classified into several subtypes based upon their electrophysiological and pharmacological characteristics. L-type Ca²⁺ channels are predominantly expressed in heart and skeletal muscle where they play an essential role in excitation-contraction coupling. T-type channels are important for cardiac pacemaker activity, while N-type and P/Q-type channels are involved in the control of neurotransmitter release in the central and peripheral nervous system. The L-type and N-type voltage-gated Ca²⁺ channels have been purified and, though their functions differ dramatically, they have similar subunit compositions. The channels are composed of

WO 02/40541

PCT/US01/46055

three subunits. The α_1 subunit forms the membrane pore and voltage sensor, while the $\alpha_2\beta$ and β subunits modulate the voltage-dependence, gating properties, and the current amplitude of the channel. These subunits are encoded by at least six α_1 , one $\alpha_2\beta$, and four β genes. A fourth subunit, γ , has been identified in skeletal muscle (Walker, D. et al. (1998) *J. Biol. Chem.* 273:2361-2367;

5 McCleskey, B.W. (1994) *Curr. Opin. Neurobiol.* 4:304-312).

The high-voltage-activated Ca^{2+} channels that have been characterized biochemically include complexes of a pore-forming α_1 subunit of approximately 190-250 kDa; a transmembrane complex of α_2 and β subunits; an intracellular β subunit; and in some cases a transmembrane γ subunit. A variety of α_1 subunits, $\alpha_2\beta$ complexes, β subunits,

10 and γ subunits are known. The Cav1 family of α_1 subunits conduct L-type Ca^{2+} currents, which initiate muscle contraction, endocrine secretion, and gene transcription, and are regulated primarily by second messenger-activated protein phosphorylation pathways. The Cav2 family of α_1 subunits conduct N-type, P/Q-type, and R-type Ca^{2+} currents, which initiate rapid synaptic transmission and are regulated primarily by direct interaction with G proteins and SNARE proteins
15 and secondarily by protein phosphorylation. The Cav3 family of α_1 subunits conduct T-type Ca^{2+} currents, which are activated and inactivated more rapidly and at more negative membrane potentials than other Ca^{2+} current types. The distinct structures and patterns of regulation of these three families of Ca^{2+} channels provide an array of Ca^{2+} entry pathways in response to changes in membrane potential and a range of possibilities for regulation of Ca^{2+} entry by second messenger
20 pathways and interacting proteins (Catterall, W.A. (2000) *Annu. Rev. Cell Dev. Biol.* 16:521-555).

The α_2 subunit of the voltage-gated Ca^{2+} -channel may include one or more Cache domains. An extracellular Cache domain may be fused to an intracellular catalytic domain, such as the histidine kinase, PP2C phosphatase, GGDH (a predicted diguanylate cyclase), HD-GYP (a predicted phosphodiesterase) or adenylyl cyclase domain, or to a noncatalytic domain, like the
25 methyl-accepting, DNA-binding winged helix-turn-helix, GAF, PAS or HAMP (a domain found in histidine kinases, adenylyl cyclases, ethyl-binding proteins and phosphatases). Small molecules are bound via the Cache domain and this signal is converted into diverse outputs depending on the intracellular domains (Anantharaman, V. and Aravind, L. (2000) *Trends Biochem. Sci.* 25:535-537).

The transient receptor family (Trp) of calcium ion channels are thought to mediate

30 capacitative calcium entry (CCE). CCE is the Ca^{2+} influx into cells to resupply Ca^{2+} stores depleted by the action of inositol triphosphate (IP3) and other agents in response to numerous hormones and growth factors. Trp and Trp-like were first cloned from *Drosophila* and have similarity to voltage gated Ca^{2+} channels in the S3 through S6 regions. This suggests that Trp and/or related proteins may form mammalian CCE channels (Zhu, X. et al. (1996) *Cell* 85:661-671; Boulay, G. et al. (1997) *J. Biol. Chem.* 272:29672-29680). Melastatin is a gene isolated in both the mouse and human, whose
35

WO 02/40541

PCT/US01/46055

expression in melanoma cells is inversely correlated with melanoma aggressiveness *in vivo*. The human cDNA transcript corresponds to a 1533-amino acid protein having homology to members of the Trp family. It has been proposed that the combined use of malastatin mRNA expression status and tumor thickness might allow for the determination of subgroups of patients at both low and high risk for developing metastatic disease (Duncan, L.M. et al (2001) J. Clin. Oncol. 19:568-576).

Chloride channels are necessary in endocrine secretion and in regulation of cytosolic and organelle pH. In secretory epithelial cells, Cl^- enters the cell across a basolateral membrane through an Na^+ , K^+ / Cl^- cotransporter, accumulating in the cell above its electrochemical equilibrium concentration. Secretion of Cl^- from the apical surface, in response to hormonal stimulation, leads to flow of Na^+ and water into the secretory lumen. The cystic fibrosis transmembrane conductance regulator (CFTR) is a chloride channel encoded by the gene for cystic fibrosis, a common fatal genetic disorder in humans. CFTR is a member of the ABC transporter family, and is composed of two domains each consisting of six transmembrane domains followed by a nucleotide-binding site. Loss of CFTR function decreases transepithelial water secretion and, as a result, the layers of mucus that coat the respiratory tree, pancreatic ducts, and intestine are dehydrated and difficult to clear. The resulting blockage of these sites leads to pancreatic insufficiency, "meconium ileus", and devastating "chronic obstructive pulmonary disease" (Al-Awqati, Q. et al. (1992) J. Exp. Biol. 172:245-266).

The voltage-gated chloride channels (CLC) are characterized by 10-12 transmembrane domains, as well as two small globular domains known as CBS domains. The CLC subunits probably function as homotetramers. CLC proteins are involved in regulation of cell volume, membrane potential stabilization, signal transduction, and transepithelial transport. Mutations in CLC-1, expressed predominantly in skeletal muscle, are responsible for autosomal recessive generalized myotonia and autosomal dominant myotonia congenita, while mutations in the kidney channel CLC-5 lead to kidney stones (Jentsch, T.J. (1996) Curr. Opin. Neurobiol. 6:303-310).

Ligand-gated channels open their pores when an extracellular or intracellular mediator binds to the channel. Neurotransmitter-gated channels are channels that open when a neurotransmitter binds to their extracellular domain. These channels exist in the postsynaptic membrane of nerve or muscle cells. There are two types of neurotransmitter-gated channels. Sodium channels open in response to excitatory neurotransmitters, such as acetylcholine, glutamate, and serotonin. This opening causes an influx of Na^+ and produces the initial localized depolarization that activates the voltage-gated channels and starts the action potential. Chloride channels open in response to inhibitory neurotransmitters, such as γ -aminobutyric acid (GABA) and glycine, leading to hyperpolarization of the membrane and the subsequent generation of an action potential. Neurotransmitter-gated ion channels have four transmembrane domains and probably function as pentamers (Jentsch, *supra*). Amino acids in the second transmembrane domain appear to be important

WO 02/40541

PCT/US01/46055

in determining channel permeation and selectivity (Sather, W.A. et al. (1994) *Curr. Opin. Neurobiol.* 4:313-323).

Ligand-gated channels can be regulated by intracellular second messengers. For example, calcium-activated K^+ channels are gated by internal calcium ions. In nerve cells, an influx of calcium during depolarization opens K^+ channels to modulate the magnitude of the action potential (Ishi et al., *supra*). The large conductance (BK) channel has been purified from brain and its subunit composition determined. The α subunit of the BK channel has seven rather than six transmembrane domains in contrast to voltage-gated K^+ channels. The extra transmembrane domain is located at the subunit N-terminus. A 28-amino-acid stretch in the C-terminal region of the subunit (the "calcium bowl" region) contains many negatively charged residues and is thought to be the region responsible for calcium binding. The β subunit consists of two transmembrane domains connected by a glycosylated extracellular loop, with intracellular N- and C-termini (Kaczorowski, *supra*; Vergara, C. et al. (1998) *Curr. Opin. Neurobiol.* 8:321-329).

Cyclic nucleotide-gated (CNG) channels are gated by cytosolic cyclic nucleotides. The best examples of these are the cAMP-gated Na^+ channels involved in olfaction and the cGMP-gated cation channels involved in vision. Both systems involve ligand-mediated activation of a G-protein coupled receptor which then alters the level of cyclic nucleotide within the cell. CNG channels also represent a major pathway for Ca^{2+} entry into neurons, and play roles in neuronal development and plasticity. CNG channels are tetramers containing at least two types of subunits, an α subunit which can form functional homomeric channels, and a β subunit, which modulates the channel properties. All CNG subunits have six transmembrane domains and a pore forming region between the fifth and sixth transmembrane domains, similar to voltage-gated K^+ channels. A large C-terminal domain contains a cyclic nucleotide binding domain, while the N-terminal domain confers variation among channel subtypes (Zufall, F. et al. (1997) *Curr. Opin. Neurobiol.* 7:404-412).

The activity of other types of ion channel proteins may also be modulated by a variety of intracellular signalling proteins. Many channels have sites for phosphorylation by one or more protein kinases including protein kinase A, protein kinase C, tyrosine kinase, and casein kinase II, all of which regulate ion channel activity in cells. Kir channels are activated by the binding of the $G\beta\gamma$ subunits of heterotrimeric G-proteins (Reimann, F. and P.M. Ashcroft (1999) *Curr. Opin. Cell. Biol.* 11:503-508). Other proteins are involved in the localization of ion channels to specific sites in the cell membrane. Such proteins include the PDZ domain proteins known as MAGUKs (membrane-associated guanylate kinases) which regulate the clustering of ion channels at neuronal synapses (Craven, S.E. and D.S. Brecht (1998) *Cell* 93:495-498).

Disease Correlation

The etiology of numerous human diseases and disorders can be attributed to defects in the

WO 02/40541

PCT/US01/46055

transport of molecules across membranes. Defects in the trafficking of membrane-bound transporters and ion channels are associated with several disorders, e.g., cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, von Gierke disease, and certain forms of diabetes mellitus. Single-gene defect diseases resulting in an inability to transport small molecules across membranes include, e.g., cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease (van't Hoff, W.G. (1996) *Exp. Nephrol.* 4:253-262; Talente, G.M. et al. (1994) *Ann. Intern. Med.* 120:218-226; and Chillon, M. et al. (1995) *New Engl. J. Med.* 332:1475-1480).

Human diseases caused by mutations in ion channel genes include disorders of skeletal muscle, cardiac muscle, and the central nervous system. Mutations in the pore-forming subunits of sodium and chloride channels cause myotonia, a muscle disorder in which relaxation after voluntary contraction is delayed. Sodium channel myotonias have been treated with channel blockers. Mutations in muscle sodium and calcium channels cause forms of periodic paralysis, while mutations in the sarcoplasmic calcium release channel, T-tubule calcium channel, and muscle sodium channel cause malignant hyperthermia. Cardiac arrhythmia disorders such as the long QT syndromes and idiopathic ventricular fibrillation are caused by mutations in potassium and sodium channels (Cooper, E.C. and L.Y. Jan (1998) *Proc. Natl. Acad. Sci. USA* 96:4759-4766). All four known human idiopathic epilepsy genes code for ion channel proteins (Berkovic, S.P. and L.E. Scheffer (1999) *Curr. Opin. Neurology* 12:177-182). Other neurological disorders such as ataxias, hemiplegic migraine and hereditary deafness can also result from mutations in ion channel genes (Jen, J. (1999) *Curr. Opin. Neurobiol.* 9:274-280; Cooper, *supra*).

Ion channels have been the target for many drug therapies. Neurotransmitter-gated channels have been targeted in therapies for treatment of insomnia, anxiety, depression, and schizophrenia. Voltage-gated channels have been targeted in therapies for arrhythmia, ischemic stroke, head trauma, and neurodegenerative disease (Taylor, C.P. and L.S. Narasimhan (1997) *Adv. Pharmacol.* 39:47-98). Various classes of ion channels also play an important role in the perception of pain, and thus are potential targets for new analgesics. These include the vanilloid-gated ion channels, which are activated by the vanilloid capsaicin, as well as by noxious heat. Local anesthetics such as lidocaine and mexiletine which blockade voltage-gated Na⁺ channels have been useful in the treatment of neuropathic pain (Eglen, *supra*).

Ion channels in the immune system have recently been suggested as targets for immunomodulation. T-cell activation depends upon calcium signaling, and a diverse set of T-cell specific ion channels has been characterized that affect this signaling process. Channel blocking agents can inhibit secretion of lymphokines, cell proliferation, and killing of target cells. A peptide antagonist of the T-cell potassium channel Kv1.3 was found to suppress delayed-type hypersensitivity and allogenic responses in pigs, validating the idea of channel blockers as safe and efficacious

WO 02/40541

PCT/US01/46055

immunosuppressants (Cahalan, M.D. and K.G. Chaudy (1997) Curr. Opin. Biotechnol. 8:749-756).

The discovery of new transporters and ion channels, 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 transport, neurological, muscle, immunological and cell proliferative disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of transporters and ion channels.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, transporters and ion channels, referred to collectively as "TRICH" and individually as "TRICH-1," "TRICH-2," "TRICH-3," "TRICH-4," "TRICH-5," "TRICH-6," "TRICH-7," "TRICH-8," "TRICH-9," "TRICH-10," "TRICH-11," "TRICH-12," "TRICH-13," "TRICH-14," "TRICH-15," "TRICH-16," "TRICH-17," "TRICH-18," "TRICH-19," and "TRICH-20." 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: 1-20, 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: 1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 1-20. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO: 1-20.

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: 1-20, 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: 1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 1-20. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO: 1-20. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO: 21-40.

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: 1-20, b) a polypeptide comprising a naturally occurring amino acid sequence at least

WO 02/40541

PCT/US01/46055

90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, 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:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

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

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:21-40, 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:21-40, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40, b) a polynucleotide comprising a naturally occurring polynucleotide

WO 02/40541

PCT/US01/46055

sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40, 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:21-40, 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:21-40, 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:1-20, 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:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional TRICH, 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:1-20, b) a polypeptide comprising a

WO 02/40541

PCT/US01/46055

naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. 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 TRICH, 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:1-20, 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:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. 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 TRICH, 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:1-20, 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:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino

WO 02/40541

PCT/US01/46055

acid sequence selected from the group consisting of SEQ ID NO:1-20, 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:1-20, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20. 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:21-40, 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:21-40, 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:21-40, 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:21-40, 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:21-40, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in

WO 02/40541

PCT/US01/46055

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

- 5 Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.
- Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability scores for the matches between each polypeptide and its homolog(s) are also shown.
- 10 Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.
- Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide
- 15 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
- 20 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

25 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,"

30 and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same

35 meanings as commonly understood by one of ordinary skill in the art to which this invention belongs.

WO 02/0541

PCT/US01/46055

Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

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

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

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

Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

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

WO 02/40541

PCT/US01/46055

Amino acids with uncharged side chains having similar hydrophobicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring 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 TRICH. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of TRICH either by directly interacting with TRICH or by acting on components of the biological pathway in which TRICH 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 TRICH 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 (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

WO 02/40541

PCT/US01/46055

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

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

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide

WO 02/40541

PCT/US01/46055

or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding TRICH or fragments of TRICH 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
20	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
25	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
30	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Ile, Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
35	Thr	Ser, Val
	Trp	Phe, Tyr
	Tyr	His, Phe, Trp
	Val	Ile, Leu, Thr

40 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

WO 02/40541

PCT/US01/46055

the side chain.

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

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

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

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

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

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

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

A fragment of SEQ ID NO:21-40 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:21-40, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:21-40 is useful, for
35 example, in hybridization and amplification technologies and in analogous methods that distinguish

WO 02/40541

PCT/US01/46055

SEQ ID NO:21-40 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:21-40 and the region of SEQ ID NO:21-40 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:1-20 is encoded by a fragment of SEQ ID NO:21-40. A fragment of SEQ ID NO:1-20 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-20. For example, a fragment of SEQ ID NO:1-20 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-20. The precise length of a fragment of SEQ ID NO:1-20 and the region of SEQ ID NO:1-20 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 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

WO 02/40541

PCT/US01/46055

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/blast/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: 30

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.

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

WO 02/40541

PCT/US01/46055

polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

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

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

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

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

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

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific
 30 hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive
 35 conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill

WO 02/40541

PCT/US01/46055

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

WO 02/40541

PCT/US01/46055

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of TRICH 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 TRICH 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 TRICH. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of TRICH.

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 ending in lysine. The terminal lysine confers solubility to the composition.

PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an TRICH 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 TRICH.

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

"Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target

WO 02/40541

PCT/US01/46055

polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

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

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

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

WO 02/40541

PCT/US01/46055

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.

5 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
10 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
15 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,
20 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.

25 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 TRICH,
30 nucleic acids encoding TRICH, 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
35 synthetic binding composition. The interaction is dependent upon the presence of a particular

WO 02/40541

PCT/US01/46055

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

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.

10 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.

25 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 02/40541

PCT/US01/46055

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.

30 THE INVENTION

The invention is based on the discovery of new human transporters and ion channels (TRICH), the polynucleotides encoding TRICH, and the use of these compositions for the diagnosis, treatment, or prevention of transport, neurological, muscle, immunological and cell proliferative disorders.

35 Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide

WO 02/40541

PCT/US01/46055

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

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (GenBank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability scores for the matches between each polypeptide and its homolog(s). Column 5 shows the annotation of the GenBank homologs 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 transporters and ion channels. For example, SEQ ID NO:5 is 61% identical to *Drosophila* sodium-hydrogen exchanger NHE1 (GenBank ID g4894991) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 6.0e-139, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:5 also contains a sodium/hydrogen exchanger family 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 analysis provides further corroborative evidence that SEQ ID NO:5 is a sodium/hydrogen exchanger. In an alternative example, SEQ ID NO:6 is about 50% identical to human citrin, the adult-onset type II citrullinemia protein, (GenBank ID g5052319) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is

WO 02/40541

PCT/US01/46055

6.0e-51, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:6 also contains mitochondrial carrier protein 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, MOTIFS, and

5 PROFILESSCAN analyses provide further corroborative evidence that SEQ ID NO:6 is a mitochondrial carrier protein. In an alternative example, SEQ ID NO:7 is 27% identical to *Synechocystis sp.* melibiose carrier protein (GenBank ID g1653342) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.8e-16, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance.

10 Additional BLAST data from DOMO and PRODOM analyses provide further corroborative evidence that SEQ ID NO:7 is a symporter protein. In an alternative example, SEQ ID NO:9 is 26% identical to an *Arabidopsis* ABC transporter (GenBank ID g4262239) and is 99% identical, from residue M1 to residue W374, to human sterolin-2 (GenBank ID g15146444) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability scores are 4.1e-25 and 0.0

15 respectively, which indicate the probabilities of obtaining the observed polypeptide sequence alignments by chance. SEQ ID NO:9 contains two transmembrane domains as determined by hidden Markov model (HMM) analysis, as well as a white/scarlet ABC transporter domain. (See Table 3.) These data provide further corroborative evidence that SEQ ID NO:9 is an ABC transporter. In an alternative example, SEQ ID NO:12 is 93% identical to rat neuronal glutamine transporter (GenBank

20 ID g6978016) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 4.4e-239, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:12 also contains a transmembrane amino acid transporter 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.)

25 These data provide corroborative evidence that SEQ ID NO:12 is an amino acid transporter protein. In an alternative example, SEQ ID NO:14 is 52% identical to mouse multidrug resistance protein (GenBank ID g387426) 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:14 also contains an ABC

30 transporter domain and an ABC transporter transmembrane region 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 PROFILESSCAN analyses provide further corroborative evidence that SEQ ID NO:14 is a multidrug resistance ABC transporter. In an alternative example, SEQ ID NO:18 is 41% identical to

35 *Arabidopsis* putative membrane transporter (GenBank ID g2289003) and is 99% identical, from

WO 02/40541

PCT/US01/46055

residue M20 to residue E648, to human proton myo-inositol transporter (GenBank ID g15211933) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability scores are 1.4e-94 and 0.0 respectively, which indicate the probabilities of obtaining the observed polypeptide sequence alignments by chance. SEQ ID NO:18 also contains a sugar (and other) transporter 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:18 is a sugar transporter. SEQ ID NO:1-4, SEQ ID NO:8, SEQ ID NO:10-11, SEQ ID NO:13, SEQ ID NO:15-17, and SEQ ID NO:19-20 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-20 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:21-40 or that distinguish between SEQ ID NO:21-40 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, 6122382H1 is the identification number of an Incyte cDNA sequence, and BRAHNON05 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., 72008374V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g2077361) 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

WO 02/40541

PCT/US01/46055

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_YYYY_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 YYYY is the number of the prediction generated by the algorithm, and N₁, N₂, ..., 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,

FLXXXXX_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).

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

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

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

Table 5 shows the representative cDNA libraries for those full length polynucleotide

WO 02/40541

PCT/US01/46055

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.

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

The invention also encompasses polynucleotides which encode TRICH. In a particular
10 embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:21-40, which encodes TRICH. The polynucleotide sequences of SEQ ID NO:21-40, 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.

15 The invention also encompasses a variant of a polynucleotide sequence encoding TRICH. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding TRICH. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID
20 NO:21-40 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:21-40. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of TRICH.

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant
25 of a polynucleotide sequence encoding TRICH. A splice variant may have portions which have significant sequence identity to the polynucleotide sequence encoding TRICH, 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%
30 polynucleotide sequence identity to the polynucleotide sequence encoding TRICH 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 TRICH. Any one of the splice variants described above can encode an amino acid sequence which contains at least one functional or
35 structural characteristic of TRICH.

WO 02/40541

PCT/US01/46055

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

Although nucleotide sequences which encode TRICH and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring TRICH under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding TRICH 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 TRICH and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode TRICH and TRICH 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 TRICH 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:21-40 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

WO 02/40541

PCT/US01/46055

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

5 (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 TRICH may be extended utilizing a partial nucleotide

10 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

15 sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme

20 digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060.) Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries

25 and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

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

35 Capillary electrophoresis systems which are commercially available may be used to analyze

WO 02/40541

PCT/US01/46055

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

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

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter TRICH-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of TRICH, 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

WO 02/40541

PCT/US01/46055

manner.

In another embodiment, sequences encoding TRICH 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.)

- 5 Alternatively, TRICH 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 Koberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the
- 10 amino acid sequence of TRICH, 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. Reguier (1990) *Methods Enzymol.* 182:392-421.)

- 15 The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, *supra*, pp. 28-33.)

In order to express a biologically active TRICH, the nucleotide sequences encoding TRICH or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding

20 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 TRICH. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding TRICH. Such signals include the ATG initiation codon and adjacent sequences, e.g. the

- 25 Kozak sequence. In cases where sequences encoding TRICH 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
- 30 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 TRICH and appropriate transcriptional and translational

- 35 control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques,

WO 02/40541

PCT/US01/46055

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.)

5 A variety of expression vector/host systems may be utilized to contain and express sequences encoding TRICH. 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);
10 plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, *supra*; Ausubel, *supra*; Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509; Engolhard, 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
15 York NY, pp. 191-196; Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; and Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al. (1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344; Buller, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al. (1994) *Mol. Immunol.* 31(3):219-226; and Verma, I.M. and N. Sornis (1997) *Nature* 389:239-242.)
20 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 TRICH. For example, routine cloning,
25 subcloning, and propagation of polynucleotide sequences encoding TRICH can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding TRICH 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
30 *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 TRICH are needed, e.g. for the production of antibodies, vectors which direct high level expression of TRICH may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

35 Yeast expression systems may be used for production of TRICH. A number of vectors

WO 02/06541

PCT/US01/46055

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, 5 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 TRICH. Transcription of sequences encoding TRICH 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.* 10 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* 15 (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 TRICH 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 20 infective virus which expresses TRICH in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of 25 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 30 of TRICH in cell lines is preferred. For example, sequences encoding TRICH 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 35 confer resistance to a selective agent, and its presence allows growth and recovery of cells which

WO 02/40541

PCT/US01/46055

successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apv* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to 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 TRICH is inserted within a marker gene sequence, transformed cells containing sequences encoding TRICH can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding TRICH 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 TRICH and that express TRICH 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 TRICH using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on TRICH is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See,

WO 02/40541

PCT/US01/46055

e.g., Hampton, R. et al. (1990) *Serological Methods, a Laboratory Manual*, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) *Current Protocols in Immunology*, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) *Immunocytochemical Protocols*, Humana Press, Totowa NJ.)

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

Host cells transformed with nucleotide sequences encoding TRICH 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
20 and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode TRICH may be designed to contain signal sequences which direct secretion of TRICH through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of
25 the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the
30 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 TRICH 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 TRICH protein
35 containing a heterologous moiety that can be recognized by a commercially available antibody may

WO 02/40541

PCT/US01/46055

facilitate the screening of peptide libraries for inhibitors of TRICH activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the TRICH encoding sequence and the heterologous protein sequence, so that TRICH may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

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

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

In one embodiment, the compound thus identified is closely related to the natural ligand of TRICH, 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 TRICH 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 TRICH, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing TRICH or cell membrane fractions which contain TRICH are then contacted with a test compound and binding, stimulation, or inhibition of activity of either TRICH or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is

WO 02/40541

PCT/US01/46055

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 TRICH, either in solution or affixed to a solid support, and detecting the binding of TRICH 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.

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

In another embodiment, polynucleotides encoding TRICH 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, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (*neo*; Capecchi, M.R. (1989) *Science* 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. et al. (1997) *Nucleic Acids Res.* 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding TRICH may also be manipulated *in vitro* in ES cells derived from

WO 02/40541

PCT/US01/46055

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).

5 Polynucleotides encoding TRICH 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 TRICH 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
10 treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress TRICH, e.g., by secreting TRICH 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
15 between regions of TRICH and transporters and ion channels. In addition, the expression of TRICH is closely associated with tumorous tissues such as spleen tumor tissue, esophageal tumor tissue, brain tumor tissue, and myxoma from atrium tissue; and normal tissues such as kidney, liver, nasal polyp, prostate, thyroid, umbilical cord blood, neuronal, digestive, uterine endometrial tissue, and normal brain tissue such as the tissues from striatum, globus pallidus, and posterior putamen.
20 Therefore, TRICH appears to play a role in transport, neurological, muscle, immunological and cell proliferative disorders. In the treatment of disorders associated with increased TRICH expression or activity, it is desirable to decrease the expression or activity of TRICH. In the treatment of disorders associated with decreased TRICH expression or activity, it is desirable to increase the expression or activity of TRICH.

25 Therefore, in one embodiment, TRICH 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 TRICH. Examples of such disorders include, but are not limited to, a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus,
30 diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, cataplexy, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer, cardiac disorders associated with transport, e.g., angina, bradyarrhythmia, tachyarrhythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline
35 myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy,

WO 02/40541

PCT/US01/46055

ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis, neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia, and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, glycogen storage disease, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, pseudohypaldosteronism type 1, Liddle's syndrome, cystinuria, iminoglycinuria, Hartup disease, Fanconi disease, and Bartter syndrome; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental 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, hemiplegic migraine, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a muscle disorder such as cardiomyopathy, myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, ethanol myopathy, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, pheochromocytoma, and myopathies including encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, ophthalmoplegia, acid maltase

WO 02/40541

PCT/US01/46055

deficiency (AMD, also known as Pompe's disease), generalized myotonia, and myotonia congenita; an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune

5 polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxicity, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis,

10 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; and a

15 cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall

20 bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

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

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

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

In a further embodiment, an antagonist of TRICH may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TRICH. Examples of such disorders include, but are not limited to, those transport, neurological, muscle, immunological and

35 cell proliferative disorders described above. In one aspect, an antibody which specifically binds

WO 02/40541

PCT/US01/46055

TRICH 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 TRICH.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding TRICH may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TRICH 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 TRICH may be produced using methods which are generally known in the art. In particular, purified TRICH may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind TRICH. Antibodies to TRICH 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 TRICH or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to TRICH 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 TRICH 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 TRICH 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

WO 02/40541

PCT/US01/46055

technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Coto, 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
 5 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
 10 chain antibodies may be adapted, using methods known in the art, to produce TRICH-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
 15 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 TRICH 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
 20 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
 25 polyclonal or monoclonal antibodies with established specificities are well known in the art. Such
 immunoassays typically involve the measurement of complex formation between TRICH and its
 specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies
 reactive to two non-interfering TRICH epitopes is generally used, but a competitive binding assay
 may also be employed (Pound, *supra*).

30 Various methods such as Scatchard analysis in conjunction with radioimmunoassay
 techniques may be used to assess the affinity of antibodies for TRICH. Affinity is expressed as an
 association constant, K_a , which is defined as the molar concentration of TRICH-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
 35 affinities for multiple TRICH epitopes, represents the average affinity, or avidity, of the antibodies

WO 02/40541

PCT/US01/46055

- for TRICH. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular TRICH 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 TRICH-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of TRICH, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.B. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).
- The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of TRICH-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, *supra*, and Coligan et al. *supra*.)
- In another embodiment of the invention, the polynucleotides encoding TRICH, 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 TRICH. 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 TRICH. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)
- In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, *supra*; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.*

WO 02/40541

PCT/US01/46055

25(14);2730-2736.)

In another embodiment of the invention, polynucleotides encoding TRICH may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in TRICH expression or regulation causes disease, the expression of TRICH 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 TRICH are treated by constructing mammalian expression vectors encoding TRICH and introducing these vectors by mechanical means into TRICH-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J.-L. and H. R  capon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of TRICH include, but are not limited to, the pCDNA 3.1, pPITAG, 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). TRICH 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.*

WO 02/06541

PCT/US01/46055

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 TRICH from a normal individual.

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

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to TRICH expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding TRICH under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver

WO 02/40541

PCT/US01/46055

polynucleotides encoding TRICH to cells which have one or more genetic abnormalities with respect to the expression of TRICH. 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
 5 (Csote, 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. Sonja (1997) *Nature* 389:239-242, both incorporated by reference herein.

10 In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding TRICH to target cells which have one or more genetic abnormalities with respect to the expression of TRICH. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing TRICH 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
 15 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
 20 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
 25 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
 30 deliver polynucleotides encoding TRICH to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA,
 35 resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity

WO 02/40541

PCT/US01/46055

(e.g., protease and polymerase). Similarly, inserting the coding sequence for TRICH into the alphavirus genome in place of the capsid-coding region results in the production of a large number of TRICH-coding RNAs and the synthesis of high levels of TRICH 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 (SN) 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 TRICH 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.J. 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 TRICH.

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

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques

WO 02/40541

PCT/US01/46055

for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding TRICH. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

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

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

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

WO 02/40541

PCT/US01/46055

biochemical system. Alterations in the expression of a polynucleotide encoding TRICH 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 TRICH. The amount of hybridization may be quantified, thus
 5 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
 10 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.J. 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
 15 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.
 20 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
 25 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
 30 *Remington's Pharmaceutical Sciences* (Maack Publishing, Eason PA). Such compositions may consist of TRICH, antibodies to TRICH, and mimetics, agonists, antagonists, or inhibitors of TRICH.

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,
 35 enteral, topical, sublingual, or rectal means.

WO 02/40541

PCT/US01/46055

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 TRICH or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, TRICH 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 TRICH or fragments thereof, antibodies of TRICH, and agonists, antagonists or inhibitors of TRICH, 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 02/40541

PCT/US01/46055

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. These 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 TRICH may be used for the diagnosis of disorders characterized by expression of TRICH, or in assays to monitor patients being treated with TRICH or agonists, antagonists, or inhibitors of TRICH. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for TRICH include methods which utilize the antibody and a label to detect TRICH 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 TRICH, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of TRICH expression. Normal or standard values for TRICH expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to TRICH under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of TRICH 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 TRICH 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 TRICH may be correlated

WO 02/40541

PCT/US01/46055

with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of TRICH, and to monitor regulation of TRICH levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding TRICH or closely related molecules may be used to identify nucleic acid sequences which encode TRICH. 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 TRICH, 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 TRICH 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:21-40 or from genomic sequences including promoters, enhancers, and introns of the TRICH gene.

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

Polynucleotide sequences encoding TRICH may be used for the diagnosis of disorders associated with expression of TRICH. Examples of such disorders include, but are not limited to, a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, cataplexy, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer, cardiac disorders associated with transport, e.g., angina, bradyarrhythmia, tachyarrhythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis, neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia, and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease,

WO 02/40541

PCT/US01/46055

cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, glycogen storage disease, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease,

5 pseudohypoadosteronism type 1, Liddle's syndrome, cystinuria, urologlycinuria, Hartup disease, Fanconi disease, and Bartter syndrome; 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

10 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,

15 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

20 toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, hemiplegic migraine, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a muscle disorder such as cardiomyopathy,

25 myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, ethanol myopathy, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine,

30 pheochromocytoma, and myopathies including encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, ophthalmoplegia, acid maltase deficiency (AMD, also known as Pompe's disease), generalized myotonia, and myotonia congenita; an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis,

35 autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-

WO 02/40541

PCT/US01/46055

candidiasis-ecrodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxicity, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hyper eosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, peoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding TRICH 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 TRICH expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding TRICH may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding TRICH 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 TRICH 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 TRICH, 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 TRICH, under conditions suitable for hybridization or

WO 02/40541

PCT/US01/46055

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

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

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

Additional diagnostic uses for oligonucleotides designed from the sequences encoding TRICH 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 TRICH, or a fragment of a polynucleotide complementary to the polynucleotide encoding TRICH, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

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

WO 02/40541

PCT/US01/46055

methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

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

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

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

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See 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

WO 02/40541

PCT/US01/46055

transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

5 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.P. 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/cc/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.

35 Another particular embodiment relates to the use of the polypeptide sequences of the present

WO 02/40541

PCT/US01/46055

invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at 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 TRICH to quantify the levels of TRICH 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; Mendez, 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

WO 02/06541

PCT/US01/46055

cases.

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

WO 02/40541

PCT/US01/46055

with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lauder, 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 TRICH on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

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

In another embodiment of the invention, TRICH, 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 TRICH 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 TRICH, or fragments thereof, and washed. Bound TRICH is then detected by methods well known in the art. Purified TRICH 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 TRICH specifically compete with a test compound for binding TRICH.

WO 02/40541

PCT/US01/46055

In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with TRICH.

In additional embodiments, the nucleotide sequences which encode TRICH may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, including U.S. Ser. No. 60/243,989, U.S. Ser. No. 60/245,904, U.S. Ser. No. 60/249,661, U.S. Ser. No. 60/247,673, U.S. Ser. No. 60/252,232, and U.S. Ser. No. 60/250,790, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

Incye cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incye 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 SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic

WO 02/40541

PCT/US01/46055

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), PSPORI1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pGEN (Incyte Genomics, Palo Alto CA), 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 DH5a, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaitersburg 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

WO 02/40541

PCT/US01/46055

ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

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

WO 02/40541

PCT/US01/46055

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:21-40. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative transporters and ion channels 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 transporters and ion channels, the encoded polypeptides were analyzed by querying against PFAM models for transporters and ion channels. Potential transporters and ion channels were also identified by homology to Incyte cDNA sequences that had been annotated as transporters and ion channels. 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

35 "Stitched" Sequences

WO 02/40541

PCT/US01/46055

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. 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.

20 "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 TRICU Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:21-40 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other

WO 02/40541

PCT/US01/46055

implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:21-40 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 Génethon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Génethon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

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

WO 02/40541

PCT/US01/46055

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

VIII. Extension of TRICH Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

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

WO 02/40541

PCT/US01/46055

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

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (w/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

WO 02/06541

PCT/US01/46055

primers and the DYNAMIC 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:21-40 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ - ^{32} P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech).

An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

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 (Schna (1999), *supra*). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers.

Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schna, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)

WO 02/40541

PCT/US01/46055

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

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μ l 5X SSC/0.2% SDS.

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

WO 02/40541

PCT/US01/46055

- 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
- 5 110°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.
- 10 Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.
- 15 **Hybridization**
- Hybridization reactions contain 9 μ l of sample mixture consisting of 0.2 μ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just
- 20 slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.
- 25 **Detection**
- Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide
- 30 containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.
- In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate
- 35

WO 02/40541

PCT/US01/46055

filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

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

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC
15 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.

20 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

25 Sequences complementary to the TRICH-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring TRICH. 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 TRICH. To
30 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 TRICH-encoding transcript.

XII. Expression of TRICH

35 Expression and purification of TRICH is achieved using bacterial or virus-based expression

WO 02/40541

PCT/US01/46055

systems. For expression of TRICH 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 TRICH upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of TRICH 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 TRICH by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945.)

In most expression systems, TRICH 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 TRICH 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 TRICH obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, and XVIII, where applicable.

XIII. Functional Assays

TRICH function is assessed by expressing the sequences encoding TRICH 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

WO 02/40541

PCT/US01/46055

formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) *Flow Cytometry*, Oxford, New York NY.

The influence of TRICH on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding TRICH and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding TRICH and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIV. Production of TRICH Specific Antibodies

TRICH 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 TRICH amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI431A peptide synthesizer (Applied Biosystems) using Fmoc chemistry and coupled to KJ.H (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to

WO 02/40541

PCT/US01/46055

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

XV. Purification of Naturally Occurring TRICH Using Specific Antibodies

Naturally occurring or recombinant TRICH is substantially purified by immunoaffinity chromatography using antibodies specific for TRICH. An immunoaffinity column is constructed by covalently coupling anti-TRICH antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing TRICH are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of TRICH (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/TRICH binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and TRICH is collected.

XVI. Identification of Molecules Which Interact with TRICH

Molecules which interact with TRICH may include transporter substrates, agonists or antagonists, modulatory proteins such as G $\beta\gamma$ proteins (Reimann, *supra*) or proteins involved in TRICH localization or clustering such as MAGUKs (Craven, *supra*). TRICH, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled TRICH, washed, and any wells with labeled TRICH complex are assayed. Data obtained using different concentrations of TRICH are used to calculate values for the number, affinity, and association of TRICH with the candidate molecules.

Alternatively, proteins that interact with TRICH are isolated using the yeast 2-hybrid system (Fields, S. and O. Song (1989) *Nature* 340:245-246). TRICH, or fragments thereof, are expressed as fusion proteins with the DNA binding domain of Gal4 or LexA, and potential interacting proteins are expressed as fusion proteins with an activation domain. Interactions between the TRICH fusion protein and the TRICH interacting proteins (fusion proteins with an activation domain) reconstitute a transactivation function that is observed by expression of a reporter gene. Yeast 2-hybrid systems are commercially available, and methods for use of the yeast 2-hybrid system with ion channel proteins are discussed in Niethammer, M. and M. Söberg (1998, *Meth. Enzymol.* 293:104-122).

TRICH 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

WO 02/00541

PCT/US01/46055

between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Potential TRICH agonists or antagonists may be tested for activation or inhibition of TRICH ion channel activity using the assays described in section XVIII.

5 XVII. Demonstration of TRICH Activity

Ion channel activity of TRICH is demonstrated using an electrophysiological assay for ion conductance. TRICH can be expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector encoding TRICH. 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 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 TRICH and β -galactosidase.

15 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 are tested for differences in membrane conductance by electrophysiological techniques that are well known in the art. Untransformed cells, and/or cells transformed with either vector sequences alone or β -galactosidase sequences alone, are used as controls and tested in parallel. Cells expressing TRICH will have higher anion or cation conductance relative to control cells. The contribution of TRICH to conductance can be confirmed by incubating the cells using antibodies specific for TRICH. The antibodies will bind to the extracellular side of TRICH, thereby blocking the pore in the ion channel, and the associated conductance.

Alternatively, ion channel activity of TRICH is measured as current flow across a TRICH-containing *Xenopus laevis* oocyte membrane using the two-electrode voltage-clamp technique (Ishii et al., *supra*; Jegla, T. and L. Salkoff (1997) J. Neurosci. 17:32-44). TRICH is subcloned into an appropriate *Xenopus* oocyte expression vector, such as pBF, and 0.5-5 ng of mRNA is injected into mature stage IV oocytes. Injected oocytes are incubated at 18 °C for 1-5 days. Inside-out macropatches are excised into an intracellular solution containing 116 mM K-gluconate, 4 mM KCl, and 10 mM Hapes (pH 7.2). The intracellular solution is supplemented with varying concentrations of the TRICH mediator, such as cAMP, cGMP, or Ca^{2+} (in the form of CaCl_2), where appropriate. Electrode resistance is set at 2-5 M Ω and electrodes are filled with the intracellular solution lacking mediator. Experiments are performed at room temperature from a holding potential of 0 mV. Voltage ramps (2.5 s) from -100 to 100 mV are acquired at a sampling frequency of 500 Hz. Current measured is proportional to the activity of TRICH in the assay.

WO 02/0541

PCT/US01/46055

In particular, the activity of TRICH-2 is measured as voltage-gated Ca^{2+} or Na^+ conductance, the activity of TRICH-15 is measured as Ca^{2+} conductance, and the activity of TRICH-16 is measured as K^+ conductance.

Transport activity of TRICH is assayed by measuring uptake of labeled substrates into

- 5 *Xenopus laevis* oocytes. Oocytes at stages V and VI are injected with TRICH mRNA (10 ng per oocyte) and incubated for 3 days at 18°C in OR2 medium (82.5mM NaCl, 2.5 mM KCl, 1mM CaCl_2 , 1mM MgCl_2 , 1mM Na_2HPO_4 , 5 mM Hepes, 3.8 mM NaOH, 50µg/ml gentamycin, pH 7.8) to allow expression of TRICH. Oocytes are then transferred to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM CaCl_2 , 1mM MgCl_2 , 10 mM Hepes/Tris pH 7.5). Uptake of various substrates (e.g.,
- 10 amino acids, sugars, drugs, ions, and neurotransmitters) is initiated by adding labeled substrate (e.g. radiolabeled with ^3H , fluorescently labeled with rhodamine, etc.) to the oocytes. After incubating for 30 minutes, uptake is terminated by washing the oocytes three times in Na^+ -free medium, measuring the incorporated label, and comparing with controls. TRICH activity is proportional to the level of internalized labeled substrate. In particular, test substrates include tricarboxylates for TRICH-1, H^+
- 15 for TRICH-3, sulfate for TRICH-4, Na^+ for TRICH-5, anionic metabolites for TRICH-6, glucose-6-phosphate for TRICH-8, and amino acids for TRICH-10.

- ATPase activity associated with TRICH can be measured by hydrolysis of radiolabeled ATP- $[\gamma\text{-}^{32}\text{P}]$, separation of the hydrolysis products by chromatographic methods, and quantitation of the recovered ^{32}P using a scintillation counter. The reaction mixture contains ATP- $[\gamma\text{-}^{32}\text{P}]$ and varying
- 20 amounts of TRICH in a suitable buffer incubated at 37°C for a suitable period of time. The reaction is terminated by acid precipitation with trichloroacetic acid and then neutralized with base, and an aliquot of the reaction mixture is subjected to membrane or filter paper-based chromatography to separate the reaction products. The amount of ^{32}P liberated is counted in a scintillation counter. The amount of radioactivity recovered is proportional to the ATPase activity of TRICH in the assay.

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

XVIII. Identification of TRICH Agonists and Antagonists

- 35 TRICH is expressed in a eukaryotic cell line such as CHO (Chinese Hamster Ovary) or HEK

WO 02/40541

PCT/US01/46055

(Human Embryonic Kidney) 293. Ion channel activity of the transformed cells is measured in the presence and absence of candidate agonists or antagonists. Ion channel activity is assayed using patch clamp methods well known in the art or as described in Example XVII. Alternatively, ion channel activity is assayed using fluorescent techniques that measure ion flux across the cell membrane (Velicelehi, G. et al. (1999) *Meth. Enzymol.* 294:20-47; West, M.R. and C.R. Molloy (1996) *Anal. Biochem.* 241:51-58). These assays may be adapted for high-throughput screening using microplates. Changes in internal ion concentration are measured using fluorescent dyes such as the Ca^{2+} indicator Fluo-4 AM, sodium-sensitive dyes such as SBFI and sodium green, or the Cl^- indicator MQAE (all available from Molecular Probes) in combination with the FLIPR fluorimetric plate reading system (Molecular Devices). In a more generic version of this assay, changes in membrane potential caused by ionic flux across the plasma membrane are measured using oxonol dyes such as DiBAC₄ (Molecular Probes). DiBAC₄ equilibrates between the extracellular solution and cellular sites according to the cellular membrane potential. The dye's fluorescence intensity is 20-fold greater when bound to hydrophobic intracellular sites, allowing detection of DiBAC₄ entry into the cell (Gonzalez, J.E. and P.A. Negulescu (1998) *Curr. Opin. Biotechnol.* 9:624-631). Candidate agonists or antagonists may be selected from known ion channel agonists or antagonists, peptide libraries, or combinatorial chemical libraries.

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 02/40541

PCT/US01/46055

Table 1

Inocyte Project ID	Polypeptide SEQ ID NO:	Inocyte Polypeptide ID	Polynucleotide SEQ ID NO:	Inocyte Polynucleotide ID
1626101	1	1626101CD1	21	1626101CB1
2907828	2	2907828CD1	22	2907828CB1
3968527	3	3968527CD1	23	3968527CB1
7472732	4	7472732CD1	24	7472732CB1
7476938	5	7476938CD1	25	7476938CB1
8128531	6	8128531CD1	26	8128531CB1
7476957	7	7476957CD1	27	7476957CB1
766243	8	766243CD1	28	766243CB1
5585710	9	5585710CD1	29	5585710CB1
7487592	10	7487592CD1	30	7487592CB1
2507246	11	2507246CD1	31	2507246CB1
3033505	12	3033505CD1	32	3033505CB1
4027893	13	4027893CD1	33	4027893CB1
7472030	14	7472030CD1	34	7472030CB1
7476089	15	7476089CD1	35	7476089CB1
6438177	16	6438177CD1	36	6438177CB1
7477243	17	7477243CD1	37	7477243CB1
7473042	18	7473042CD1	38	7473042CB1
7482460	19	7482460CD1	39	7482460CB1
1578772	20	1578772CD1	40	1578772CB1

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank NO.	Probability score	GenBank Homology
1	1526101CD1	G13785618	1.008-105	[flj] [Mus musculus] side-of-exon 4 Fleming, M. D. et al. (2001) A mutation in a mitochondrial transmembrane protein is responsible for the pleiotropic hematological and skeletal phenotype of floxex-tail [f/f] mice. <i>Genes Dev.</i> 15:652-657
2	2307838CD1	G12004581	0.0	[Rattus sp.] t5-carboxylate carrier Azzi, A. et al. (1993) The mitochondrial tri-carboxylate carrier. <i>J. Bioenerg. Biomembr.</i> 25:5-5-524
3	3965927CD1	G6434968	0.0	[flj] [Mus musculus] calcium channel [Mus musculus] putative K1-2g_Apase Walleck, M.S. et al. (1999) Differential expression of putative translocator amphipathic transporters, Zyg101, Genomica (Online) 11:139-150
4	7475732CD1	G15341532	0.0	[flj] [Homo sapiens] (AF331531) putative anion transporter [Mus musculus] sulfate transporter Kobayashi, T. et al. (1997) Cloning of mouse diastrophic dysplasia sulfate transporter gene induced during osteoblast differentiation by bone morphogenetic protein-2. <i>Gene</i> 198:341-349
5	7475938CD1	G4894891	6.008-139	[Xenopus laevis] sodium-hydrogen exchanger NH81 [Homo sapiens] citrinin, adult-onset type II citrullinemia protein Kobayashi, K. et al. (1999) The gene mutated in adult- onset type II citrullinemia encodes a putative mitochondrial carrier protein. <i>Nat. Genet.</i> 22:159-163
7	7476751CD1	G1653342	1.808-16	[Xenopus laevis] galactose carrier protein Kazuo, T. et al. (1995) DNA Res. 3:109-136
8	286243CD1	G7229675	6.908-39	[Arabidopsis thaliana] glucose 6 phosphate/phosphate translocator

WO 02/40541

PCT/US01/46055

Table 2 (cont.)

Polyptide Seq ID NO:	Incyte Polyptide ID	GenBank ID MO:	Probability score	GenBank Homolog
9	6385710CD1	G15146444	0.0	[fll][Homo sapiens] sterolin-2 Lu, K. et al. (2001) Two genes that map to the stx1 locus cause sitosterolemia: genomic structure and spectrum of mutations involving sterolin-1 and sterolin-2, encoded by ABCG8 and ABCG9, respectively. Am. J. Hum. Genet. 69:278-290
11	2507246CD1	G472900	3.10E-56	[Caenorhabditis elegans] carrier protein (c2) Furukawa, M. et al. (1994) Extension of the mitochondrial transport superfamily: sequences of five members from the nematode worm <i>Caenorhabditis elegans</i> . DNA Seq. 4:281-291
12	3033505CD1	G6978016	4.40E-239	[Rattus norvegicus] neuronal glutamine transporter Vaccani, H. et al. (2000) Cloning and functional identification of a neuronal glutamine transporter. J. Biol. Chem. 275:4049-4054
13	4027603CD1	G2198807	1.10E-53	[Gallus gallus] monocarboxylate transporter 3 Yoon, H. et al. (1997) Biochem. Biophys. Res. Commun. 234:90-94; Yoon, H. and Philo, N. (1994) J. Exp. Eye Res. 57:417-424; Yoon, H. et al. (1999) Genomics 60:366-370
14	7472030CD1	G387435	0.0	[Mus musculus] multidrug resistance protein Gros, P. et al. (1986) Cell 47:371-380
15	7476389CD1	G2836759	2.50E-11	[Caenorhabditis elegans] sodium/calcium exchanger domain: PF02214 (K+ channel tetramerization domain)
16	6428177CD1	G3889445	1.70E-14	[Caenorhabditis elegans] contains similarity to Pfam domain: PF02214 (K+ channel tetramerization domain)
17	7477243CD1	G6457274	0.0	[Mus musculus] putative K+-ATPase Thalass, K.S. et al. (1999) Physiol. Genomics (Online) 1:139-150
18	7471042CD1	G15211933	0.0	[fll][Homo sapiens] proton myo-inositol transporter Udelsky, R. et al. (2001) Identification of a neuronal N(+)-myo-inositol symporter expressed predominantly in the brain. The EMBO Journal 20:4467-4477

WO 02/40541

PCT/US01/46055

Table 2 (cont.)

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
19	7482060CD1	g5606493	6.80E-83	(Homo sapiens) cardiac potassium channel subunit (Kv6.2)
20	1578772CD1	g11933435 g11907976	2.00E-05 3.00E-05	(fl)(Arabidopsis thaliana) sulfate transporter (fl)(Solanum tuberosum) high affinity sulfate transporter type 1

WO 02/40541

PCT/US01/46055

Table 3

Seq ID	Incyte ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	1626101CD1	337	S22, S117, S124, T58, T34, T71	N154	POTATY TRANSMEMBRANE CHROMOSOME XV READING FRAME PD05886:A20-P264	BLAST_PRODOM
2	2507828CD1	816	S123, S264, S351, S359, S375, S395, S4, S54, S697, S703, S716, S745, S749, T14, T322, T382, T559, T602, T618, T639, T764, T824	N599, N511, N616, N695	Transmembrane domains: L107-L134, L235-F254, C257-F120, G306-L523, M550-F577, L666-L686 Ion transport protein domain: L437-L686 Sodium channel signature PR00170: Q227-F254, S296-L325 Transmembrane domains: V300-V319, T353-K380, V1004-S1023 EI-E2 ATPase domain: G146-T174, N256-E279 EI-E2 ATPases phosphorylation site Signature R100154: G158-F175, T185-F403, D653-L693 EI-E2 ATPase phosphorylation site: T571-D419 P-type cation-transporting ATPase superfamily signature PR00115: L289-F403, S659-D673 ATPase HYDROLASE TRANSMEMBRANE PHOSPHORYLATION BINDING PROTEIN PROBABLE CALCIUM-TRANSPORTING CALCIUM TRANSPORTER PD01657:R817-L909 PD14950: C757-F818 PD06817:R145-D245	HMMER HMMER_PPAM BLIIPS_BLOCKS PROFILESAN BLIIPS_PRINTS HMMER HMMER_PPAM BLIIPS_BLOCKS PROFILESAN BLIIPS_PRINTS BLAST_PRODOM
3	3968527CD1	1647	S1038, S119, S146, S166, S417, S453, S491, S498, S499, S548, S559, S605, S624, S629, S835, S920, T143, T147, T207, T212, T240, T276, T377, T390, T397, T445, T528, T634, T649, T665, T687, T707, T743, T776, T881, T811	N182, N285, N335		

WO 02/06541

PCT/US01/46055

Table 3 (cont.)

SEQ ID NO	Incyte ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					do ATPase; CALCIUM TRANSPORTING; DMO2405 Q10309 127-845:Y495-S981, S132-A464, S81-V370 DMO2405 P40527 208-977:R201-S981, S91-A492 DMO2405 Q09891 206-1107:E715-L851, L444-F703, M154-K327, L336-G414 DMO2405 P39524 236-1049:V148-V869, Q92-D365, I936-S1038	BLAST_DMO
					EL-E2 ATPase motif: D391-T397	NOTES
4	7472732CD1	671	S10 S138 S225 S311 S345 S352 S454 S556 S640 S658 Y507 Y595	N125 N131 N661	Transmembrane domains: L228-N248, L399-Y417, V470-P488 Sulfate transporter Family Domain: M162-F487 SPAS (Sulfate Transporter and Anti sigma factor antagonist) domain: E509-A652 Sulfate transporter proteins signature E401130: A159-M201, S53-L106 PROTEIN TRANSPORT SULFATE TRANSPORTER TRANSMEMBRANE PERMEASE INTERGENIC REGION AFFINITY GLYCOPROTEIN P0001255:V164-R485 P000121:P33-G168 SULFATE TRANSPORTERS DMO1229 P40879 5-463:P33-M446 DMO1229 P50443 49-505:L32-N447 DMO1229 P45180 10-468:P6-M446 DMO1229 P33193 11-427:P33-M201, E303-M446	HMWR HMWR_FFAM HMWR_FFAM BLAST_BLOCKS BLAST_FFAM BLAST_DMO

WO 02/04541

PCT/US01/46055

Table 3 (cont.)

SEQ ID NO:	Incyte ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical methods and databases
5	7476938CD1	671	S15 S288 S582 S661 S663 S68 S85 T104 T120 T334 T440 T503 T584 T600 T630 T634 T649	M103 N109 N176 N337 N47 N480	Transmembrane domains: M150-G171, Y172-I194, T241-I263, G390-G418, V452-I480, I537-G555 Sodium/hydrogen exchanger family domain: I152-D568 Na ⁺ /H ⁺ exchanger signature PRO1084: M215-P226, G229-S243, I244-T252, G284-L294 + TRANSPORT EXCHANGER NA P001672: M215-I263, A297-I333 Na ⁺ /H ⁺ PROTEIN TRANSMEMBRANE TRANSPORT ANTIPORTER SYMPORT SODIUM EXCHANGER GLYCOPROTEIN SODIUM/HYDROGEN P000363: G149-E557	HMER HMER PFAM BLIMES_PRINTS BLIMES_PRODOM BLAST_PRODOM
6	8128531CD1	315	S165 T149 T160 T2 T240 T251 T35 Y261	N71	40 BEHA: EXCHANGER NA: DM02572 Q01345 12-703:157-N580 DM02572 P48751 17-738:315-N580 DM02572 P26434 14-716:1356-R332 DM02572 P48764 10-734:1356-R336 Mitochondrial carrier proteins domain: S7-Q39, N101-S219 Mitochondrial energy transfer proteins signature P000415: T13-Q37, M173-G185 Mitochondrial energy transfer proteins signature: H3-T35, L102-S150, Y221-L271 PROTEIN TRANSPORT TRANSMEMBRANE REPEAT MITOCHONDRIAL CARRIER NUCLEONAR INNER P000317: S7-Y302	HMER PFAM BLIMES_BLOCKS PROFILES CAN BLAST_PRODOM

WO 02/04541

PCT/US01/46055

Table 3 (cont.)

SPO ID NO.	Tricyste Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					MITOCHONDRIAL ENERGY TRANSFER PROTEINS DMC0026 S5495 534-520:G15-Q97 DMC0026 S5495 522-719:E105-Q131, A170-L207 DMC0026 S60949 16-113:K13-L92 Mitochondrial carrier protein motif: P28-L36 P241-L249	BLAST_DOMO
7	7476757CDL	445	S247 S26 S389 S393 S440 T181 T355 T380 T405		SODIUM GALACTOSIDE SYMPORTER FAMILY DX01084 P10868 1-456:L178-K382 (p=6.1e-07) TRANSFERRIN PROTEIN TRANSFERASE SYMPORT SUGAR SYMPORTER PERMEASE INNER MEMBRANE CARRIER EDC03362:R123-L375 (p=5.7e-09) Signal peptide signal_peptide:R1-G23	BLAST_DOMO BLAST_PRODOM HMMER HMMER_PFIAM
8	266243CDL	410	S162 S208 S356 S363 T140 T180 T297 T335 T351	N295 N333	Integral membrane protein DUF6 DUF6_A39-V181 INTERMEMBRANE SPACE DOMAIN DX02684 S53128 1-401:V12-K136, P15-L63 DX02684 S57450 1-407:R10-K145, L26-V88 DX02684 S57457 1-409:R21-K145, H10-Q90 DX02684 P52172 1-406:R10-K145, V5-S89 TRANSLOCATOR PRECURSOR TRANSFERASE CYT PEROXIDASE/PHOSPHOTRANSFERASE STIMULATORY PHOSPHATASE/PHOSPHATE NONREVERSIBLE PHOSPHATASE Transmembrane domain: L159-T186, S316-V334	BLAST_DOMO BLAST_PRODOM HMMER

WO 02/40541

PCT/US01/46055

Table 3 (cont.)

SBO ID	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
9	63857A0001	374	S145, S174, S310, S332, S362, T76, T313, T73, T80	N280, N320	do WH1B; FRUIT; FY; SCARLET; LXC5200 P45844 289-650; M1-V348 LXC5200 P10090 317-666; V2-L346 do PUMPBASE; DEPENDENT; ATP; SDR10; LXC5200 P51553 406-797; VA-P340 LXC5200 S55517 406-797; VA-P340 transmembrane_domain: L418-E137, Y319-Y351	BLAST_DOMO
10	7483595CD1	443	S412, S59, T26, T369, T388	N259	transmembrane amino acid transporter aa_trans: A104-P438 ACID AMINO ACID TRANSPORTER PERMEASE TRANSMEMBRANE INHERGENTIC REGION PROLINE PD001875; C738-L367 transmembrane_domain: V115-C114, V177-T193, T211-F258, F292-L310, L326-Q344 signal_cleavage: A1-Q25	BLAST_DOMO
11	2307246CD1	521	S209, S307, S41, S50, T226, T274, T298	N234, N229	mitochondrial carrier proteins N10-PL25, S127-N220, S232-K322 mitochondrial energy transfer proteins RUC0213:P18-Q40, V177-Q189 mitochondrial energy transfer proteins signature: Q6-V101, Q6-G100, Q6-V99, Q6-L98, Q6-R97, Q6-P96, Q6-G95, Q6-L94, Q6-R94, F233-G283 TRANSPORT TRANSMEMBRANE MITOCHONDRIUM CARRIER INNER MITOCHONDRIAL ADP/ATP PD000117:SL71-E317, R14-E216, S127-F292, S127-E264, S12-E232, R184-T320, N10-E117, T231-E310	BLAST_DOMO

WO 02/00541

PCT/US01/46055

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
					MITOCHONDRIAL ENERGY TRANSFER PROTEINS DM00026 S44092 201-284:P125-E214, G71-I122, M237-E310, N10-Q40 DM00026 S44092 302-380:N237-E316, S134-E208, T68-Y112, L15-L82 DM00026 P38127 56-163:A17-F119, E263-Y309 DM00026 P38127 291-375:A239-E316, S134-L214, G73-Y112, A17-S42 signature: Mitochondrial energy transfer proteins	BLAST_D0X0
12	3033303CD1	487	S303 S347 S378 S413 S45 S481 S482 S49 S56 S6 T145 T17 T359 T265 T32 T332 T355 T374 T442 T450	N15 N23 N251 N257 N26 N312 N79	P31-L39 P253-L261 Transmembrane domains: M85-L107, V198-P221, F224-I243, F316-N336, L352-P372, T399-V421, L457-W477 Transmembrane amino acid transporter: A95-S469 ACID AMINO TRANSPORTER PERMEASE TRANSMEMBRANE INTERGENIC PROLINE TRANSPORTER PROTEIN PD13374:H341-H487	MOTIFS HMMER HMMER_PPAM BLAST_PRODOM BLAST_PRODOM

WO 02/40541

PCT/US01/46055

Table 3 (cont.)

SEQ ID NO	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signatures Sequences, Domains and Motifs	Analytical Methods and Databases
13	402/593CD1	509	S233 T267 T419 V277	N411	Transmembrane domain: P344-P366, L400-V418, F464-L483 Monocarboxylate transporter: S19-S467 TRANSPORTER: LINKED DM05037 Q03064 1-475:P7-M244, V298-L480 DM05037 P53988 1-455:P7-S194, A299-L480 DM05037 P36021 155-612:P7-V211, L300-L483	EMBL EMBL-PIR BLAST-DMO
14	747030CD1	1232	S159 S155 S253 S369 S401 S426 S554 S651 S654 S673 S887 T4178 T1186 T1211 T137 T20 T233 T491 T586 T595 T640 T728 T886 Y505 Y975	N1079 N1163 N139 N190 N172 N191 N424 N703 N764 N794 N86 N885 N92	Transmembrane domains: V112-L131, T735-V731, M812-N834 ABC transporter transmembrane region: M49-L140, V693-T943 ABC transporter: G105-G1204, S415-G539 ABC transporters family signatures: M576-S540, L4131-L4145 ATP/GTP-binding site motif A (P-loop): G422-S429, G425-S432 ABC transporters family signatures: I508-D557, I1113-D1162 ABC transporters family BL00211: L420-V431, L4191-D1162 ATP-binding transport transmembrane protein PD00111: G167-D171, C1039-I1092, G1179-L1216 MALK protein DM00150 P21448 53-386: G52-G385, F748-G988, G826-N723 MALK protein DM00150 S55692 70-399: L54-G385, F748-G988, G697-N723	HMMER HMMER-PIR HMMER-PIR MOTIFS MOTIFS ProfileScan BLMPFS-BLOCKS BLMPFS-PRODOM BLAST-DMO BLAST-DMO

Table 3 (cont.)

Seq ID	Incyte Polyptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
15	747609CD1	759	S128 S161 S336 S515 S349 S405 S47 S505 S644 S46 S663 S710 S77 S84 T12 T130 NS21 NS45 T371 T435 T628 T735		MAK protein DM00130 P21439 [61-391: L54-G385, F748-G2988 MAK protein DM00130 P23174 [61-391: L54-G385, F748-G5988, T659-T724 AUP-binding transmembrane transporter, multidrug resistance, ABC transporter, DM00130: L48-I334, F748-F933, V693-I718 P-glycoprotein, multidrug resistance, ATP-binding transporter: FD16702: T465-S524 ATP-binding transmembrane transporter DM00101: E1055-G1126 Transmembrane domain: T251-V767 Sodium/calcium exchanger: DM03297 P48376 8-969: V117-I346 (P-value = 6.0e-10)	BLAST-DMO BLAST-DMO BLAST-DMO BLAST-PRODOM BLAST-PRODOM EXMR BLAST-DMO
16	6438177CD1	283	S163 S96 T119 Y112	N26 N87 N73 N184 N218 N250 N306 N334 N393 N461 N72 N78 N79	K ⁺ channel tetramerisation domain: V58-Q155 Potassium channel signature: H100-T119 Potassium channel CDK6: SHAW: DM00490 P17972 1-102: V58-I413 (P-value = 2.8e-7)	HMUVER-PFAM BLIKMPS-PRINTS BLAST-DMO

Table 3 (cont.)

SEQ ID	Inocyte ID	Amino Acid Residues	Postnatal Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Assays
17	7477744C01	1129	S30 S47 S138 S278 S282 S443 S454 S495 S548 S708 S713 S716 S762 S813 S924 S982 S1036 S1100 S1105 S1109 S1113 S1204 T250 T254 T264 T208 T328 T334 T408 T413 T449 T645 T680 T693 T701 T704 T709 T1121 Y258 Y747		Transmembrane domains: P995-A1012, T1070-K1038 E1-22 ATPases Phosphorylation site proteins EL0015c: G144-L161, V403-7421, K563-V573, L250-L650, T811-K831 K-22 ATPases phosphorylation site: A199-V438 A199-V438	BLMPS_BLOCKS BLMPS_PRINTS BLAST_PRODUM BLAST_PRODUM BLAST_PRODUM BLAST_PRODUM

WO 02/40541

PCT/US01/46055

Table 3 (cont.)

SEQ ID NO.	Protein ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
					RTP/GTP-binding site motif A (P-loop) A271-S278	MOTIFS
					E1-E2 ATPases phosphorylation site D409-T415	MOTIFS
13	7473042CD1	548	S2 S6 S224 S262 S294 S355 S419 S474 S607 S619 S640 S645 T12 T273 T287 T327 T394 T423 T461 V11	N285 N433 N458 N485	Transmembrane domains: A79-P95, L366-V385, F395-S415 Sugar (and other) transporter: V84-F609 Sugar transport proteins E500216: G92-S103, L174-K223 Sugar transport proteins signatures: L366-S421, S162-V225 Sugar transporter signature P800171: G92-V102, L175-V194, Q336-V346, L510-V531, S533-N545 Glucose transporter signature P800172: L326-V347, L354-V385, L50- K110, L510-S533, T543-S561, S574-L594 SUGAR TRANSPORTER PROTEIN P800817: K296-R391 SUGAR TRANSPORT PROTEINS D400135 Q01440 101-433; V178-S419, L513-G589 D400135 P54723 120-434; S378-T44, D400135 S25009 121-478; G161-Q417, D400135 S10-K396 D400135 S43230 170-502; S178-T447, S506-R500	HMDBR HMDBR_YFAM BLIMPS_BLOCKS PROFLASCAN BLIMPS_PRINTS BLIMPS_PRINTS BLIMPS_PRINTS BLAST_PRODOM BLAST_PRODOM
					Sugar transport proteins signature 1: G281-G396	MOTIFS

104

SEQ ID	Incyte ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
19	7482060CDL	545	S36 S43 S128 S298 S335 S405 T185 Y269 Y297 T441 Y55	M198	Sugar transport proteins signature 2: V180-R205 Transmembrane domains: T165-L182, V269-A295, V306-L325 Ion transport protein: T175-T390 K+ channel tetramerisation domain: A9-L416 Potassium channel signature P400169: B60-G79, A157-T185, I208-K228, F231-V251, Y275-C301, E304-E327, F339-N351, G368-F394 CHANNEL IONIC PROTEIN POTASSIUM SUBUNIT VOLUNTEERED TRANSMEMBRANE CALCIUM TRANSPORT ION P400144: F231-Y298 CHANNEL; POTASSIUM; CDSK; SHAM; DM00490 JH0595 26-142: V11-R115 DM00490 P15337 13-134: R5-R115 DM00436 JH0595 144-307: A163-L278 DM00490 P17970 268-384: V11-S115 Signal cleavage: M-345 Signal peptide: M-A15 SVAS domain (Sulphate transporter and Acetasma factor antagonist): H110-A236 SULFATE TRANSPORTER PROTEIN TRANSPORT TRANSMEMBRANE AFFINITY GLYCOPROTEIN SULFATE HIGH DISEASE P0001755: H110-A236 E-value: 3.0e-08	HMMER HMMER_P7AM BLAST_PRODOM
20	1578772CDL	262	S57 S154 S232 S249 W53 T178		SPSCAN HMMER	

WO 02/40541

PCT/US01/46055

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
21	1626101C01	1373	577-657	6123382H1 (BRACHOR05) 2822668H1 (ADRETR06) 22065374V1 2378333W6 (SLTNOR01) 7276683H1 (LIVRDIS04) 5508488M1 51134156V1 7602865J1 (BSOCTM01) 7592760J1 (SLDPT001) 7448609H1 (GOLNOR05) 7782700J1 (ORRT0001) 8308495H1 (SLTNOR01) 7764212H1 (ESOOT002)	220 1 322 822 2591 1325 468 1914 2573 2578 1281 1	880 305 1036 1373 2331 1931 1137 1023 2573 2582 3108 1833 582
22	2907828C01	3231	1-94, 2624-2681, 807-866	GRI_610377937.edit 7068888H1 (BRACHOR01) 7756687H1 (SPLNTR01) 70697015L (BRACHOR02) 7039903H1 (UTRSTMR02) 7759687J1 (SPLNTR01) 8104892H1 (MIXDIE02) 55052339H1 7097441H1 (BRACHOR02) 7032041H1 (BRACHOR02) GRI_68748693_000007.edit GRI_68748693_000006_00 0003.edit GRI_68748693_000003_00 0004.resequenc.edit	1 2443 1389 2679 155 3047 43 380 2098 1919 677 1371 1705 1463	429 3108 2005 3160 778 2527 363 4204 2643 1580 2732 1384 3538 2116 1871
23	3968527C01	3160	2860-3160, 1-434, 1498-1790			
24	7472732C01	2848	2653-2848, 2455-2571, 917-1609, 1-72, 150- 195			

WO 02/40541

PCT/US01/46055

Table 4 (cont.)

Polynucleotide Seq ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
25	7476538CBI	3727	1-1490	55061546H1	1	427
				55061546H1	116	728
				2467913W6 (TEYRNO708)	2098	2755
				2467913W6 (CYXINO708)	1405	1945
				688280HE (MIXONB01)	2236	2848
				4107126H1 (BRSTFUT1)	3169	3443
				3554480H1 (BRATFNO5)	2598	3198
				8244920H1 (DQNEUMR01)	1879	2510
				7106310V1	1741	2403
				5292823H1 (SLNTHNR02)	567	1257
				717144H1 (BRSTFNO1)	156	613
				7228355H1 (LQNGNCO1)	2496	3097
				8148855H1 (MIXTNEB1)	424	840
				6799132H1 (COLERNR03)	3188	3727
				8190281H1 (BRATFNR03)	1329	1773
26	8128531CBI	2571	1-925	7963983H1 (SPLNPEA02)	1201	1697
				QNN_g9187762_004_edit	1	539
				3467659H1 (TATRNOR04)	2235	2571
				70010270D1	909	5112
				4560138F6 (845F1UT22)	953	1509
				2237089F6 (PROSROT16)	1644	2204
				4341662H1 (BRAUTROT2)	1	266
				6883871V1 (BRAUTROT3)	477	945
				2332084F6 (PROSROT16)	1991	2478
				4001357F6 (BRW2AZS07)	1372	1963
27	7476575CBI	1660	1107-1660, 490-803	8128531H1 (SCOMETC01)	150	803
				GNM_g712223_000033_00	667	1660
				5511643H1	1	717

WO 02/40541

PCT/US01/46055

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyle Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
28	266243CB1	2743	1-155, 2720-2743, 2490-2605	7765596d1 (URETU0E01) 7629030h1 (GELADIE01) 71153607v1 7279379h1 (BMAKXE01) 6618283j1 (BRAUTDR01) 7629030j1 (GELADIE01) GNN-98575919_008	2067 555 1709 1169 563 5 1	2743 1237 2171 1778 1096 660 1333
29	668571CCB1	3239	1-899, 2163-2217, 2516-2663	72460968d1 72458459d1 71976815v1 72463146d1 72461236d1 71977010v1 72462439d1 71875953v1 G1077361	1862 1332 3433 580 1 2584 968 2740 1261	1862 2068 3280 1190 669 3231 1611 3239 1096
30	7483599CB1	1615	749-823, 114-353	EN970000023927 FL7483599_97708819_000 O10_97293314_1_2-3 FL7483599_97708819_000 O10_97293314_1_3-4 FL7483599_97708819_000 O10_97293314_1_4-5 FL7483599_97708819_000 O10_97293314_1_7 FL7483599_97708819_000 O10_97293314_1_5-6 FL7483599_97708819_000 O10_97293314_1_9 g2077387	826 943 134 421 421 679 521 925 1015	420 517 606 825 663 1332 1418

WO 02/40541

PCT/US01/46055

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
31	2507246CE1	1245	923-1245	CRM: g7417485_000010_00 2	1	1332
				FL7483599_g7708819_000 010_g7293314_1_1-2	115	333
				71424095V1	300	1021
				50493686 (TMR3DT02)	739	1345
32	3033505CE1	4169	793-1236 2297-2327 4074-4169 3031-3429	354532F1 (RAT390T01)	619	1737
				7222943_01	1	481
				5045025_01 (BRABDI02)	3278	3674
				71530266V1 (BRAIF08)	656	1046
				4527568V2 (PROSWT01)	55	158
				410431356 (BRSTU01)	3125	3636
				4529404H1 (DTHXT01)	1	286
				71029133V1	2442	3042
				48760586 (HNTAGT01)	1194	1700
				5859947H1 (BRAIF08)	1945	2572
				4324588H1 (TLYMNT01)	557	810
				7103133H1 (BRAWTR02)	877	1301
				7288253H1 (BRAIF06)	3584	4169
				5913364J1 (PTUDIR01)	2564	3174
33	4027691CE1	3440	2635-2665, 558-1546	3033505F6 (OLYXOT05)	1354	1913
				7124694V1	1864	2523
				2866257F6 (KIDNKT02)	2069	2700
				7365929H1 (OVARIC01)	2923	3400
				7617003H1 (KIDNV001)	2423	3031
				5501287H1 (BRABDI01)	450	614
				21485147	1	451
				70503458V1	1705	2322
				7648766V1 (STOMYD01)	1193	1833
				70618523V1	1489	2263
34	4027691CE1	3440	2635-2665, 558-1546	3107576H1 (PTAUDI02)	582	1536

WO 02/40541

PCT/US01/46055

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
34	7472030CBL	3599	2728-3235, 2347-2550, 1-982, 3286-3699, 1712-2259	7163817H1 (OVAR1001)	733	1373
				2866257H1 (KIDN020)	2068	2402
				63117430	211	561
				124560386 (LUNG0003)	1359	1710
				217472030_G10445386_g3	2347	2651
				07181_1_7-8		
				PT7472030_G10445386_g3	2551	2792
				07181_1_8-9		
				5600293H1 (STES00128)	1815	2263
				8180328H1 (MYELO001)	541	1552
35	7476089CBL	2428	1-851, 1225-2428	G21-8308130_G1045386	1	5699
				edit		
				227476089_58656012_g59	1300	2346
				Q2358		
				58018675J1	1702	2428
				5002353J1	1	583
				5200775J1	897	1207
				72855686 (BRALFER06)	512	567
				193118756 (CORNF002)	1264	1808
				7924964H1 (COLN0002)	1	592
36	8428177CBL	2243	1052-1089, 1737-2243	8055028J1 (LIVAT001)	616	1137
				2132181H1 (OVAR00703)	1995	2243
				6437511H1 (BRAL00702)	1412	1968
				739342688 (BRAL01203)	351	986
				3234007H2 (COLN00103)	1862	2112
				8218958J2 (SINTFER02)	496	1699
				55110512J1	252	1056
				55110612J1	1325	2135
				56000471J1	1	498
				8228559621 (BRAL0002)	2417	3044
37	7477243CBL	3711	1-673, 2313-3058, 3678-3711	49833287 (BRALFER05)	1911	2668

WO 02/06541

PCT/US01/46055

Table 4 (cont.)

Polynucleotide Seq. ID NO.	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
38	7473042CB1	2717	1-607, 1986-2406	3863115T6 (XIMNOT020)	1051	3711
				55155912J1	1730	2139
				5845787F8 (LIVR03008)	2646	3327
				6356661F6 (EMART006)	556	1444
				56008775J1	965	1762
39	7482056CB1	2235	827-859, 1642-2235, 1525-1551	72622010V1	532	1238
				GRN 87008856_000017_00 2	34	776
				21797555V1	1628	2287
				6195488B2 (BRAID0601)	316	1087
				72457143DA	2059	2717
				5514028H1	1	253
				55061745J1	282	293
				GRN 69454649_000007_00	428	2065
				0012	1	871
				6770140H8 (BRAID0601)	1	871
40	1578772CB1	2563	710-845, 1- 138, 2543- 2563	5923423H1 (BRAID0601)	1932	2235
				6770140F8 (BRAID0601)	2416	1648
				5402544H1 (BRAID0601)	549	807
				1619843T6 (BRAID0601)	1940	2563
				70880785V1	769	1319
				6449038H1 (BRAID0601)	1275	1715
				2083536H1 (UTASNOT08)	1122	1380
				6536573H1 (OVAR002)	1	475
				275594686 (THP1AZS08)	1445	1833
				6553014H1 (BRAID0601)	1531	2121
41	7239234H1 (BRAID0601)	238		7239234H1 (BRAID0601)	238	689

WO 02/40541

PCT/US01/46055

Table 5

Polynucleotide SEQ ID NO:	Locate Project ID	Representative Library
21	1625101CB1	NOSED101
22	2907828CB1	DCCL5101
23	3969527CB1	SFLNT001
24	7472732CB1	THVRN008
25	7475918CB1	LIVRN003
26	8128531CB1	THVRN008
27	7476757CB1	ESGXT001
28	266243CB1	EMART001
29	6585710CB1	SINTN021
31	2507246CB1	LAUHU002
32	3331505CB1	THVRN005
33	4027633CB1	KIERN010
34	7472030CB1	LUGRN003
35	7476039CB1	ESGXT006
36	6428177CB1	ESGXT002
37	7477265CB1	EMART006
38	7473042CB1	CTTEN010
39	7482060CB1	BRANR001
40	1578772CB1	BRAITUT12

WO 02/40541

PCT/US01/46055

Table 6

Library	Vector	Library Description
BRATME01	pINCY	This 5' biased random primed library was constructed using RNA isolated from treated SH-SY5Y cells derived from a metastatic bone marrow neuroblastoma, removed from a 4-year-old Caucasian female (Scherling AG). The medium was MEM/10% F12 with 10% fetal calf serum. After reaching about 80% confluency cells were treated with 5-hydroxydecanal (5-OHDA) at 100 μ M for 8 hours.
BRATNN06	pINCY	Library was constructed using RNA isolated from an untreated SH-SY5Y cell line derived from bone marrow neuroblastoma tumor cells removed from a 4-year-old Caucasian female.
BRABDTE02	pINCY	This 3' biased random primed library was constructed using RNA isolated from diseased cerebellum tissue removed from the brain of a 57-year-old Caucasian male who died from a cerebrovascular accident. Serologies were negative. Patient history included Huntington's disease, emphysema, and tobacco abuse (3-4 packs per day, for 40 years).
BRATPER06	PCDNA2.1	This random primed library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus who was stillborn with a hypoplastic left heart at 23 weeks' gestation. Serologies were negative.
BRATUTL2	pINCY	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 40-year-old Caucasian female during excision of a cerebral meningioma lesion. Pathology indicated grade 1 meningioma.
BRATNG01	pINCY	This random primed library was constructed using RNA isolated from striatum, globus pallidus and posterior putamen tissue removed from an 81-year-old Caucasian female who died from a hemorrhage and ruptured thoracic aorta due to atherosclerosis. Pathology indicated moderate atherosclerosis involving the internal carotids, bilaterally; microscopic infarcts of the frontal cortex and hippocampus; and scattered diffuse amyloid plaques and neurofibrillary tangles, consistent with age. Grossly, the septum showed only mild thickening and hyalinization along the superior sagittal sinus. The remainder of the septum and hippocampus were unremarkable and contained some congested blood vessels. Mild atrophy was noted in the frontal poles and lobes, and temporal lobes, bilaterally. Microscopically there were small, type II Alzheimer type II astrocytes within the deep layers of the neocortex. There was increased satellitosis around neurons in the deep grey matter in the middle frontal cortex. The amygdala contained rare diffuse plaques and neurofibrillary tangles. The posterior hippocampus contained a

WO 02/40541

PCT/US01/46055

Table 6 (cont.)

Library	Vector	Library Description
ES02UE01	pINCY	microscopic area of cystic cavitation with hemosiderin-laden macrophages surrounded by reactive gliosis. Patient history included sepsis, cholangitis, post-operative atelectasis, pneumonia QAD, cardiomegaly due to left ventricular hypertrophy, splenomegaly, arteriolonephrosclerosis, nodular colloid goiter, emphysema, CHF, hypothyroidism, and peripheral vascular disease. This 5' biased random primed library was constructed using RNA isolated from esophageal tumor tissue removed from a 61-year-old Caucasian male during a partial esophagectomy, proximal gastrectomy, pyloromyotomy, and regional lymph node excision. Pathology indicated an invasive grade 3 adenocarcinoma in the esophagus extending distally to involve the gastroesophageal junction. The tumor extended through the muscularis to involve the periesophageal and benign gastric soft tissues. One perigastric and two periesophageal lymph nodes were positive for tumor. There were multiple perigastric and periesophageal tumor implants. The patient presented with deficiency anemia and myelodysplasia. Patient history included hyperlipidemia, and tobacco and alcohol abuse in remission. Previous surgeries included adenotomyllectomy, rhinoplasty, vasectomy, and hemorrhoidectomy. A previous bone marrow aspiration found the marrow to be hypercellular for age and had a cellularity-to-fat ratio of 95:5. The marrow was focally densely fibrotic. Granulocytic precursors were slightly increased with normal maturation. The estimate of blast cells was greater than 5%. Megakaryocytes were increased and appeared atypical in clusters. Stomach cells and granulomata were absent. Patient medications included Enoxacin, Danocrine, Berocca Plus tablets, Selenium, vitamin B6 phosphate, vitamins E & C, and beta carotene. Family history included alcohol abuse, atherosclerotic coronary artery disease, type II diabetes, chronic liver disease, and primary cardiomyopathy in the father, and benign hypertension and cerebrovascular disease in the mother.
KIDEN0020	pINCY	Library was constructed using RNA isolated from left kidney tissue removed from a 43-year-old Caucasian male during nephroureterectomy, regional lymph node excision, and unilateral left adrenalectomy. Pathology for the associated tumor tissue indicated a grade 2 renal cell carcinoma. Family history included atherosclerotic coronary artery disease.

WO 02/40541

PCT/US01/46055

Table 6 (cont.)

Library	Vector	Library Description
LAWR0002	pINCY	Library was constructed using RNA isolated from a myxoma removed from the left atrium of a 43-year-old Caucasian male during amnioplasty. Pathology indicated atrial myxoma. Patient history included pulmonary insufficiency, acute myocardial infarction, atherosclerotic coronary artery disease, hyperlipidemia, and tobacco use. Family history included benign hypertension, acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
LIVR0003	pINCY	Library was constructed using RNA isolated from liver tissue removed from a Caucasian male fetus, who died from Patau's syndrome (trisomy 13) at 20 weeks' gestation.
LUNG0003	PSPORT1	Library was constructed using RNA isolated from lung tissue of a 79-year-old Caucasian male. Pathology for the associated tumor tissue indicated grade 4 carcinoma. Patient history included a benign prostate neoplasm and atherosclerosis.
NOSEDING1	pINCY	This normalized nasal polyp tissue library was constructed from 1.08 million independent clones from a pooled nasal polyp tissue library. Starting RNA was made from pooled cDNA from two donors. cDNA was generated using mRNA isolated from a nasal polyp removed from a 78-year-old Caucasian male during nasal polypectomy (donor A) and from nasal polyps from another donor (donor B). Pathology (A) indicated a nasal polyp and striking eosinophilia, especially deep in the epithelium. In many instances, eosinophils were undergoing frank necrosis with striking deposition of Charcot-Leyden crystals. Foci of eosinophil infiltration in appearance of cells were seen in certain areas and those areas closer to the appearance of respiratory epithelium showed loss of surface epithelium. This examination of respiratory epithelium showed loss of surface epithelium. This nasal polyp showed typical histology for polypoid change associated with allergic disease. Patient history included asthma allergy tests (which were positive for histamine but negative for common substances), eosinophilic function test (EFT) which showed reduction in the forced expiratory volume (FEV ₁) with increased use of a bronchodilator, and nasal polyps. Patient history (A) included asthma. Previous surgery (A) included a nasal polypectomy. The patient was not using glucocorticoids in treatment for asthma. The library was normalized in 1 round using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Donald

WO 02/40541

PCT/US01/46055

Table 6 (cont.)

Library	Vector	Library Description
SINT0021	PINCY	et al., Genome Research 6 (1996):751, except that a significantly longer (48 hours/round) reannealing hybridization was used.
SPUN0001	PCDNA2.1	Library was constructed using RNA isolated from small intestine tissue obtained from a 8-year-old Black male, who died from anoxia. Serology was negative. This 5' biased random primed library was constructed using RNA isolated from spleen tumor tissue removed from a 28-year-old male during total splenectomy. Pathology indicated malignant lymphoma, diffuse large cell type, B-cell phenotype with abundant reactive T-cells and marked granulomatous response involving the spleen, where it formed approximately 45 nodules, liver, and multiple lymph nodes.
THY00008	PINCY	Library was constructed using RNA isolated from thymus tissue removed from a 4-month-old Caucasian male during a total thymectomy and open heart repair of aortic ventricular canal defect using hypothermia. Pathology indicated a grossly normal thymus. The patient presented with a congenital heart anomaly, congestive heart failure, and Down's syndrome. Patient history included abnormal thyroid function study and premature birth. Previous procedures included right and left heart angiography. Patient medications included Digoxin, Synthroid, and Lasix.
THY00008	PINCY	Library was constructed using RNA isolated from the diseased left thyroid tissue removed from a 33-year-old Caucasian female during a complete thyroidectomy. Pathology indicated lymphocytic thyroiditis. Pathology for the matched tumor tissue indicated grade 1 papillary carcinoma. Multiple lymph nodes from the right, left, and midline section of the neck were negative for tumor. Fragments of the thymus were benign. Fibroadipose tissue was identified in the right anterior and superior parathyroid regions. Multiple lymph nodes (2 of 6) from the right side of the neck contained microscopic foci of metastatic papillary carcinoma. Patient history included attention deficit disorder with hyperactivity. Previous surgeries included an operative procedure on the external ear. Patient medications included Prozac. Family history included chronic obstructive aschma in the mother; alcohol abuse, benign hypertension, and depressive disorder in the grandparents; and attention deficit disorder with hyperactivity in the siblings.

WO 02/40541

PCT/US01/46055

Table 6 (cont.)

Library	Vector	Library Description
UWMC005	pUNCY	Library was constructed using RNA isolated from nonactivated Tp2 cells. These cells were differentiated from umbilical cord CD4 ⁺ cells with IL-4 in the presence of anti-IL-12 antibodies and 87-transfected COS cells.
UWMC5701	PL4DESCRIFT	Library was constructed using RNA isolated from mononuclear cells obtained from the umbilical cord blood of 12 individuals. The cells were cultured for 12 days with IL-5 before RNA was obtained from the pooled lysates.
UTWEND010	pUNCY	Library was constructed using RNA isolated from pooled uterine endometrial tissue removed from three adult females during endometrial biopsy. Pathology indicated normal endometrium. All three patients were positive for Her2-3 in-situ.

WO 02/40541

PCT/US01/46055

Table 7

Program	Description	Reference	Parameter Threshold
ABI FAC TURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PRACCEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA;	Mismatch <5%
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
PASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. PASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, 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=1.06E-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: fastx score=1.00 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. 19:5565-5572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Atwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nussell, Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater

WO 02/40541

PCT/US01/46055

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileSam	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Grisham, M. et al. (1988) CABIOS 4:61-66; Grisham, M. et al. (1989) Methods Enzymol. 183:166-199; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:2717-2721.	Normalized quality scores GCG-Specified "HIGLI" value for that particular Prosite motif. Generally, score=1-42.1.
Pfired	A base-calling algorithm that examines automated sequence traces with high sensitivity and probability.	Erving, B. et al. (1998) Genome Res. 8:175-183; Erving, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-490; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= (20 or greater; Match length= 56 or greater
Corsed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1996) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nelson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	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.	
TDHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth 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. 25:2717-2721; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 02/40541

PCT/US01/46055

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:1-20,
 - 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:1-20,
 - c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, and
 - d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.
2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.
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:21-40.
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 02/40541

PCT/US01/46055

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:1-20.

5

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

12. An isolated polynucleotide selected from the group consisting of:

- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:21-40,
- 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:21-40,
- c) a polynucleotide complementary to a polynucleotide of a),
- d) a polynucleotide complementary to a polynucleotide of b), and
- e) an RNA equivalent of a)-d).

10

15

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:

- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

25

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:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain

35

WO 02/40541

PCT/US01/46055

- 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:1-20.
- 10 19. A method for treating a disease or condition associated with decreased expression of functional TRICH, 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 TRICH, comprising administering to a patient in need of such treatment a composition of
- 25 claim 21.
23. A method of screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:
- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- 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 02/40541

PCT/US01/46055

TRICH, 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 amounts of the compound and in the absence of the compound.

30 29. A method of assessing toxicity of a test compound, the method comprising:

- a) treating a biological sample containing nucleic acids with the test compound,
- b) hybridizing the nucleic acids of the treated biological sample with a probe comprising
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 02/40541

PCT/US01/46055

- 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 TRICH 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 TRICH 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 TRICH in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

WO 02/40541

PCT/US01/46055

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:
- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
 - b) isolating antibodies from said animal, and
 - c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.
37. A polyclonal antibody produced by a method of claim 36.
38. A composition comprising the polyclonal antibody of claim 37 and a suitable carrier.
39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:
- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-20, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
 - b) isolating antibody producing cells from the animal,
 - c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
 - d) culturing the hybridoma cells, and
 - e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.
40. A monoclonal antibody produced by a method of claim 39.
41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.
42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.
43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant

WO 02/40541

PCT/US01/46055

immunoglobulin library.

44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-20 in a sample, the method comprising:

- 5 a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-20 in the sample.

10

45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-20 from a sample, the method comprising:

- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
- 15 b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-20.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 20 13.

47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:

- a) labeling the polynucleotides of the sample,
- 25 b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and
- c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

35

WO 02/40541

PCT/US01/46055

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

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

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

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

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

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

WO 02/40541

PCT/US01/46055

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.
63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.
- 5 64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.
65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.
66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.
- 10 67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
- 15 70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.
71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.
- 20 72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.
73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.
- 25 74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.
75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.
76. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
30 NO:21.
77. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:22.
- 35 78. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID

WO 02/40541

PCT/US01/46055

NO:23.

79. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:24.

5

80. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:25.

81. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
10 NO:26.

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
15 NO:28.

84. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:29.

20

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
25 NO:31.

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
30 NO:33.

89. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:34.

35

WO 02/40541

PCT/US01/46055

90. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:35.

91. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
5 NO:36.

92. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:37.

10 93. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:38.

94. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:39.

15 95. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID
NO:40.

WO 02/40541

PCT/US01/46055

<110> INCVTE GENOMICS, INC.

TANG, Y. Tom
 YUB, Henry
 NGUYEN, Darniel B.
 HAFALIA, April J.A.
 ELLIOTT, Vicki S.
 LU, Yan
 WALIA, Narinder K.
 YAO, Monique G.
 BAUGHN, Mariah R.
 GANDHI, Ameena R.
 DING, Li
 SRINIVASIA, Madhusudan
 RAMKUMAR, Jayalaxmi
 ARVIZU, Chandra
 GIETZEN, Kimberly J.
 LAL, Preeti G.
 AZHMEAI, Valda
 KHAN, Farrah A.
 THANGAVELU, Kavitha
 THORNTON, Michael
 LU, Dyung Aina M.
 TRIBOULEY, Catherine M.
 WARREN, Bridget A.
 TSON, H. Craig
 DAS, Debopriya
 RAUWANA, Brigitte E.
 POLICKY, Jennifer L.
 KEARNEY, Liam

<120> TRANSPORTERS AND ION CHANNELS

<130> PI-0270 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/243,989; 60/245,904; 60/247,673; 60/249,661; 60/252,332
 60/250,790

<151> 2000-10-27; 2000-11-03; 2000-11-09; 2000-11-17; 2000-11-20;
 2000-12-01

<160> 40

<170> PERL Program

<210> 1

<211> 337

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1626101CD1

<400> 1

Met	Ser	Leu	Glu	Gln	Glu	Glu	Glu	Thr	Gln	Pro	Gly	Arg	Leu	Leu
1			5						10				15	
Gly	Arg	Arg	Asp	Ala	Val	Pro	Ala	Phe	Ile	Glu	Pro	Asn	Val	Arg
			20						25				30	
Phe	Trp	Ile	Thr	Glu	Arg	Gln	Ser	Phe	Ile	Arg	Arg	Phe	Leu	Gln
			35						40				45	
Trp	Thr	Glu	Leu	Leu	Asp	Pro	Thr	Asn	Val	Phe	Ile	Ser	Val	Glu

WO 02/06541

PCT/US01/46055

```

50          55          60
Ser Ile Glu Asn Ser Arg Gln Leu Leu Cys Thr Asn Glu Asp Val
65          70          75
Ser Ser Pro Ala Ser Ala Asp Gln Arg Ile Gln Glu Ala Trp Lys
80          85          90
Arg Ser Leu Ala Thr Val His Pro Asp Ser Ser Asn Leu Ile Pro
95          100          105
Lys Leu Phe Arg Pro Ala Ala Phe Leu Pro Phe Met Ala Pro Thr
110          115          120
Val Phe Leu Ser Met Thr Pro Leu Lys Gly Ile Lys Ser Val Ile
125          130          135
Leu Pro Gln Val Phe Leu Cys Ala Tyr Met Ala Ala Phe Asn Ser
140          145          150
Ile Asn Gly Asn Arg Ser Tyr Thr Cys Lys Pro Leu Glu Arg Ser
155          160          165
Leu Leu Met Ala Gly Ala Val Ala Ser Ser Thr Phe Leu Gly Val
170          175          180
Ile Pro Gln Phe Val Gln Met Lys Tyr Gly Leu Thr Gly Pro Trp
185          190          195
Ile Lys Arg Leu Leu Pro Val Ile Phe Leu Val Gln Ala Ser Gly
200          205          210
Met Asn Val Tyr Met Ser Arg Ser Leu Glu Ser Ile Lys Gly Ile
215          220          225
Ala Val Met Asp Lys Glu Gly Asn Val Leu Gly His Ser Arg Ile
230          235          240
Ala Gly Thr Lys Ala Val Arg Glu Thr Leu Ala Ser Arg Ile Val
245          250          255
Leu Phe Gly Thr Ser Ala Leu Ile Pro Glu Val Phe Thr Tyr Phe
260          265          270
Phe Lys Arg Thr Gln Tyr Phe Arg Lys Asn Pro Gly Ser Leu Trp
275          280          285
Ile Leu Lys Leu Ser Cys Thr Val Leu Ala Met Gly Leu Met Val
290          295          300
Pro Phe Ser Phe Ser Ile Phe Pro Gln Ile Gly Gln Ile Gln Tyr
305          310          315
Cys Ser Leu Glu Glu Lys Ile Gln Ser Pro Thr Glu Glu Thr Glu
320          325          330
Ile Phe Tyr His Arg Gly Val
335

```

```

<210> 2
<211> 816
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2907828CD1

```

```

<400> 2
Met Ala Val Ser Leu Asp Asp Asp Val Pro Leu Ile Leu Thr Leu
1          5          10          15
Asp Glu Gly Gly Ser Ala Pro Leu Ala Pro Ser Asn Gly Leu Gly
20          25          30
Gln Glu Glu Leu Pro Ser Lys Asn Gly Gly Ser Tyr Ala Ile His
35          40          45
Asp Ser Gln Ala Pro Ser Leu Ser Ser Gly Gly Glu Ser Ser Pro
50          55          60
Ser Ser Pro Ala His Asn Trp Glu Met Asn Tyr Gln Glu Ala Ala
65          70          75
Ile Tyr Leu Gln Glu Gly Glu Asn Asn Asp Lys Phe Phe Thr His
80          85          90
Pro Lys Asp Ala Lys Ala Leu Ala Ala Tyr Leu Phe Ala His Asn

```

WO 02/0541

PCT/US01/46055

```

95      100      105
His Leu Phe Tyr Leu Met Glu Leu Ala Thr Ala Leu Leu Leu Leu
110      115      120
Leu Leu Ser Leu Cys Glu Ala Pro Ala Val Pro Ala Leu Arg Leu
125      130      135
Gly Ile Tyr Val His Ala Thr Leu Glu Leu Phe Ala Leu Met Val
140      145      150
Val Val Phe Glu Leu Cys Met Lys Leu Arg Trp Leu Gly Leu His
155      160      165
Thr Phe Ile Arg His Lys Arg Thr Met Val Lys Thr Ser Val Leu
170      175      180
Val Val Gln Phe Val Glu Ala Ile Val Val Leu Val Arg Gln Met
185      190      195
Ser His Val Arg Val Thr Arg Ala Leu Arg Cys Ile Phe Leu Val
200      205      210
Asp Cys Arg Tyr Cys Gly Gly Val Arg Arg Asn Leu Arg Gln Ile
215      220      225
Phe Gln Ser Leu Pro Pro Phe Met Asp Ile Leu Leu Leu Leu Leu
230      235      240
Phe Phe Met Ile Ile Phe Ala Ile Leu Gly Phe Tyr Leu Phe Ser
245      250      255
Pro Asn Pro Ser Asp Pro Tyr Phe Ser Thr Leu Glu Asn Ser Ile
260      265      270
Val Ser Leu Phe Val Leu Leu Thr Thr Ala Asn Phe Pro Asp Val
275      280      285
Met Met Pro Ser Tyr Ser Arg Asn Pro Trp Ser Cys Val Phe Phe
290      295      300
Ile Val Tyr Leu Ser Ile Glu Leu Tyr Phe Ile Met Asn Leu Leu
305      310      315
Leu Ala Val Val Phe Asp Thr Phe Asn Asp Ile Glu Lys Arg Lys
320      325      330
Phe Lys Ser Leu Leu Leu His Lys Arg Thr Ala Ile Gln His Ala
335      340      345
Tyr Arg Leu Leu Ile Ser Gln Arg Arg Pro Ala Gly Ile Ser Tyr
350      355      360
Arg Gln Phe Glu Gly Leu Met Arg Phe Tyr Lys Pro Arg Met Ser
365      370      375
Ala Arg Glu Arg Tyr Leu Thr Phe Lys Ala Leu Asn Gln Asn Asn
380      385      390
Thr Pro Leu Leu Ser Leu Lys Asp Phe Tyr Asp Ile Tyr Glu Val
395      400      405
Ala Ala Leu Lys Trp Lys Ala Lys Lys Asn Arg Glu His Trp Phe
410      415      420
Asp Glu Leu Pro Arg Thr Ala Leu Leu Ile Phe Lys Gly Ile Asn
425      430      435
Ile Leu Val Lys Ser Lys Ala Phe Gln Tyr Phe Met Tyr Leu Val
440      445      450
Val Ala Val Asn Gly Val Trp Ile Leu Val Glu Thr Phe Met Leu
455      460      465
Lys Gly Gly Asn Phe Phe Ser Lys His Val Pro Trp Ser Tyr Leu
470      475      480
Val Phe Leu Thr Ile Tyr Gly Val Glu Leu Phe Leu Lys Val Ala
485      490      495
Gly Leu Gly Pro Val Glu Tyr Leu Ser Ser Gly Trp Asn Leu Phe
500      505      510
Asp Phe Ser Val Thr Val Phe Ala Phe Leu Gly Leu Leu Ala Leu
515      520      525
Ala Leu Asn Met Glu Pro Phe Tyr Phe Ile Val Val Leu Arg Pro
530      535      540
Leu Gln Leu Leu Arg Leu Phe Lys Leu Lys Glu Arg Tyr Arg Asn
545      550      555
Val Leu Asp Thr Met Phe Glu Leu Leu Pro Arg Met Ala Ser Leu
560      565      570

```

WO 02/40541

PCT/US01/46055

Gly Leu Thr Leu Leu Ile Phe Tyr Tyr Ser Phe Ala Ile Val Gly
 575 580 585
 Met Glu Phe Phe Cys Gly Ile Val Phe Pro Asn Cys Cys Asn Thr
 590 595 600
 Ser Thr Val Ala Asp Ala Tyr Arg Trp Arg Asn His Thr Val Gly
 605 610 615
 Asn Arg Thr Val Val Glu Glu Gly Tyr Tyr Tyr Leu Asn Asn Phe
 620 625 630
 Asp Asn Ile Leu Asn Ser Phe Val Thr Leu Phe Glu Leu Thr Val
 635 640 645
 Val Asn Asn Trp Tyr Ile Ile Met Glu Gly Val Thr Ser Gln Thr
 650 655 660
 Ser His Trp Ser Arg Leu Tyr Phe Met Thr Phe Tyr Ile Val Thr
 665 670 675
 Met Val Val Met Thr Ile Ile Val Ala Phe Ile Leu Glu Ala Phe
 680 685 690
 Val Phe Arg Met Asn Tyr Ser Arg Lys Asn Gln Asp Ser Glu Val
 695 700 705
 Asp Gly Gly Ile Thr Leu Glu Lys Glu Ile Ser Lys Glu Glu Leu
 710 715 720
 Val Ala Val Leu Glu Leu Tyr Arg Glu Ala Arg Gly Ala Ser Ser
 725 730 735
 Asp Val Thr Arg Leu Leu Glu Thr Leu Ser Gln Met Glu Arg Tyr
 740 745 750
 Gln Gln His Ser Met Val Phe Leu Gly Arg Arg Ser Arg Thr Lys
 755 760 765
 Ser Asp Leu Ser Leu Lys Met Tyr Gln Glu Glu Ile Gln Glu Trp
 770 775 780
 Tyr Glu Glu His Ala Arg Glu Gln Glu Gln Gln Arg Gln Leu Ser
 785 790 795
 Ser Ser Ala Ala Pro Ala Ala Gln Gln Pro Pro Gly Ser Arg Gln
 800 805 810
 Arg Ser Gln Thr Val Thr
 815

<210> 3

<211> 1047

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3968527CD1

<400> 3

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

WO 02/06541

PCT/US01/46055

Ser Gln Val Tyr Ser Arg Leu Thr Ala Arg Gly Thr Val Lys Val
 140 145 150
 Lys Ser Ser Asn Ile Gln Val Gly Asp Leu Ile Ile Val Glu Lys
 155 160 165
 Asn Gln Arg Val Pro Ala Asp Met Ile Phe Leu Arg Thr Ser Glu
 170 175 180
 Lys Asn Gly Ser Cys Phe Leu Arg Thr Asp Gln Leu Asp Gly Glu
 185 190 195
 Thr Asp Trp Lys Leu Arg Leu Pro Val Ala Cys Thr Gln Arg Leu
 200 205 210
 Pro Thr Ala Ala Asp Leu Leu Gln Ile Arg Ser Tyr Val Tyr Ala
 215 220 225
 Glu Glu Pro Asn Ile Asp Ile His Asn Phe Val Gly Thr Phe Thr
 230 235 240
 Arg Glu Asp Ser Asp Pro Pro Ile Ser Glu Ser Leu Ser Ile Glu
 245 250 255
 Asn Thr Leu Trp Ala Gly Thr Val Val Ala Ser Gly Thr Val Val
 260 265 270
 Gly Val Val Leu Tyr Thr Gly Arg Glu Leu Arg Ser Val Met Asn
 275 280 285
 Thr Ser Asn Pro Arg Ser Lys Ile Gly Leu Phe Asp Leu Glu Val
 290 295 300
 Asn Cys Leu Thr Lys Ile Leu Phe Gly Ala Leu Val Val Val Ser
 305 310 315
 Leu Val Met Val Ala Leu Gln His Phe Ala Gly Arg Trp Tyr Leu
 320 325 330
 Gln Ile Ile Arg Phe Leu Leu Leu Phe Ser Asn Ile Ile Pro Ile
 335 340 345
 Ser Leu Arg Val Asn Leu Asp Met Gly Lys Ile Val Tyr Ser Trp
 350 355 360
 Val Ile Arg Arg Asp Ser Lys Ile Pro Gly Thr Val Val Arg Ser
 365 370 375
 Ser Thr Ile Pro Gln Gln Leu Gly Arg Ile Ser Tyr Leu Leu Thr
 380 385 390
 Asp Lys Thr Gly Thr Leu Thr Gln Asn Glu Met Ile Phe Lys Arg
 395 400 405
 Leu His Leu Gly Thr Val Ala Tyr Gly Leu Asp Ser Met Asp Glu
 410 415 420
 Val Gln Ser His Ile Phe Ser Ile Tyr Thr Gln Gln Ser Gln Asp
 425 430 435
 Pro Pro Ala Gln Lys Gly Pro Thr Leu Thr Thr Lys Val Arg Arg
 440 445 450
 Thr Met Ser Ser Arg Val His Glu Ala Val Lys Ala Ile Ala Leu
 455 460 465
 Cys His Asn Val Thr Pro Val Tyr Glu Ser Asn Gly Val Thr Asp
 470 475 480
 Gln Ala Glu Ala Glu Lys Gln Tyr Glu Asp Ser Cys Arg Val Tyr
 485 490 495
 Gln Ala Ser Ser Pro Asp Glu Val Ala Leu Val Gln Trp Thr Glu
 500 505 510
 Ser Val Gly Leu Thr Leu Val Gly Arg Asp Gln Ser Ser Met Gln
 515 520 525
 Leu Arg Thr Pro Gly Asp Gln Ile Leu Asn Phe Thr Ile Leu Gln
 530 535 540
 Ile Phe Pro Phe Thr Tyr Glu Ser Lys Arg Met Gly Ile Ile Val
 545 550 555
 Arg Asp Glu Ser Thr Gly Glu Ile Thr Phe Tyr Met Lys Gly Ala
 560 565 570
 Asp Val Val Met Ala Gly Ile Val Gln Tyr Asn Asp Trp Leu Glu
 575 580 585
 Glu Glu Cys Gly Asn Met Ala Arg Glu Gly Leu Arg Val Leu Val
 590 595 600
 Val Ala Lys Lys Ser Leu Ala Glu Glu Gln Tyr Gln Asp Phe Glu

WO 02/40541

PCT/US01/46055

605
 Ala Arg Tyr Val Gln Ala Lys Leu Ser Val His Asp Arg Ser Leu
 620
 Lys Val Ala Thr Val Ile Glu Ser Leu Glu Met Glu Met Glu Leu
 635
 Leu Cys Leu Thr Gly Val Glu Asp Gln Leu Gln Ala Asp Val Arg
 650
 Pro Thr Leu Glu Thr Leu Arg Asn Ala Gly Ile Lys Val Trp Met
 665
 Leu Thr Gly Asp Lys Leu Glu Thr Ala Thr Cys Thr Ala Lys Asn
 680
 Ala His Leu Val Thr Arg Asn Gln Asp Ile His Val Phe Arg Leu
 695
 Val Thr Asn Arg Gly Glu Ala His Leu Glu Leu Asn Ala Phe Arg
 710
 Arg Lys His Asp Cys Ala Leu Val Ile Ser Gly Asp Ser Leu Glu
 725
 Val Cys Leu Lys Tyr Tyr Glu Tyr Glu Phe Met Glu Leu Ala Cys
 740
 Gln Cys Pro Ala Val Val Cys Cys Arg Cys Ala Pro Thr Gln Lys
 755
 Ala Gln Ile Val Arg Leu Leu Gln Glu Arg Thr Gly Lys Leu Thr
 770
 Cys Ala Val Gly Asp Gly Gly Asn Asp Val Ser Met Ile Gln Glu
 785
 Ser Asp Cys Gly Val Gly Val Glu Gly Lys Glu Gly Lys Gln Ala
 800
 Ser Leu Ala Ala Asp Phe Ser Ile Thr Gln Phe Lys His Leu Gly
 815
 Arg Leu Leu Met Val His Gly Arg Asn Ser Tyr Lys Arg Ser Ala
 830
 Ala Leu Ser Gln Phe Val Ile His Arg Ser Leu Cys Ile Ser Thr
 845
 Met Gln Ala Val Phe Ser Ser Val Phe Tyr Phe Ala Ser Val Pro
 860
 Leu Tyr Gln Gly Phe Leu Ile Ile Gly Tyr Ser Thr Ile Tyr Thr
 875
 Met Phe Pro Val Phe Ser Leu Val Leu Asp Lys Asp Val Lys Ser
 890
 Glu Val Ala Met Leu Tyr Pro Glu Leu Tyr Lys Asp Leu Leu Lys
 905
 Gly Arg Pro Leu Ser Tyr Lys Thr Phe Leu Ile Trp Val Leu Ile
 920
 Ser Ile Tyr Gln Gly Ser Thr Ile Met Tyr Gly Ala Leu Leu Leu
 935
 Phe Glu Ser Glu Phe Val His Ile Val Ala Ile Ser Phe Thr Ser
 950
 Leu Ile Leu Thr Glu Leu Leu Met Val Ala Leu Thr Ile Gln Thr
 965
 Trp His Trp Leu Met Thr Val Ala Glu Leu Leu Ser Leu Ala Cys
 980
 Tyr Ile Ala Ser Leu Val Phe Leu His Glu Phe Ile Asp Val Tyr
 995
 Phe Ile Ala Thr Leu Ser Phe Leu Trp Lys Val Ser Val Ile Thr
 1010
 Leu Val Ser Cys Leu Pro Leu Tyr Val Leu Lys Tyr Leu Arg Arg
 1025
 Arg Phe Ser Pro Pro Ser Tyr Ser Lys Leu Thr Ser
 1040

<210> 4
 <211> 671
 <212> PRT

WO 02/06541

PCT/US01/46055

<213> Homo sapiens

<220>

<221> misc_feature

<223> Truncyte ID No: 7472732CD1

<400> 4

```

Met Thr Gly Ala Lys Arg Lys Lys Lys Ser Met Leu Trp Ser Lys
1      5      10      15
Met His Thr Pro Gln Cys Glu Asp Ile Ile Gln Trp Cys Arg Arg
20     25     30
Arg Leu Pro Ile Leu Asp Trp Ala Pro His Tyr Asn Leu Lys Glu
35     40     45
Asn Leu Leu Pro Asp Thr Val Ser Gly Ile Met Leu Ala Val Gln
50     55     60
Gln Val Thr Gln Gly Leu Ala Phe Ala Val Leu Ser Ser Val His
65     70     75
Pro Val Phe Gly Leu Tyr Gly Ser Leu Phe Pro Ala Ile Ile Tyr
80     85     90
Ala Ile Phe Gly Met Gly His His Val Ala Thr Gly Thr Phe Ala
95     100    105
Leu Thr Ser Leu Ile Ser Ala Asn Ala Val Glu Arg Ile Val Pro
110    115    120
Gln Asn Met Gln Asn Leu Thr Thr Gln Ser Asn Thr Ser Val Leu
125    130    135
Gly Leu Ser Asp Phe Glu Met Gln Arg Ile His Val Ala Ala Ala
140    145    150
Val Ser Phe Leu Gly Gly Val Ile Gln Val Ala Met Phe Val Leu
155    160    165
Gln Leu Gly Ser Ala Thr Phe Val Val Thr Glu Pro Val Ile Ser
170    175    180
Ala Met Thr Thr Gly Ala Ala Thr His Val Val Thr Ser Gln Val
185    190    195
Lys Tyr Leu Leu Gly Met Lys Met Pro Tyr Ile Ser Gly Pro Leu
200    205    210
Gly Phe Phe Tyr Ile Tyr Ala Tyr Val Phe Glu Asn Ile Lys Ser
215    220    225
Val Arg Leu Glu Ala Leu Leu Leu Ser Leu Leu Ser Ile Val Val
230    235    240
Leu Val Leu Val Lys Glu Leu Asn Glu Gln Phe Lys Arg Lys Ile
245    250    255
Lys Val Val Leu Pro Val Asp Leu Val Leu Ala Pro Asn Thr Ser
260    265    270
Pro Leu His His His Tyr Asp Cys Leu Phe Ala Asn Phe Leu Glu
275    280    285
Pro Pro Trp Glu Asp Gly Leu Pro Glu Gly Ala Phe Asn Gln Ala
290    295    300
Glu Gly His Leu Arg Arg Asn Ile Ile Pro Ser Pro Arg Ala Pro
305    310    315
Pro Met Asn Ile Leu Ser Ala Val Ile Thr Glu Ala Phe Gly Val
320    325    330
Ala Leu Val Gly Tyr Val Ala Ser Leu Ala Leu Ala Gln Gly Ser
335    340    345
Ala Lys Lys Phe Lys Tyr Ser Ile Asp Asp Asn Gln Glu Phe Leu
350    355    360
Ala His Gly Leu Ser Asn Ile Val Ser Ser Phe Phe Phe Cys Ile
365    370    375
Pro Ser Ala Ala Ala Met Gly Arg Thr Ala Gly Leu Tyr Ser Thr
380    385    390
Gly Ala Lys Thr Gln Val Ala Cys Leu Ile Ser Cys Ile Phe Val
395    400    405
Leu Ile Val Ile Tyr Ala Ile Gly Pro Leu Leu Tyr Trp Leu Pro
410    415    420

```

WO 02/06541

PCT/US01/46055

```

Met Cys Val Leu Ala Ser Ile Ile Val Val Gly Leu Lys Gly Met
425 430 435
Leu Ile Gln Phe Arg Asp Leu Lys Lys Tyr Trp Asn Val Asp Lys
440 445 450
Ile Asp Trp Gly Ile Trp Val Ser Thr Tyr Val Phe Thr Ile Cys
455 460 465
Phe Ala Ala Asn Val Gly Leu Leu Phe Gly Val Val Cys Thr Ile
470 475 480
Ala Ile Val Ile Gly Arg Phe Pro Arg Ala Met Thr Val Ser Ile
485 490 495
Lys Asn Met Lys Glu Met Glu Phe Lys Val Lys Thr Glu Met Asp
500 505 510
Ser Glu Thr Leu Gln Gln Val Lys Ile Ile Ser Ile Asn Asn Pro
515 520 525
Leu Val Phe Leu Asn Ala Lys Lys Phe Tyr Thr Asp Leu Met Asn
530 535 540
Met Ile Gln Lys Glu Asn Ala Cys Asn Gln Pro Leu Asp Asp Ile
545 550 555
Ser Lys Cys Glu Gln Asn Thr Leu Leu Asn Ser Leu Ser Asn Gly
560 565 570
Asn Cys Asn Glu Glu Ala Ser Gln Ser Cys Pro Asn Glu Lys Cys
575 580 585
Tyr Leu Ile Leu Asp Cys Ser Gly Phe Thr Phe Phe Asp Tyr Ser
590 595 600
Gly Val Ser Met Leu Val Glu Val Tyr Met Asp Cys Lys Gly Arg
605 610 615
Ser Val Asp Val Leu Leu Ala His Cys Thr Ala Ser Leu Ile Lys
620 625 630
Ala Met Thr Tyr Tyr Gly Asn Leu Asp Ser Glu Lys Pro Ile Phe
635 640 645
Phe Glu Ser Val Ser Ala Ala Ile Ser His Ile His Ser Asn Lys
650 655 660
Asn Leu Ser Lys Leu Ser Asp His Ser Glu Val
665 670

```

<210> 5

<211> 671

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7476938CD1

<400> 5

```

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

```

WO 02/06541

PCT/US01/46055

Gln Thr Gly Glu Gln Ala Gln Gln Glu Glu Gln Ser Ser Gly Met
 140 145 150
 Thr Ile Phe Phe Ser Leu Leu Val Leu Ala Ile Cys Ile Ile Leu
 155 160 165
 Val His Leu Leu Ile Arg Tyr Arg Leu His Phe Leu Pro Glu Ser
 170 175 180
 Val Ala Val Val Ser Leu Gly Ile Leu Met Gly Ala Val Ile Lys
 185 190 195
 Ile Ile Glu Phe Lys Lys Leu Ala Asn Trp Lys Glu Glu Glu Met
 200 205 210
 Phe Arg Pro Asn Met Phe Phe Leu Leu Leu Leu Pro Pro Ile Ile
 215 220 225
 Phe Glu Ser Gly Tyr Ser Leu His Lys Gly Asn Phe Phe Gln Asn
 230 235 240
 Ile Gly Ser Ile Thr Leu Phe Ala Val Phe Gly Thr Ala Ile Ser
 245 250 255
 Ala Phe Val Val Gly Gly Ile Tyr Phe Leu Gly Gln Ala Asp
 260 265 270
 Val Ile Ser Lys Leu Asn Met Thr Asp Ser Phe Ala Phe Gly Ser
 275 280 285
 Leu Ile Ser Ala Val Asp Pro Val Ala Thr Ile Ala Ile Phe Asn
 290 295 300
 Ala Leu His Val Asp Pro Val Leu Asn Met Leu Val Phe Gly Glu
 305 310 315
 Ser Ile Leu Asn Asp Ala Val Ser Ile Val Leu Thr Asn Thr Ala
 320 325 330
 Glu Gly Leu Thr Arg Lys Asn Met Ser Asp Val Ser Gly Trp Gln
 335 340 345
 Thr Phe Leu Gln Ala Leu Asp Tyr Phe Leu Lys Met Phe Phe Gly
 350 355 360
 Ser Ala Ala Leu Gly Thr Leu Thr Gly Leu Ile Ser Ala Leu Val
 365 370 375
 Leu Lys His Ile Asp Leu Arg Lys Thr Pro Ser Leu Glu Phe Gly
 380 385 390
 Met Met Ile Ile Phe Ala Tyr Leu Pro Tyr Gly Leu Ala Glu Gly
 395 400 405
 Ile Ser Leu Ser Gly Ile Met Ala Ile Leu Phe Ser Gly Ile Val
 410 415 420
 Met Ser His Tyr Thr His His Asn Leu Ser Pro Val Thr Gln Ile
 425 430 435
 Leu Met Gln Gln Thr Leu Arg Thr Val Ala Phe Leu Cys Glu Thr
 440 445 450
 Cys Val Phe Ala Phe Leu Gly Leu Ser Ile Phe Ser Phe Pro His
 455 460 465
 Lys Phe Glu Ile Ser Phe Val Ile Trp Cys Ile Val Leu Val Leu
 470 475 480
 Phe Gly Arg Ala Val Asn Ile Phe Pro Leu Ser Tyr Leu Leu Asn
 485 490 495
 Phe Phe Arg Asp His Lys Ile Thr Pro Lys Met Met Phe Ile Met
 500 505 510
 Trp Phe Ser Gly Leu Arg Gly Ala Ile Pro Tyr Ala Leu Ser Leu
 515 520 525
 His Leu Asp Leu Glu Pro Met Glu Lys Arg Gln Leu Ile Gly Thr
 530 535 540
 Thr Thr Ile Val Ile Val Leu Phe Thr Ile Leu Leu Leu Gly Gly
 545 550 555
 Ser Thr Met Pro Leu Ile Arg Leu Met Asp Ile Glu Asp Ala Lys
 560 565 570
 Ala His Arg Arg Asn Lys Lys Asp Val Asn Leu Ser Lys Thr Glu
 575 580 585
 Lys Met Gly Asn Thr Val Glu Ser Glu His Leu Ser Glu Leu Thr
 590 595 600
 Glu Glu Glu Tyr Glu Ala His Tyr Ile Arg Arg Gln Asp Leu Lys

WO 02/00541

PCT/US01/46055

```

      605      610      615
Gly Phe Val Trp Leu Asp Ala Lys Tyr Leu Asn Pro Phe Phe Thr
      620      625      630
Arg Arg Leu Thr Gln Glu Asp Leu His His Gly Arg Ile Gln Met
      635      640      645
Lys Thr Leu Thr Asn Lys Trp Tyr Glu Glu Val Arg Gln Gly Pro
      650      655      660
Ser Gly Ser Glu Asp Asp Glu Gln Glu Leu Leu
      665      670

```

```

<210> 6
<211> 315
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 8128531CD1

```

```

<400> 6
Met Thr His Gln Asp Leu Ser Ile Thr Ala Lys Leu Ile Asn Gly
 1      5      10      15
Gly Val Ala Gly Leu Val Gly Val Thr Cys Val Phe Pro Ile Asp
 20      25      30
Leu Ala Lys Thr Arg Leu Gln Asn Gln His Gly Lys Ala Met Tyr
 35      40      45
Lys Gly Met Ile Asp Cys Leu Met Lys Thr Ala Arg Ala Glu Gly
 50      55      60
Phe Phe Gly Met Tyr Arg Gly Ala Ala Val Asn Leu Thr Leu Val
 65      70      75
Thr Pro Glu Lys Ala Ile Lys Leu Ala Ala Asn Asp Phe Phe Arg
 80      85      90
Arg Leu Leu Met Glu Asp Gly Met Gln Arg Asn Leu Lys Met Glu
 95      100      105
Met Leu Ala Gly Cys Gly Ala Gly Met Cys Gln Val Val Val Thr
 110      115      120
Cys Pro Met Glu Met Leu Lys Ile Gln Leu Gln Asp Ala Gly Arg
 125      130      135
Leu Ala Val His His Gln Gly Ser Ala Ser Ala Pro Ser Thr Ser
 140      145      150
Arg Ser Tyr Thr Thr Gly Ser Ala Ser Thr His Arg Arg Pro Ser
 155      160      165
Ala Thr Leu Ile Ala Trp Glu Leu Leu Arg Thr Gln Gly Leu Ala
 170      175      180
Gly Leu Tyr Arg Gly Leu Gly Ala Thr Leu Leu Arg Asp Ile Pro
 185      190      195
Phe Ser Ile Ile Tyr Phe Pro Leu Phe Ala Asn Leu Asn Asn Leu
 200      205      210
Gly Phe Asn Glu Leu Ala Gly Lys Ala Ser Phe Ala His Ser Phe
 215      220      225
Val Ser Gly Cys Val Ala Gly Ser Ile Ala Ala Val Ala Val Thr
 230      235      240
Pro Leu Asp Val Leu Lys Thr Arg Ile Gln Thr Leu Lys Lys Gly
 245      250      255
Leu Gly Glu Asp Met Tyr Ser Gly Ile Thr Asp Cys Ala Arg Lys
 260      265      270
Leu Trp Ile Gln Glu Gly Pro Ser Ala Phe Met Lys Gly Ala Gly
 275      280      285
Cys Arg Ala Leu Val Ile Ala Pro Leu Phe Gly Ile Ala Gln Gly
 290      295      300
Val Tyr Phe Ile Gly Ile Gly Glu Arg Ile Leu Lys Cys Phe Asp
 305      310      315

```

WO 02/00541

PCT/US01/46055

<210> 7
 <211> 445
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7476757CD1

<400> 7
 Met Pro Trp Val Leu Gly Cys Thr Pro Phe Ile Ala Leu Ala Tyr
 1 5 10 15
 Phe Phe Leu Trp Phe Leu Pro Pro Phe Thr Ser Leu Arg Gly Leu
 20 25 30
 Trp Tyr Thr Thr Phe Tyr Cys Leu Phe Gln Ala Leu Ala Thr Phe
 35 40 45
 Phe Gln Val Pro Tyr Thr Ala Leu Thr Met Leu Leu Thr Pro Cys
 50 55 60
 Pro Arg Glu Arg Asp Ser Ala Thr Ala Ile Pro Asp Asp Cys Gly
 65 70 75
 Asp Gly Gly Asn Thr Asp Gly Gly His Cys Pro Arg Ala His Arg
 80 85 90
 Val Arg Arg Pro Gln Thr Pro Gln Val Arg Gly His Cys Asp Pro
 95 100 105
 Gly Ala Ser His Cys Leu Pro Glu Cys Ser His Leu Tyr Cys Ile
 110 115 120
 Ala Ala Ala Val Val Val Val Thr Tyr Pro Val Cys Ile Ser Leu
 125 130 135
 Leu Cys Leu Gly Val Lys Glu Arg Pro Gly Phe Ala Phe Glu Leu
 140 145 150
 Cys Glu Ala Lys Val Thr Arg Phe Cys Val Ala Asp Pro Ser Ala
 155 160 165
 Pro Ala Ser Gly Pro Gly Leu Ser Phe Leu Ala Gly Leu Ser Leu
 170 175 180
 Thr Thr Arg His Pro Pro Tyr Leu Lys Leu Val Ile Ser Phe Leu
 185 190 195
 Phe Ile Ser Ala Ala Val Gln Val Glu Gln Ser Tyr Leu Val Leu
 200 205 210
 Phe Cys Thr His Ala Ser Gln Leu His Asp His Val Gln Gly Leu
 215 220 225
 Val Ser Ala Val Leu Ser Thr Pro Leu Trp Glu Trp Val Leu Gln
 230 235 240
 Arg Phe Gly Lys Lys Thr Ser Ala Phe Gly Ile Phe Ala Met Val
 245 250 255
 Pro Phe Ala Ile Leu Leu Ala Ala Val Pro Thr Ala Pro Val Ala
 260 265 270
 Tyr Val Val Ala Phe Val Ser Gly Val Ser Ile Ala Val Ser Leu
 275 280 285
 Leu Leu Pro Trp Ser Met Leu Pro Asp Val Val Asp Asp Phe Gln
 290 295 300
 Leu Gln His Arg His Gly Pro Gly Leu Glu Thr Ile Phe Tyr Ser
 305 310 315
 Ser Tyr Val Phe Phe Thr Lys Leu Ser Gly Ala Cys Ala Leu Gly
 320 325 330
 Ile Ser Thr Leu Ser Leu Glu Phe Ser Gly Tyr Lys Ala Gly Val
 335 340 345
 Cys Lys Gln Ala Glu Glu Val Val Val Thr Leu Lys Val Leu Ile
 350 355 360
 Gly Ala Val Pro Thr Cys Met Ile Leu Ala Gly Leu Cys Ile Leu
 365 370 375
 Met Val Gly Ser Thr Pro Lys Thr Pro Ser Arg Asp Ala Ser Ser
 380 385 390
 Arg Leu Ser Leu Arg Arg Arg Ala Gln Ala Pro Asn Val His Thr

WO 02/40541

PCT/US01/46055

```

395          400          405
Ser Lys Val His Glu His Ala His Ile Met Gln Ala His Ala Gly
410          415          420
Gln Ala Val Gly Gly Leu Val Ile Ser His Ser Leu Leu Arg Val
425          430          435
Thr Ala Ser Gly Ser Ala Ala Glu Arg Tyr
440          445

```

```

<210> 8
<211> 410
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 266243CD1

```

```

<400> 8
Met Ala Ala Ala Val Gly Ala Gly His Gly Ala Gly Gly Pro
1      5      10      15
Gly Ala Ala Ser Ser Ser Gly Gly Ala Arg Glu Gly Ala Arg Val
20     25     30
Ala Ala Leu Cys Leu Leu Trp Tyr Ala Leu Ser Ala Gly Gly Asn
35     40     45
Val Val Asn Lys Val Ile Leu Ser Ala Phe Pro Phe Pro Val Thr
50     55     60
Val Ser Leu Cys His Ile Leu Ala Leu Cys Ala Gly Leu Pro Pro
65     70     75
Leu Leu Arg Ala Trp Arg Val Pro Pro Ala Pro Pro Val Ser Gly
80     85     90
Pro Gly Pro Ser Pro His Pro Ser Ser Gly Pro Leu Leu Pro Pro
95     100    105
Arg Phe Tyr Pro Arg Tyr Val Leu Pro Leu Ala Phe Gly Lys Tyr
110    115    120
Phe Ala Ser Val Ser Ala His Val Ser Ile Trp Lys Val Pro Val
125    130    135
Ser Tyr Ala His Thr Val Lys Ala Thr Met Pro Ile Trp Val Val
140    145    150
Leu Leu Ser Arg Ile Ile Met Lys Glu Lys Gln Ser Thr Lys Val
155    160    165
Tyr Leu Ser Leu Ile Pro Ile Ile Ser Gly Val Leu Leu Ala Thr
170    175    180
Val Thr Glu Leu Ser Phe Asp Met Trp Gly Leu Val Ser Ala Leu
185    190    195
Ala Ala Thr Leu Cys Phe Ser Leu Gln Asn Ile Phe Ser Lys Lys
200    205    210
Val Leu Arg Asp Ser Arg Ile His His Leu Arg Leu Leu Asn Ile
215    220    225
Leu Gly Cys His Ala Val Phe Phe Met Ile Pro Thr Trp Val Leu
230    235    240
Val Asp Leu Ser Ala Phe Leu Val Ser Ser Asp Leu Thr Tyr Val
245    250    255
Tyr Gln Trp Pro Trp Thr Leu Leu Leu Leu Ala Val Ser Gly Phe
260    265    270
Cys Asn Phe Ala Gln Asn Val Ile Ala Phe Ser Ile Leu Asn Leu
275    280    285
Val Ser Pro Leu Ser Tyr Ser Val Ala Asn Ala Thr Lys Arg Ile
290    295    300
Met Val Ile Thr Val Ser Leu Ile Met Leu Arg Asn Pro Val Thr
305    310    315
Ser Thr Asn Val Leu Gly Met Met Thr Ala Ile Leu Gly Val Phe
320    325    330
Leu Tyr Asn Lys Thr Lys Tyr Asp Ala Asn Gln Gln Ala Arg Lys

```


WO 02/40541

PCT/US01/46055

```

335      340      345
His Leu Leu Pro Val Thr Thr Ala Asp Leu Ser Ser Lys Glu Arg
350      355      360
His Arg Ser Pro Leu Glu Lys Pro His Asn Gly Leu Leu Phe Pro
365      370      375
Gln His Gly Asp Tyr Gln Tyr Gly Arg Asn Asn Ile Leu Thr Asp
380      385      390
His Phe Gln Tyr Ser Arg Gln Ser Tyr Pro Asn Ser Tyr Ser Leu
395      400      405
Asn Arg Tyr Asp Val
410

```

```

<210> 9
<211> 374
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 6585710CD1

```

```

<400> 9
Met Val His Tyr Phe Thr Ala Ile Gly Tyr Pro Cys Pro Arg Tyr
1      5      10      15
Ser Asn Pro Ala Asp Phe Tyr Val Asp Leu Thr Ser Ile Asp Arg
20      25      30
Arg Ser Arg Glu Gln Glu Leu Ala Thr Arg Glu Lys Ala Gln Ser
35      40      45
Leu Ala Ala Leu Phe Leu Glu Lys Val Arg Asp Leu Asp Asp Phe
50      55      60
Leu Trp Lys Ala Glu Thr Lys Asp Leu Asp Glu Asp Thr Cys Val
65      70      75
Glu Ser Ser Val Thr Pro Leu Asp Thr Asn Cys Leu Pro Ser Pro
80      85      90
Thr Lys Met Pro Gly Ala Val Gln Gln Phe Thr Thr Leu Ile Arg
95      100      105
Arg Gln Ile Ser Asn Asp Phe Arg Asp Leu Pro Thr Leu Leu Ile
110      115      120
His Gly Ala Glu Ala Cys Leu Met Ser Met Thr Ile Gly Phe Leu
125      130      135
Tyr Phe Gly His Gly Ser Ile Gln Leu Ser Phe Met Asp Thr Ala
140      145      150
Ala Leu Leu Phe Met Ile Gly Ala Leu Ile Pro Phe Asn Val Ile
155      160      165
Leu Asp Val Ile Ser Lys Cys Tyr Ser Glu Arg Ala Met Leu Tyr
170      175      180
Tyr Glu Leu Glu Asp Gly Leu Tyr Thr Thr Gly Pro Tyr Phe Phe
185      190      195
Ala Lys Ile Leu Gly Glu Leu Pro Glu His Cys Ala Tyr Ile Ile
200      205      210
Ile Tyr Gly Met Pro Thr Tyr Trp Leu Ala Asn Leu Arg Pro Gly
215      220      225
Leu Gln Pro Phe Leu Leu His Phe Leu Leu Val Trp Leu Val Val
230      235      240
Phe Cys Cys Arg Ile Met Ala Leu Ala Ala Ala Ala Leu Leu Pro
245      250      255
Thr Phe His Met Ala Ser Phe Phe Ser Asn Ala Leu Tyr Asn Ser
260      265      270
Phe Tyr Leu Ala Gly Gly Phe Met Ile Asn Leu Ser Ser Leu Trp
275      280      285
Thr Val Pro Ala Trp Ile Ser Lys Val Ser Phe Leu Arg Trp Cys
290      295      300
Phe Glu Gly Leu Met Lys Ile Gln Phe Ser Arg Arg Thr Tyr Lys

```

WO 02/40541

PCT/US01/46055

```

305           310           315
Met Pro Leu Gly Asn Leu Thr Ile Ala Val Ser Gly Asp Lys Ile
320           325           330
Leu Ser Ala Met Glu Leu Asp Ser Tyr Pro Leu Tyr Ala Ile Tyr
335           340           345
Leu Ile Val Ile Gly Leu Ser Gly Gly Phe Met Val Leu Tyr Tyr
350           355           360
Val Ser Leu Arg Phe Ile Lys Gln Lys Pro Ser Gln Asp Trp
365           370

<210> 10
<211> 443
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7483599CDL

<400> 10
Met Asp Lys Phe Leu Asp Thr Tyr Asn Leu Pro Arg Leu Asn Gln
1      5      10      15
Glu Glu Ile Gln Asn Leu Lys Arg Pro Ile Thr Ser Asn Glu Ile
20     25     30
Lys Ala Ile Ile Lys Ser Leu Gln Met Ser Leu Leu Gly Arg Asp
35     40     45
Tyr Asn Ser Glu Leu Asn Ser Leu Asp Asn Gly Pro Gln Ser Pro
50     55     60
Ser Glu Ser Ser Ser Ser Ile Thr Ser Glu Asn Val His Pro Ala
65     70     75
Gly Glu Ala Gly Leu Ser Met Met Gln Thr Leu Ile His Leu Leu
80     85     90
Lys Cys Asn Ile Gly Thr Gly Leu Leu Gly Leu Pro Leu Ala Ile
95     100    105
Lys Asn Ala Gly Leu Leu Val Gly Pro Val Ser Leu Leu Ala Ile
110    115    120
Gly Val Leu Thr Val His Cys Met Val Ile Leu Leu Asn Cys Ala
125    130    135
Gln His Leu Ser Gln Pro Arg Leu Gln Lys Thr Phe Val Asn Tyr
140    145    150
Gly Glu Ala Thr Met Tyr Gly Leu Glu Thr Cys Pro Asn Thr Trp
155    160    165
Leu Arg Ala His Ala Val Trp Gly Arg Tyr Thr Val Ser Phe Leu
170    175    180
Leu Val Ile Thr Gln Leu Gly Phe Cys Ser Val Tyr Phe Met Phe
185    190    195
Met Ala Asp Asn Leu Gln Gln Met Val Glu Lys Ala His Val Thr
200    205    210
Ser Asn Ile Cys Gln Pro Arg Glu Ile Leu Thr Leu Thr Pro Ile
215    220    225
Leu Asp Ile Arg Phe Tyr Met Leu Ile Ile Leu Pro Phe Leu Ile
230    235    240
Leu Leu Val Phe Ile Gln Asn Leu Lys Val Leu Ser Val Phe Ser
245    250    255
Thr Leu Ala Asn Ile Thr Thr Leu Gly Ser Met Ala Leu Ile Phe
260    265    270
Glu Tyr Ile Met Glu Gly Ile Pro Tyr Pro Ser Asn Leu Pro Leu
275    280    285
Met Ala Asn Trp Lys Thr Phe Leu Leu Phe Phe Gly Thr Ala Ile
290    295    300
Phe Thr Phe Glu Gly Val Gly Met Val Leu Pro Leu Lys Asn Gln
305    310    315
Met Lys His Pro Gln Gln Phe Ser Phe Val Leu Tyr Leu Gly Met

```

WO 02/06541

PCT/US01/46055

```

320
Ser Ile Val Ile Ile Leu Tyr Ile Leu Leu Gly Thr Leu Gly Tyr
335
Met Lys Phe Gly Ser Asp Thr Gln Ala Ser Ile Thr Leu Asn Leu
350
Pro Asn Cys Trp Tyr Val Leu Pro Thr Ser Gly Glu Ile Gly Arg
365
Asp Thr Gly Thr Val Leu Val Val Ile Ala Glu Ser Thr Ala Lys
380
Leu Ser His Glu Ala Gly Asn Pro Ser Leu Glu Val Thr Tyr Val
395
Ser Pro Ala His Thr Ala Ser Val Lys Ala Ser His Met Ala Ala
410
Pro His Ser Lys Gly Ala Gly Lys Cys Asn Ser Ala Met Cys Leu
425
Glu Val Phe Gly Glu Gln His Lys
440

```

<210> 11

<211> 321

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2507246CD1

<400> 11

```

Met Ala Thr Gly Gly Gln Gln Lys Glu Asn Thr Leu Leu His Leu
1 5 10 15
Phe Ala Gly Gly Cys Gly Gly Thr Val Gly Ala Ile Phe Thr Cys
20 25 30
Pro Leu Glu Val Ile Lys Thr Arg Leu Gln Ser Ser Arg Leu Ala
35 40 45
Leu Arg Thr Val Tyr Tyr Pro Gln Val His Leu Gly Thr Ile Ser
50 55 60
Gly Ala Gly Met Val Arg Pro Thr Ser Val Thr Pro Gly Leu Phe
65 70 75
Gln Val Leu Lys Ser Ile Leu Glu Lys Glu Gly Pro Lys Ser Leu
80 85 90
Phe Arg Gly Leu Gly Pro Asn Leu Val Gly Val Ala Pro Ser Arg
95 100 105
Ala Val Tyr Phe Ala Cys Tyr Ser Lys Ala Lys Glu Gln Phe Asn
110 115 120
Gly Ile Phe Val Pro Asn Ser Asn Ile Val His Ile Phe Ser Ala
125 130 135
Gly Ser Ala Ala Phe Ile Thr Asn Ser Leu Met Asn Pro Ile Trp
140 145 150
Met Val Lys Thr Arg Met Gln Leu Glu Gln Lys Val Arg Gly Ser
155 160 165
Lys Gln Met Asn Thr Leu Gln Cys Ala Arg Tyr Val Tyr Gln Thr
170 175 180
Glu Gly Ile Arg Gly Phe Tyr Arg Gly Leu Thr Ala Ser Tyr Ala
185 190 195
Gly Ile Ser Glu Thr Ile Ile Cys Phe Ala Ile Tyr Glu Ser Leu
200 205 210
Lys Lys Tyr Leu Lys Glu Ala Pro Leu Ala Ser Ser Ala Asn Gly
215 220 225
Thr Glu Lys Asn Ser Thr Ser Phe Phe Gly Leu Met Ala Ala Ala
230 235 240
Ala Leu Ser Lys Gly Cys Ala Ser Cys Ile Ala Tyr Pro His Glu
245 250 255
Val Ile Arg Thr Arg Leu Arg Glu Glu Gly Thr Lys Tyr Lys Ser

```

WO 02/40541

PCT/US01/46055

260
 Phe Val Gln Thr Ala Arg Leu Val Phe Arg Glu Glu Gly Tyr Leu
 275
 Ala Phe Tyr Arg Gly Leu Phe Ala Gln Leu Ile Arg Gln Ile Pro
 290
 Asn Thr Ala Ile Val Leu Ser Thr Tyr Glu Leu Ile Val Tyr Leu
 305
 Leu Glu Asp Arg Thr Gln
 320

<210> 12
 <211> 487
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3033505CD1

<400> 12
 Met Met His Phe Lys Ser Gly Leu Glu Leu Thr Glu Leu Gln Asn
 1 5 10 15
 Met Thr Val Pro Glu Asp Asp Asn Ile Ser Asn Asp Ser Asn Asp
 20 25 30
 Phe Thr Glu Val Glu Asn Gly Gln Ile Asn Ser Lys Phe Ile Ser
 35 40 45
 Asp Arg Glu Ser Arg Arg Ser Leu Thr Asn Ser His Leu Glu Lys
 50 55 60
 Lys Lys Cys Asp Glu Tyr Ile Pro Gly Thr Thr Ser Leu Gly Met
 65 70 75
 Ser Val Phe Asn Leu Ser Asn Ala Ile Met Gly Ser Gly Ile Leu
 80 85 90
 Gly Leu Ala Phe Ala Leu Ala Asn Thr Gly Ile Leu Leu Phe Leu
 95 100 105
 Val Leu Leu Thr Ser Val Thr Leu Leu Ser Ile Tyr Ser Ile Asn
 110 115 120
 Leu Leu Leu Ile Cys Ser Lys Glu Thr Gly Cys Met Val Tyr Glu
 125 130 135
 Lys Leu Gly Glu Gln Val Phe Gly Thr Thr Gly Lys Phe Val Ile
 140 145 150
 Phe Gly Ala Thr Ser Leu Gln Asn Thr Gly Ala Met Leu Ser Tyr
 155 160 165
 Leu Phe Ile Val Lys Asn Glu Leu Pro Ser Ala Ile Lys Phe Leu
 170 175 180
 Met Gly Lys Glu Glu Thr Phe Ser Ala Trp Tyr Val Asp Gly Arg
 185 190 195
 Val Leu Val Val Ile Val Thr Phe Gly Ile Ile Leu Pro Leu Cys
 200 205 210
 Leu Leu Lys Asn Leu Gly Tyr Leu Gly Tyr Thr Ser Gly Phe Ser
 215 220 225
 Leu Ser Cys Met Val Phe Phe Leu Ile Val Val Ile Tyr Lys Lys
 230 235 240
 Phe Gln Ile Pro Cys Ile Val Pro Glu Leu Asn Ser Thr Ile Ser
 245 250 255
 Ala Asn Ser Thr Asn Ala Asp Thr Cys Thr Pro Lys Tyr Val Thr
 260 265 270
 Phe Asn Ser Lys Thr Val Tyr Ala Leu Pro Thr Ile Ala Phe Ala
 275 280 285
 Phe Val Cys His Pro Ser Val Leu Pro Ile Tyr Ser Glu Leu Lys
 290 295 300
 Asp Arg Ser Gln Lys Lys Met Gln Met Val Ser Asn Ile Ser Phe
 305 310 315
 Phe Ala Met Phe Val Met Tyr Phe Leu Thr Ala Ile Phe Gly Tyr

WO 02/40541

PCT/US01/46055

```

320      325      330
Leu Thr Phe Tyr Asp Asn Val Gln Ser Asp Leu Leu His Lys Tyr
335      340      345
Gln Ser Iys Asp Asp Ile Leu Ile Leu Thr Val Arg Leu Ala Val
350      355      360
Ile Val Ala Val Ile Leu Thr Val Pro Val Leu Phe Phe Thr Val
365      370      375
Arg Ser Ser Leu Phe Glu Leu Ala Lys Lys Thr Lys Phe Asn Leu
380      385      390
Cys Arg His Thr Val Val Thr Cys Ile Leu Leu Val Val Ile Asn
395      400      405
Leu Leu Val Ile Phe Ile Pro Ser Met Lys Asp Ile Phe Gly Val
410      415      420
Val Gly Val Thr Ser Ala Asn Met Leu Ile Phe Ile Leu Pro Ser
425      430      435
Ser Leu Tyr Leu Lys Ile Thr Asp Gln Asp Gly Asp Lys Gly Thr
440      445      450
Gln Arg Ile Trp Ala Ala Leu Phe Leu Gly Leu Gly Val Leu Phe
455      460      465
Ser Leu Val Ser Ile Pro Leu Val Ile Tyr Asp Trp Ala Cys Ser
470      475      480
Ser Ser Ser Asp Glu Gly His
485

```

```

<210> 13
<211> 509
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4027693CD1

```

```

<400> 13
Met Glu Leu Lys Lys Ser Pro Asp Gly Gly Trp Gly Trp Val Ile
1      5      10      15
Val Phe Val Ser Phe Leu Thr Gln Phe Leu Cys Tyr Gly Ser Pro
20      25      30
Leu Ala Val Gly Val Leu Tyr Ile Glu Trp Leu Asp Ala Phe Gly
35      40      45
Glu Gly Lys Gly Lys Thr Ala Trp Val Gly Ser Leu Ala Ser Gly
50      55      60
Val Gly Leu Leu Ala Ser Pro Val Cys Ser Leu Cys Val Ser Ser
65      70      75
Phe Gly Ala Arg Pro Val Thr Ile Phe Ser Gly Phe Met Val Ala
80      85      90
Gly Gly Leu Met Leu Ser Ser Phe Ala Pro Asn Ile Tyr Phe Leu
95      100      105
Phe Phe Ser Tyr Gly Ile Val Val Gly Leu Gly Cys Gly Leu Leu
110      115      120
Tyr Thr Ala Thr Val Thr Ile Thr Cys Gln Tyr Phe Asp Asp Arg
125      130      135
Arg Gly Leu Ala Leu Gly Leu Ile Ser Thr Gly Ser Ser Val Gly
140      145      150
Leu Phe Ile Tyr Ala Ala Leu Gln Arg Met Leu Val Glu Phe Tyr
155      160      165
Gly Leu Asp Gly Cys Leu Leu Ile Val Gly Ala Leu Ala Leu Asn
170      175      180
Ile Leu Ala Cys Gly Ser Leu Met Arg Pro Leu Gln Ser Ser Asp
185      190      195
Cys Pro Leu Pro Lys Lys Ile Ala Pro Glu Asp Leu Pro Asp Lys
200      205      210
Tyr Ser Ile Tyr Asn Glu Lys Gly Lys Asn Leu Glu Glu Asn Ile

```

WO 02/40541

PCT/US01/46055

```

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

```

<210> 14

<211> 1232

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7472030CD1

<400> 14

```

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

```

WO 02/40541

PCT/US01/46055

95 100 105
 Leu Leu Thr Leu Tyr Tyr Val Gly Ile Gly Val Ala Ala Leu Ile
 110 115 120
 Phe Gly Tyr Ile Gln Ile Ser Leu Trp Ile Ile Thr Ala Ala Arg
 125 130 135
 Gln Thr Lys Arg Ile Arg Lys Gln Phe Phe His Ser Val Leu Ala
 140 145 150
 Gln Asp Ile Gly Trp Phe Asp Ser Cys Asp Ile Gly Glu Leu Asn
 155 160 165
 Thr Arg Met Thr Asp Asp Ile Asp Lys Ile Ser Asp Gly Ile Gly
 170 175 180
 Asp Lys Ile Ala Leu Leu Phe Gln Asn Met Ser Thr Phe Ser Ile
 185 190 195
 Gly Leu Ala Val Gly Leu Val Lys Gly Trp Lys Leu Thr Leu Val
 200 205 210
 Thr Leu Ser Thr Ser Pro Leu Ile Met Ala Ser Ala Ala Ala Cys
 215 220 225
 Ser Arg Met Val Ile Ser Leu Thr Ser Lys Glu Leu Ser Ala Tyr
 230 235 240
 Ser Lys Ala Gly Ala Val Ala Glu Glu Val Leu Ser Ser Ile Arg
 245 250 255
 Thr Val Ile Ala Phe Arg Ala Gln Glu Lys Glu Leu Gln Arg Tyr
 260 265 270
 Thr Gln Asn Leu Lys Asp Ala Lys Asp Phe Gly Ile Lys Arg Thr
 275 280 285
 Ile Ala Ser Lys Val Ser Leu Gly Ala Val Tyr Phe Phe Met Asn
 290 295 300
 Gly Thr Tyr Gly Leu Ala Phe Trp Tyr Gly Thr Ser Leu Ile Leu
 305 310 315
 Asn Gly Glu Pro Gly Tyr Thr Ile Gly Thr Val Leu Ala Val Phe
 320 325 330
 Phe Ser Val Ile His Ser Ser Tyr Cys Ile Gly Ala Ala Val Pro
 335 340 345
 His Phe Glu Thr Phe Ala Ile Ala Arg Gly Ala Ala Phe His Ile
 350 355 360
 Phe Gln Val Ile Asp Lys Lys Pro Ser Ile Gly Asn Phe Ser Thr
 365 370 375
 Ala Gly Tyr Lys Pro Glu Ser Ile Glu Gly Thr Val Glu Phe Lys
 380 385 390
 Asn Val Ser Phe Asn Tyr Pro Ser Arg Pro Ser Ile Lys Ile Leu
 395 400 405
 Lys Gly Leu Asn Leu Gly Ile Lys Ser Gly Glu Thr Val Ala Leu
 410 415 420
 Val Gly Leu Asn Gly Ser Gly Lys Ser Thr Val Val Gln Leu Leu
 425 430 435
 Gln Arg Leu Tyr Asp Pro Asp Asp Gly Phe Ile Met Val Asp Glu
 440 445 450
 Asn Asp Ile Arg Ala Leu Asn Val Arg His Tyr Arg Asp His Ile
 455 460 465
 Gly Val Val Ser Gln Glu Pro Val Leu Phe Gly Thr Thr Ile Ser
 470 475 480
 Asn Asn Ile Lys Tyr Gly Arg Asp Asp Val Thr Asp Glu Glu Met
 485 490 495
 Glu Arg Ala Ala Arg Glu Ala Asn Ala Tyr Asp Phe Ile Met Glu
 500 505 510
 Phe Pro Asn Lys Phe Asn Thr Leu Val Gly Glu Lys Gly Ala Gln
 515 520 525
 Met Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Leu
 530 535 540
 Val Arg Asn Pro Lys Ile Leu Ile Leu Asp Glu Ala Thr Ser Ala
 545 550 555
 Leu Asp Ser Glu Ser Lys Ser Ala Val Gln Ala Ala Leu Glu Lys
 560 565 570

WO 02/40541

PCT/US01/46055

Ala Ser Lys Gly Arg Thr Thr Ile Val Val Ala His Arg Leu Ser
 575 580 585
 Thr Ile Arg Ser Ala Asp Leu Ile Val Thr Leu Lys Asp Gly Met
 590 595 600
 Leu Ala Glu Lys Gly Ala His Ala Glu Leu Met Ala Lys Arg Gly
 605 610 615
 Leu Tyr Tyr Ser Leu Val Met Ser Gln Asp Ile Lys Lys Ala Asp
 620 625 630
 Glu Gln Met Glu Ser Met Thr Tyr Ser Thr Glu Arg Lys Thr Asn
 635 640 645
 Ser Leu Pro Leu His Ser Val Lys Ser Ile Lys Ser Asp Phe Ile
 650 655 660
 Asp Lys Ala Glu Glu Ser Thr Gln Ser Lys Glu Ile Ser Leu Pro
 665 670 675
 Glu Val Ser Leu Leu Lys Ile Leu Lys Leu Asn Lys Pro Glu Trp
 680 685 690
 Pro Phe Val Val Leu Gly Thr Leu Ala Ser Val Leu Asn Gly Thr
 695 700 705
 Val His Pro Val Phe Ser Ile Ile Phe Ala Lys Ile Ile Thr Met
 710 715 720
 Phe Gly Asn Asn Asp Lys Thr Thr Leu Lys His Asp Ala Glu Ile
 725 730 735
 Tyr Ser Met Ile Phe Val Ile Leu Gly Val Ile Cys Phe Val Ser
 740 745 750
 Tyr Phe Met Gln Asp Ile Ala Trp Phe Asp Glu Lys Glu Asn Ser
 755 760 765
 Thr Gly Gly Leu Thr Thr Ile Leu Ala Ile Asp Ile Ala Gln Ile
 770 775 780
 Gln Gly Ala Thr Gly Ser Arg Ile Gly Val Leu Thr Gln Asn Ala
 785 790 795
 Thr Asn Met Gly Leu Ser Val Ile Ile Ser Phe Ile Tyr Gly Trp
 800 805 810
 Glu Met Thr Phe Leu Ile Leu Ser Ile Ala Pro Val Leu Ala Val
 815 820 825
 Thr Gly Met Ile Glu Thr Ala Ala Met Thr Gly Phe Ala Asn Lys
 830 835 840
 Asp Lys Gln Glu Leu Lys His Ala Gly Lys Ile Ala Thr Glu Ala
 845 850 855
 Leu Glu Asn Ile Arg Thr Ile Val Ser Leu Thr Arg Glu Lys Ala
 860 865 870
 Phe Glu Gln Met Tyr Glu Glu Met Leu Gln Thr Gln His Arg Asn
 875 880 885
 Thr Ser Lys Lys Ala Gln Ile Ile Gly Ser Cys Tyr Ala Phe Ser
 890 895 900
 His Ala Phe Ile Tyr Phe Ala Tyr Ala Ala Gly Phe Arg Phe Gly
 905 910 915
 Ala Tyr Leu Ile Gln Ala Gly Arg Met Thr Pro Glu Gly Met Phe
 920 925 930
 Ile Val Phe Thr Ala Ile Ala Tyr Gly Ala Met Ala Ile Gly Glu
 935 940 945
 Thr Leu Val Leu Ala Pro Glu Tyr Ser Lys Ala Lys Ser Gly Ala
 950 955 960
 Ala His Leu Phe Ala Leu Leu Glu Lys Lys Pro Asn Ile Asp Ser
 965 970 975
 Arg Ser Gln Glu Gly Lys Lys Pro Asp Thr Cys Glu Gly Asn Leu
 980 985 990
 Glu Phe Arg Glu Val Ser Phe Phe Tyr Pro Cys Arg Pro Asp Val
 995 1000 1005
 Phe Ile Leu Arg Gly Leu Ser Leu Ser Ile Glu Arg Gly Lys Thr
 1010 1015 1020
 Val Ala Phe Val Gly Ser Ser Gly Cys Gly Lys Ser Thr Ser Val
 1025 1030 1035
 Gln Leu Leu Gln Arg Leu Tyr Asp Pro Val Gln Gly Gln Val Leu

WO 02/00541

PCT/US01/46055

```

1040      1045      1050
Phe Asp Gly Val Asp Ala Lys Glu Leu Asn Val Gln Trp Leu Arg
1055      1060      1065
Ser Gln Ile Ala Ile Val Pro Gln Glu Pro Val Leu Phe Asn Cys
1070      1075      1080
Ser Ile Ala Glu Asn Ile Ala Tyr Gly Asp Asn Ser Arg Val Val
1085      1090      1095
Pro Leu Asp Glu Ile Lys Glu Ala Ala Asn Ala Ala Asn Ile His
1100      1105      1110
Ser Phe Ile Glu Gly Leu Pro Glu Lys Tyr Asn Thr Gln Val Gly
1115      1120      1125
Leu Lys Gly Ala Gln Leu Ser Gly Gly Gln Lys Gln Arg Leu Ala
1130      1135      1140
Ile Ala Arg Ala Leu Leu Gln Lys Pro Lys Ile Leu Leu Leu Asp
1145      1150      1155
Glu Ala Thr Ser Ala Leu Asp Asn Asp Ser Glu Lys Val Val Gln
1160      1165      1170
His Ala Leu Asp Lys Ala Arg Thr Gly Arg Thr Cys Leu Val Val
1175      1180      1185
Thr His Arg Leu Ser Ala Ile Gln Asn Ala Asp Leu Ile Val Val
1190      1195      1200
Leu His Asn Gly Lys Ile Lys Glu Gln Gly Thr His Gln Glu Leu
1205      1210      1215
Leu Arg Asn Arg Asp Ile Tyr Phe Lys Leu Val Asn Ala Gln Ser
1220      1225      1230
Val Gln

```

<210> 15

<211> 759

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7476089CD1

<400> 15

```

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

```

WO 02/40541

PCT/US01/46055

```

185      190      195
Glu Leu Asp Lys Ala Phe Ser Val Ser Val Leu Ser Val Ser Ser
200      205      210
Gly Ser Leu Gly Ala His Ile Asn Ala Thr Leu Thr Val Leu Ala
215      220      225
Ser Asp Asp Pro Tyr Gly Ile Phe Ile Phe Ser Glu Lys Asn Arg
230      235      240
Pro Val Lys Val Glu Glu Ala Thr Gln Asn Ile Thr Leu Ser Ile
245      250      255
Ile Arg Leu Lys Gly Leu Met Gly Lys Val Leu Val Ser Tyr Ala
260      265      270
Thr Leu Asp Asp Met Glu Lys Pro Pro Tyr Phe Pro Pro Asn Leu
275      280      285
Ala Arg Ala Thr Gln Gly Arg Asp Tyr Ile Pro Ala Ser Gly Phe
290      295      300
Ala Leu Phe Gly Ala Asn Gln Ser Glu Ala Thr Ile Ala Ile Ser
305      310      315
Ile Leu Asp Asp Asp Glu Pro Glu Arg Ser Glu Ser Val Phe Ile
320      325      330
Glu Leu Leu Asn Ser Thr Leu Val Ala Lys Val Gln Ser Arg Ser
335      340      345
Ile Pro Asn Ser Pro Arg Leu Gly Pro Lys Val Glu Thr Ile Ala
350      355      360
Gln Leu Ile Ile Ile Ala Asn Asp Asp Ala Phe Gly Thr Leu Gln
365      370      375
Leu Ser Ala Pro Ile Val Arg Val Ala Glu Asn His Val Gly Pro
380      385      390
Ile Ile Asn Val Thr Arg Thr Gly Gly Ala Phe Ala Asp Val Ser
395      400      405
Val Lys Phe Lys Ala Val Pro Ile Thr Ala Ile Ala Gly Glu Asp
410      415      420
Tyr Ser Ile Ala Ser Ser Asp Val Val Leu Leu Glu Gly Glu Thr
425      430      435
Ser Lys Ala Val Pro Ile Tyr Val Ile Asn Asp Ile Tyr Pro Glu
440      445      450
Leu Glu Glu Ser Phe Leu Val Gln Leu Met Asn Glu Thr Thr Gly
455      460      465
Gly Ala Arg Leu Gly Ala Leu Thr Glu Ala Val Ile Ile Ile Glu
470      475      480
Ala Ser Asp Asp Pro Tyr Gly Leu Phe Gly Phe Gln Ile Thr Lys
485      490      495
Leu Ile Val Glu Glu Pro Glu Phe Asn Ser Val Lys Val Asn Leu
500      505      510
Pro Ile Ile Arg Asn Ser Gly Thr Leu Gly Asn Val Thr Val Gln
515      520      525
Trp Val Ala Thr Ile Asn Gly Gln Leu Ala Thr Gly Asp Leu Arg
530      535      540
Val Val Ser Gly Asn Val Thr Phe Ala Pro Gly Glu Thr Ile Gln
545      550      555
Thr Leu Leu Leu Glu Val Leu Ala Asp Asp Val Pro Glu Ile Glu
560      565      570
Glu Val Ile Gln Val Gln Leu Thr Asp Ala Ser Gly Gly Gly Thr
575      580      585
Ile Gly Leu Asp Arg Ile Ala Asn Ile Ile Ile Pro Ala Asn Asp
590      595      600
Asp Pro Tyr Gly Thr Val Ala Phe Ala Gln Met Val Tyr Arg Val
605      610      615
Gln Glu Pro Leu Glu Arg Ser Ser Cys Ala Asn Ile Thr Val Arg
620      625      630
Arg Ser Gly Gly His Phe Gly Arg Leu Leu Leu Phe Tyr Ser Thr
635      640      645
Ser Asp Ile Asp Val Val Ala Leu Ala Met Glu Glu Gly Gln Asp
650      655      660

```

WO 02/40541

PCT/US01/46055

```

Leu Leu Ser Tyr Tyr Glu Ser Pro Ile Gln Gly Val Pro Asp Pro
665 670 675
Leu Trp Arg Thr Trp Met Asn Val Ser Ala Val Gly Glu Pro Leu
680 685 690
Tyr Thr Cys Ala Thr Leu Cys Leu Lys Glu Gln Ala Cys Ser Ala
695 700 705
Phe Ser Phe Phe Ser Ala Ser Glu Gly Pro Gln Cys Phe Trp Met
710 715 720
Thr Ser Trp Ile Ser Pro Ala Val Asn Asn Ser Asp Phe Trp Thr
725 730 735
Tyr Arg Lys Asn Met Thr Arg Val Ala Ser Leu Leu Val Val Arg
740 745 750
Leu Trp Leu Gly Val Thr Met Ser Leu
755

```

<210> 16

<211> 383

<212> PRT

<213> Homo sapiens

<220>

<223> misc_feature

<223> Incyte ID No: 6428177CD1

<400> 16

```

Met Pro His Arg Lys Glu Arg Pro Ser Gly Ser Ser Leu His Thr
1 5 10 15
His Gly Ser Thr Gly Thr Ala Glu Gly Gly Asn Met Ser Arg Leu
20 25 30
Ser Leu Thr Arg Ser Pro Val Ser Pro Leu Ala Ala Gln Gly Ile
35 40 45
Pro Leu Pro Ala Gln Leu Thr Lys Ser Asn Ala Pro Val His Ile
50 55 60
Asp Val Gly Gly His Met Tyr Thr Ser Ser Leu Ala Thr Leu Thr
65 70 75
Lys Tyr Pro Asp Ser Arg Ile Ser Arg Leu Phe Asn Gly Thr Glu
80 85 90
Pro Ile Val Leu Asp Ser Leu Lys Gln His Tyr Phe Ile Asp Arg
95 100 105
Asp Gly Glu Ile Phe Arg Tyr Val Leu Ser Phe Leu Arg Thr Ser
110 115 120
Lys Leu Leu Leu Pro Asp Asp Phe Lys Asp Phe Ser Leu Leu Tyr
125 130 135
Glu Glu Ala Arg Tyr Tyr Gln Leu Gln Pro Met Val Arg Glu Leu
140 145 150
Glu Arg Trp Gln Gln Glu Gln Arg Arg Arg Ser Arg Ala
155 160 165
Cys Asp Cys Leu Val Val Arg Val Thr Pro Asp Leu Gly Glu Arg
170 175 180
Ile Ala Leu Ser Gly Glu Lys Ala Leu Ile Glu Glu Val Phe Pro
185 190 195
Glu Thr Gly Asp Val Met Cys Asn Ser Val Asn Ala Gly Trp Asn
200 205 210
Gln Asp Pro Thr His Val Ile Arg Phe Pro Leu Asn Gly Tyr Cys
215 220 225
Arg Leu Asn Ser Val Gln Val Leu Glu Arg Leu Phe Gln Arg Gly
230 235 240
Phe Ser Val Ala Ala Ser Cys Gly Gly Gly Val Asp Ser Ser Gln
245 250 255
Phe Ser Glu Tyr Val Leu Cys Arg Glu Glu Arg Arg Pro Gln Pro
260 265 270
Thr Pro Thr Ala Val Arg Ile Lys Gln Glu Pro Leu Asp
275 280

```

WO 02/40541

PCT/US01/46055

```

<210> 17
<211> 1129
<212> FRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477243CD1

<400> 17
Met Phe Arg Arg Ser Leu Asn Arg Phe Cys Ala Gly Glu Glu Lys
1      5      10      15
Arg Val Gly Thr Arg Thr Val Phe Val Gly Asn His Pro Val Ser
20     25     30
Glu Thr Glu Ala Tyr Ile Ala Gln Arg Phe Cys Asp Asn Arg Ile
35     40     45
Val Ser Ser Lys Tyr Thr Leu Trp Asn Phe Leu Pro Lys Asn Leu
50     55     60
Phe Glu Gln Phe Arg Arg Ile Ala Asn Phe Tyr Phe Leu Ile Ile
65     70     75
Phe Leu Val Gln Val Thr Val Asp Thr Pro Thr Ser Pro Val Thr
80     85     90
Ser Gly Leu Pro Leu Phe Phe Val Ile Thr Val Thr Ala Ile Lys
95     100    105
Gln Gly Tyr Glu Asp Cys Leu Arg His Arg Ala Asp Asn Glu Val
110    115    120
Asn Lys Ser Thr Val Tyr Ile Ile Glu Asn Ala Lys Arg Val Arg
125    130    135
Lys Glu Ser Glu Lys Ile Lys Val Gly Asp Val Val Glu Val Gln
140    145    150
Ala Asp Glu Thr Phe Pro Cys Asp Leu Ile Leu Leu Ser Ser Cys
155    160    165
Thr Thr Asp Gly Thr Cys Tyr Val Thr Thr Ala Ser Leu Asp Gly
170    175    180
Glu Ser Asn Cys Lys Thr His Tyr Ala Val Arg Asp Thr Ile Ala
185    190    195
Leu Cys Thr Ala Glu Ser Ile Asp Thr Leu Arg Ala Ala Ile Glu
200    205    210
Cys Glu Gln Pro Gln Pro Asp Leu Tyr Lys Phe Val Gly Arg Ile
215    220    225
Asn Ile Tyr Ser Asn Ser Leu Glu Ala Val Ala Arg Ser Leu Gly
230    235    240
Pro Glu Asn Leu Leu Leu Lys Gly Ala Thr Leu Lys Asn Thr Glu
245    250    255
Lys Ile Tyr Gly Val Ala Val Tyr Thr Gly Met Glu Thr Lys Met
260    265    270
Ala Leu Asn Tyr Gln Gly Lys Ser Gln Lys Arg Ser Ala Val Glu
275    280    285
Lys Ser Ile Asn Ala Phe Leu Ile Val Tyr Leu Phe Ile Leu Leu
290    295    300
Thr Lys Ala Ala Val Cys Thr Thr Leu Lys Tyr Val Trp Gln Ser
305    310    315
Thr Pro Tyr Asn Asp Glu Pro Trp Tyr Asn Gln Lys Thr Gln Lys
320    325    330
Glu Arg Glu Thr Leu Lys Val Leu Lys Met Phe Thr Asp Phe Leu
335    340    345
Ser Phe Met Val Leu Phe Asn Phe Ile Ile Pro Val Ser Met Tyr
350    355    360
Val Thr Val Glu Met Gln Lys Phe Leu Gly Ser Phe Phe Ile Ser
365    370    375
Trp Asp Lys Asp Phe Tyr Asp Glu Glu Ile Asn Glu Gly Ala Leu
380    385    390
Val Asn Thr Ser Asp Leu Asn Glu Glu Leu Gly Gln Val Asp Tyr

```

WO 02/40541

PCT/US01/46055

```

395                               400                               405
Val Phe Thr Asp Lys Thr Gly Thr Leu Thr Glu Asn Ser Met Glu
410                               415                               420
Phe Ile Glu Cys Cys Ile Asp Gly His Lys Tyr Lys Gly Val Thr
425                               430                               435
Gln Glu Val Asp Gly Leu Ser Gln Thr Asp Gly Thr Leu Thr Tyr
440                               445                               450
Phe Asp Lys Val Asp Lys Asn Arg Glu Glu Leu Phe Leu Arg Ala
455                               460                               465
Leu Cys Leu Cys His Thr Val Glu Ile Lys Thr Asn Asp Ala Val
470                               475                               480
Asp Gly Ala Thr Glu Ser Ala Glu Leu Thr Tyr Ile Ser Ser Ser
485                               490                               495
Pro Asp Glu Ile Ala Leu Val Lys Gly Ala Lys Arg Tyr Gly Phe
500                               505                               510
Thr Phe Leu Gly Asn Arg Asn Gly Tyr Met Arg Val Glu Asn Gln
515                               520                               525
Arg Lys Glu Ile Glu Glu Tyr Glu Leu Leu His Thr Leu Asn Phe
530                               535                               540
Asp Ala Val Arg Arg Arg Met Ser Val Ile Val Lys Thr Gln Glu
545                               550                               555
Gly Asp Ile Leu Leu Phe Cys Lys Gly Ala Asp Ser Ala Val Phe
560                               565                               570
Pro Arg Val Gln Asn His Glu Ile Glu Leu Thr Lys Val His Val
575                               580                               585
Glu Arg Asn Ala Met Asp Gly Tyr Arg Thr Leu Cys Val Ala Phe
590                               595                               600
Lys Glu Ile Ala Pro Asp Asp Tyr Glu Arg Ile Asn Arg Gln Leu
605                               610                               615
Ile Glu Ala Lys Met Ala Leu Gln Asp Arg Glu Glu Lys Met Glu
620                               625                               630
Lys Val Phe Asp Asp Ile Glu Thr Asn Met Asn Leu Ile Gly Ala
635                               640                               645
Thr Ala Val Glu Asp Lys Leu Gln Asp Gln Ala Ala Glu Thr Ile
650                               655                               660
Glu Ala Leu His Ala Ala Gly Leu Lys Val Trp Val Leu Thr Gly
665                               670                               675
Asp Lys Met Glu Thr Ala Lys Ser Thr Cys Tyr Ala Cys Arg Leu
680                               685                               690
Phe Gln Thr Asn Thr Glu Leu Leu Glu Leu Thr Thr Lys Thr Ile
695                               700                               705
Glu Glu Ser Glu Arg Lys Glu Asp Arg Leu His Glu Leu Leu Ile
710                               715                               720
Glu Tyr Arg Lys Lys Leu Leu His Glu Phe Pro Lys Ser Thr Arg
725                               730                               735
Ser Phe Lys Lys Ala Trp Thr Glu His Gln Glu Tyr Gly Leu Ile
740                               745                               750
Ile Asp Gly Ser Thr Leu Ser Leu Ile Leu Asn Ser Ser Gln Asp
755                               760                               765
Ser Ser Ser Asn Asn Tyr Lys Ser Ile Phe Leu Gln Ile Cys Met
770                               775                               780
Lys Cys Thr Ala Val Leu Cys Cys Arg Met Ala Pro Leu Gln Lys
785                               790                               795
Ala Gln Ile Val Arg Met Val Lys Asn Leu Lys Gly Ser Pro Ile
800                               805                               810
Thr Leu Ser Ile Gly Asp Gly Ala Asn Asp Val Ser Met Ile Leu
815                               820                               825
Glu Ser His Val Gly Ile Gly Ile Lys Gly Lys Glu Gly Arg Gln
830                               835                               840
Ala Ala Arg Asn Ser Asp Tyr Ser Val Pro Lys Phe Lys His Leu
845                               850                               855
Lys Lys Leu Leu Leu Ala His Gly His Leu Tyr Tyr Val Arg Ile
860                               865                               870

```

WO 02/40541

PCT/US01/46055

Ala His Leu Val Gln Tyr Phe Phe Tyr Lys Asn Leu Cys Phe Ile
 875 880 885
 Leu Pro Gln Phe Leu Tyr Gln Phe Phe Cys Gly Phe Ser Gln Gln
 890 895 900
 Pro Leu Tyr Asp Ala Ala Tyr Leu Thr Met Tyr Asn Ile Cys Phe
 905 910 915
 Thr Ser Leu Pro Ile Leu Ala Tyr Ser Leu Leu Glu Gln His Ile
 920 925 930
 Asn Ile Asp Thr Leu Thr Ser Asp Pro Arg Leu Tyr Met Lys Ile
 935 940 945
 Ser Gly Asn Ala Met Leu Gln Leu Gly Pro Phe Leu Tyr Trp Thr
 950 955 960
 Phe Leu Ala Ala Phe Glu Gly Thr Val Phe Phe Phe Gly Thr Tyr
 965 970 975
 Phe Leu Phe Gln Thr Ala Ser Leu Glu Gln Asn Gly Lys Val Tyr
 980 985 990
 Gly Asn Trp Thr Phe Gly Thr Ile Val Phe Thr Val Leu Val Phe
 995 1000 1005
 Thr Val Thr Leu Lys Leu Ala Leu Asp Thr Arg Phe Trp Thr Trp
 1010 1015 1020
 Ile Asn His Phe Val Ile Trp Gly Ser Leu Ala Phe Tyr Val Phe
 1025 1030 1035
 Phe Ser Phe Phe Trp Gly Gly Ile Ile Trp Pro Phe Leu Lys Gln
 1040 1045 1050
 Gln Arg Met Tyr Phe Val Phe Ala Gln Met Leu Ser Ser Val Ser
 1055 1060 1065
 Thr Trp Leu Ala Ile Ile Leu Leu Ile Phe Ile Ser Leu Phe Pro
 1070 1075 1080
 Glu Ile Leu Leu Ile Val Leu Lys Asn Val Arg Arg Arg Ser Ala
 1085 1090 1095
 Arg Arg Asn Leu Ser Cys Arg Arg Ala Ser Asp Ser Leu Ser Ala
 1100 1105 1110
 Arg Pro Ser Val Arg Pro Leu Leu Leu Arg Thr Phe Ser Asp Glu
 1115 1120 1125
 Ser Asn Val Leu

<210> 13

<211> 648

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyta ID No: 7473042CD1

<400> 18

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

WO 02/40541

PCT/US01/46055

Glu Leu Leu Val Ser Ser Thr Val Gly Ala Ala Ala Val Ser Ala
 125 130 135
 Leu Ala Gly Gly Ala Leu Asn Gly Val Phe Gly Arg Arg Ala Ala
 140 145 150
 Ile Leu Leu Ala Ser Ala Leu Phe Thr Ala Gly Ser Ala Val Leu
 155 160 165
 Ala Ala Ala Asn Asn Lys Glu Thr Leu Leu Ala Gly Arg Leu Val
 170 175 180
 Val Gly Leu Gly Ile Gly Ile Ala Ser Met Thr Val Pro Val Tyr
 185 190 195
 Ile Ala Glu Val Ser Pro Pro Asn Leu Arg Gly Arg Leu Val Thr
 200 205 210
 Ile Asn Thr Leu Phe Ile Thr Gly Gly Gln Phe Phe Ala Ser Val
 215 220 225
 Val Asp Gly Ala Phe Ser Tyr Leu Gln Lys Asp Gly Trp Arg Tyr
 230 235 240
 Met Leu Gly Leu Ala Val Val Pro Ala Val Ile Gln Phe Phe Gly
 245 250 255
 Phe Leu Phe Leu Pro Glu Ser Pro Arg Trp Leu Ile Gln Lys Gly
 260 265 270
 Gln Thr Gln Lys Ala Arg Arg Ile Leu Ser Gln Met Arg Gly Asn
 275 280 285
 Gln Thr Ile Asp Glu Glu Tyr Asp Ser Ile Lys Asn Asn Ile Glu
 290 295 300
 Glu Glu Glu Lys Glu Val Gly Ser Ala Gly Pro Val Ile Cys Arg
 305 310 315
 Met Leu Ser Tyr Pro Gln Thr Arg Arg Ala Leu Ile Val Gly Cys
 320 325 330
 Gly Leu Gln Met Phe Glu Gln Leu Ser Gly Ile Asn Thr Ile Met
 335 340 345
 Tyr Tyr Ser Ala Thr Ile Leu Gln Met Ser Gly Val Glu Asp Asp
 350 355 360
 Arg Leu Ala Ile Trp Leu Ala Ser Val Thr Ala Phe Thr Asn Phe
 365 370 375
 Ile Phe Thr Leu Val Gly Val Trp Leu Val Glu Lys Val Gly Arg
 380 385 390
 Arg Lys Leu Thr Phe Gly Ser Leu Ala Gly Thr Thr Val Ala Leu
 395 400 405
 Ile Ile Leu Ala Leu Gly Phe Val Leu Ser Ala Gln Val Ser Pro
 410 415 420
 Arg Ile Thr Phe Lys Pro Ile Ala Pro Ser Gly Gln Asn Ala Thr
 425 430 435
 Cys Thr Arg Tyr Ser Tyr Cys Asn Glu Cys Met Leu Asp Pro Asp
 440 445 450
 Cys Gly Phe Cys Tyr Lys Met Asn Lys Ser Thr Val Ile Asp Ser
 455 460 465
 Ser Cys Val Pro Val Asn Lys Ala Ser Thr Asn Glu Ala Ala Trp
 470 475 480
 Gly Arg Cys Glu Asn Glu Thr Lys Phe Lys Thr Glu Asp Ile Phe
 485 490 495
 Trp Ala Tyr Asn Phe Cys Pro Thr Pro Tyr Ser Trp Thr Ala Leu
 500 505 510
 Leu Gly Leu Ile Leu Tyr Leu Val Phe Phe Ala Pro Gly Met Gly
 515 520 525
 Pro Met Pro Trp Thr Val Asn Ser Glu Ile Tyr Pro Leu Trp Ala
 530 535 540
 Arg Ser Thr Gly Asn Ala Cys Ser Ser Gly Ile Asn Trp Ile Phe
 545 550 555
 Asn Val Leu Val Ser Leu Thr Phe Leu His Thr Ala Glu Tyr Leu
 560 565 570
 Thr Tyr Tyr Gly Ala Phe Phe Leu Tyr Ala Gly Phe Ala Ala Val
 575 580 585
 Gly Leu Leu Phe Ile Tyr Gly Cys Leu Pro Glu Thr Lys Gly Lys

WO 02/06541

PCT/US01/46055

```

590      Lys Leu Glu Glu Ile Glu Ser Leu Phe 595      Asp Asn Arg Leu Cys 600
605      Cys Gly Thr Ser Asp Ser Asp Glu Gly Arg Tyr Ile Glu Tyr Ile 615
620      Arg Val Lys Gly Ser Asn Tyr His Leu Ser Asp Asn Asp Ala Ser 630
635      Asp Val Glu 640      Asp Asn Asp Ala Ser 645

```

```

<210> 19
<211> 545
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7402060CD1

```

```

<400> 19
Met Thr Phe Gly Arg Ser Gly Ala Ala Ser Val Val Leu Asn Val
1      5      10      15
Gly Gly Ala Arg Tyr Ser Leu Ser Arg Glu Leu Leu Lys Asp Phe
20     25     30
Pro Leu Arg Arg Val Ser Arg Leu His Gly Cys Arg Ser Glu Arg
35     40     45
Asp Val Leu Glu Val Cys Asp Asp Tyr Asp Arg Glu Arg Asn Glu
50     55     60
Tyr Phe Phe Asp Arg His Ser Glu Ala Phe Gly Phe Ile Leu Leu
65     70     75
Tyr Val Arg Gly His Gly Lys Leu Arg Phe Ala Pro Arg Met Cys
80     85     90
Glu Leu Ser Phe Tyr Asn Glu Met Ile Tyr Trp Gly Leu Glu Gly
95     100    105
Ala His Leu Glu Tyr Cys Cys Gln Arg Arg Leu Asp Asp Arg Met
110    115    120
Ser Asp Thr Tyr Thr Phe Tyr Ser Ala Asp Glu Pro Gly Val Leu
125    130    135
Gly Arg Asp Glu Ala Arg Pro Gly Ala Arg Gly Gly Ser Leu Gln
140    145    150
Ala Leu Ala Gly Ala His Ala Ala Asp Leu Arg Gly Ala His Ile
155    160    165
Leu Ala Ser Val Ser Val Val Phe Val Ile Val Ser Met Val Val
170    175    180
Leu Cys Ala Ser Thr Leu Pro Asp Trp Arg Asn Ala Ala Ala Asp
185    190    195
Asn Arg Ser Leu Asp Asp Arg Ser Arg Ile Ile Glu Ala Ile Cys
200    205    210
Ile Gly Trp Phe Thr Ala Glu Cys Ile Val Arg Phe Ile Val Ser
215    220    225
Lys Asn Lys Cys Glu Phe Val Lys Arg Pro Leu Asn Ile Ile Asp
230    235    240
Leu Leu Ala Ile Thr Pro Tyr Tyr Ile Ser Val Leu Met Thr Val
245    250    255
Phe Thr Gly Glu Asn Ser Gln Leu Gln Arg Ala Gly Val Thr Leu
260    265    270
Arg Val Leu Arg Met Met Arg Ile Phe Trp Val Ile Lys Leu Ala
275    280    285
Arg His Phe Ile Gly Leu Gln Thr Leu Gly Leu Thr Leu Lys Arg
290    295    300
Cys Tyr Arg Glu Met Val Met Leu Leu Val Phe Ile Cys Val Ala
305    310    315
Met Ala Ile Phe Ser Ala Leu Ser Gln Leu Leu Glu His Gly Leu

```


WO 02/00541

PCT/US01/46055

```

320      325      330
Asp Leu Glu Thr Ser Asn Lys Asp Phe Thr Ser Ile Pro Ala Ala
335      340      345
Cys Trp Trp Val Ile Ile Ser Met Thr Thr Val Gly Tyr Gly Asp
350      355      360
Met Tyr Pro Ile Thr Val Pro Gly Arg Ile Leu Gly Gly Val Cys
365      370      375
Val Val Ser Gly Ile Val Leu Leu Ala Leu Pro Ile Thr Phe Ile
380      385      390
Tyr His Ser Phe Val Gln Cys Tyr His Glu Leu Lys Phe Arg Ser
395      400      405
Ala Arg Ser Ile Cys Leu Thr Ser Val Thr Ser Val Leu Gly Thr
410      415      420
Val Gly Tyr Thr Glu Met Thr Ile Asn Gly Pro Cys Pro Asp Ala
425      430      435
Leu Arg Asp Pro Cys Thr Cys Lys Lys Pro Leu Lys Thr His Ser
440      445      450
Gly Val Leu Tyr Lys Ala Met Ala Asp Leu Trp Gln Ser Leu Glu
455      460      465
Gly Gly Pro Pro Val Glu Gln Leu Pro Pro Asp Pro Leu Thr Arg
470      475      480
Trp Cys Phe His Pro Ala Gly Ser Thr Leu Cys Gly Pro Ala Asn
485      490      495
Ser Met Ala Val Ala Ser Pro Gly Ser Arg Pro Ala Ala Pro Gly
500      505      510
Gly Gly Phe Leu Arg Thr Glu Ala Leu Val Leu Ile Val Ala Ala
515      520      525
Gly Pro Val Asp Gly Leu Asn Cys Glu Asn His Pro Phe Arg Gly
530      535      540
Gly Cys Lys Asp Phe
545

```

<210> 20

<211> 262

<212> PST

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1578772CD1

<400> 20

```

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

```

WO 02/40541

PCT/US01/46055

155 160 165
 Pro Val Arg Val Val Val Leu Asp Phe Ser Gly Val Thr Phe Ala
 170 175 180
 Asp Ala Ala Gly Ala Arg Glu Val Val Gln Leu Ala Ser Arg Cys
 185 190 195
 Arg Asp Ala Arg Ile Arg Leu Leu Leu Ala Gln Cys Asn Ala Leu
 200 205 210
 Val Gln Gly Thr Leu Thr Arg Val Gly Leu Leu Asp Arg Val Thr
 215 220 225
 Pro Asp Gln Leu Phe Val Ser Val Gln Asp Ala Ala Ala Tyr Ala
 230 235 240
 Leu Gly Ser Leu Val Arg Gly Ser Ser Thr Arg Ser Gly Ser Gln
 245 250 255
 Glu Ala Leu Gly Cys Gly Lys

<210> 21
 <211> 1373
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1626101CB1

<400> 21
 caggcggggc ctgggggggg cggccactct aaccagggca aactgtccct ggaacaggag 60
 gaggaagacg aactctggcg gctcttagga cgcagagacg cgtcccgcc ctccattgag 120
 cccaactgtc gctctctggat ccccgagcgc caatccctta ttgcacgatt tcttcaatgg 180
 acggaattat tagatccctac aactgtgttc atttcagttg aaggtataga aaactcgagg 240
 caactattgt gcacaaatga agatgtttcc agccctgccl cggcggacca aaggatcacg 300
 ggaagcttggg agcggagttc tgcacacagc catcccgaca gcagcaacct gatcccccag 360
 cttttttggac ctgcagcgtt cctgcctttc atggcgccca cggatatttt gtcaatgacg 420
 coactgaagc gcatcaatgc cgtgatttta cctcaggttt tctctgtgac ctacatggca 480
 gcgttcaaca gcatcaatgg aacacagcgt tacacttgta agccactaga aagatcatta 540
 ctaatggcgg gagccgttgc ttcttcaact ttcttaggg taatccctca gtttgcacg 600
 atgaagtatg gcttgacttg ccttggtatt acaagactct tacctgtgat ctctctcgtg 660
 caagccagtg gaatgaatgt ctacatgtcc cgaagtcttg aatccattaa ggggattcgg 720
 gtcatggaca aggaaggcaa tgtcctgggt caatccagaa ttgctggggc aaaggctggt 780
 aggaanaacgc tagcatccag aatagtctg tttgggacct cagctctgat tctgaagtc 840
 ttcaactact tttttaaaag gacccagtat ttccggaaaa acccagggtc attgtggatt 900
 ttgaactgtt cctgtactgt cctggcaatg ggaatgatgg tgcattttc tttagtata 960
 ttccacagca ttgtacagat acgtactgt agtcttgaag ggaanaatca gtctccaca 1020
 gaagaaacag aaacttttta tccacagagg gtgtaggcgt gatttttagg tgaatttatg 1080
 tggttcctgc ttgaaaacct tcccctctcc aggttcgggt tagagaactt tgcacacagt 1140
 cctctgggga cccacagcgt gtctgtgctg acnaggcgac ttcaagattcc atactgagat 1200
 cgttccacag ctggcgtctc tgggtgtttt aaggctggct ggaagaaca gtgggagggt 1260
 gcccgctctg acacccctgg ggttgcctgag ggaacggttg gactggggat cggctatgca 1320
 aaggatactg tgaatcaact aattaactna taaacctgtc tcaagttagg gaa 1373

<210> 22
 <211> 3231
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2907828CB1

<400> 22
 ttccgggggt gggcgggctg caacagcttc gggctcgggg ttttgggggc ggcgcggcgc 60
 ggcctaggctg cgcggtgcgg accccggcgc gggctcgggg ttgctggggc ggcgcgtaag 120
 atgcctctaa tggaggagtt tctgagcagc acccctggcc cagtgccctt gaaggggagc 180

WO 02/40541

PCT/US01/46055

```

tcaaacacaga gactatttca agccctggat atcatatcct gaggccacaa ggaagagaga 240
acatggctgt gaggtttgat gacgacgtgc cgtcatcctc gaccttggat gagggtggca 300
gtgcccacat ggtccctccc aacggcctgg gccaaagaga gctaccttagc aaaaatggcg 360
ggagctatgc catccacgac tcccagccgc ccagctctcg ctctggggat gagggttccc 420
ccttccagccc gcaacacaa cgggagatga attaccaaga ggcagcaatc tacctccagg 480
aagccgagaa caacgacacg ttcttccacc acccccagga tgcacaggcg ctggcgagct 540
acctctttgc acacacacac ctctcttacc tgatggagct ggcacaggcc ctgctgtctc 600
tgctgtctct cctgtgcgag gcccccgcgc tcccgcacat ccggcttggc atctatgtcc 660
acgcacaccc gtagctgttt gccctgatgg tggtagtgtt tgaactctgc atgaagtac 720
gctggtggg cctccacacc ttcctccggc acaaggggac catggtcaag acctgtgtgc 780
tgacccgagc actgctgtgc atttctctgg tggagctgtc gtattgcggt ggcttccggc 840
gcaacctgcg gcaatcttcc cagtcctctg ccctctctat ggaactctcc ctgctgtctc 900
tgcttctcat gatcatcttt gccatctctg gtttctactt gtttctccct aacctctcag 960
acctctactt aggcacccct gagaacagca tgcctagttt gtttgtctct ctgacacag 1020
ccaatttccc agatgtgat atgctctct actcccgga ccctggtctc tgcgtctctc 1080
tcattgtgta ccttccacac gagctgtatt tcatcatgaa cctgctctgc gctgtggtgt 1140
tcgaacccct caatgacatt gagaacgga agttcaagtc ttgtctactg cacaagcgaa 1200
ccgtctatca gcatgctac ccctgtcttc tcagccagag gaggcctctc gccatctctc 1260
acagcgaggt tgaaggccct atgctctctt acagccccc gatgagtgcc agggagcgct 1320
ctcttactct caagccctct aatcagaaca acacccccc gctcagccta agggactttt 1380
acgatctcta ggaagtttgt gctttgaagt ggaaggccaa gaaaacacga gaggacttgt 1440
ttgatgagct tcccaggagc gcgtctctca tcttcaaggt tattaatata ctgtgtgaat 1500
ccaagggcct ccagtatctc atgtacttgg tgggtggcag caacggggto tggatctctc 1560
tggagacatt tatgtgaaa ggtgggaact tcttctccaa gcaagtgccc tggagttaac 1620
tcgtcttctt aactatctat ggggtggagc tgttctctga ggttgcggc ctggggccctg 1680
tggagttaatt gtttccggga tggaaattgt ttgaactctc cgtgacagtg ttgccttctc 1740
tggagctgtc gggctctggc ctcaacatgg agccctctca ttctatctgt gctctgagc 1800
cctctcagct cctgaggttg tttaagttga aggaagctta ccgcaagtg ctggaacaa 1860
tgttctagct gctgcccgg atggccagcc tgggtctaac cctgctctat tttaataat 1920
ccttgcctat cgtgggcatg gaggttctct gcgggacgt ctccccaac tgcgcaaca 2040
cgagtcacgt ccagatgctc taccctggc gccacacac cgtgggcaac aggaacctgt 2100
tggaggaagc ctactattat ctcaataatt ttgaacacat cctcaacagc ttgtgtgccc 2160
tgctttagct cacagttgtc acaacttggt acatcatcat ggaaggcgtc acctctcaga 2220
ccttccactg gacccgctc taactctatg cctttacat tgtgacctg gtggtgatga 2280
cgatcatgtt cgcctttatc ctccagpctc tcttctctcc aatgaactac agcccgcaag 2340
accaagactc ggaagttgat ggtggcatca cctttagaaa ggaactctcc aaagaagagc 2400
tggttgccgt cctggagctc taccgggagc caacgggggc ctccctggat gtcacacagg 2460
tgctggagac cctctccag atggagagat accagcaaca ttccatggty ttctctggag 2520
ggcgtacaa gacacagagc gacctgagcc tgaagatga ccaggggag atccaggagt 2580
ggtatgagga gcatggcagg gacgaagagc agcagcgaca actcagcagc agtgagccc 2640
ccgcccacaa gtagccccc ggcagcgccc agcgtctcca gacgttacc tagccagcg 2700
cccgaaagcc gttcttcta tgcacacaa caatagatt actctactg gatgtacga 2760
actgcgtgt gtgtacacat actcactgat atgacatat ttatatcag gaagaaaaa 2820
gaagacacag atggggcttg gttataacc accttgcct gtcttctta actccagag 2880
caggtttagt gagggttggg ggtggggcca ccaggtctga gcttctcta ctgtggaag 2940
ctccagagag ccttccacaa gggagccccc cacctggate cagtcactgt ccgggctctc 3000
ccctcatgtg gcttggcctc catogggcac gtccaaagct gtcactgcta ctgottcag 3060
ctcacatccc cccgacctga tggcgtgccc gccccctctc cctggggccc atgccaagc 3120
tttctgtgtt ttgtcttagg gacagaacca cttaggaagg aaagaactcc oggtctccag 3180
ggtggtattt cagtgtctgt gataatgtca cgcacaccc cttcggggac c 3231

```

<210> 23
 <211> 3160
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3968527CB1

<400> 23
 atgacggaca acatcccgct gacgcccgtg cgcagagaaga agcggatgga cagcagggcc 60

WO 02/40541

PCT/US01/46055

```

cgccgcccgt gctgcgagtg gctgcgagtg tgcggtggag gggaggccag ccccgccact 120
gtctggctgg ggcacccgac gaggagagac cagagtgatc ctgggaatgt catcaccaat 130
cagaagtaga attttcttcc etttcttctt ggggtgctgt tcaaccagtt caataacttt 240
ttcaacctctt atttcttact tcttgcctgc tctcagtttg ttcccgaaat gagacttggt 300
gcactctata cctactgggt tccctcgggc ttctgtctgg ccgtcaactgt catcctgag 360
ccggtggagg agatccgatg ctacgtcgcc gacaaaggag tcaactccca ggtctacagc 420
cggctccacg ccagaggcac agtgaagggt anaggtctta acatccaggt lggagccctt 480
atcatcgttg aaaaagacca ggcgggtcct gccagataga tcttcttgag gacatcagaa 540
aaaaacgggt catgcttctt ggcgacggat cagctggatg gggagacgga ctggaagctg 600
cggcttcccg tggccttgca cgcagggctc ccacggcccg ccgaactctt tcagatcaga 660
tcgbatgtgt acgcagaaga gccaaatatt gacattcaca acttcgtggg aacttttaac 720
cggagaagaca ggcgaccccc gatcacggag agccttgaga tagagaaacac gctgtggggt 780
ggcaotgttg tgcacacagg taactgttgt ggtgttgtt ttacacctgg cagagaactc 840
cggagtgaca tgaatacctc aaatcccccga agtaagatcg gctgtttcga cttygaagtg 900
aactgcctca ccaagatcct ctctgggtgc ctgggtgggg tctcgctagt catgggtgac 960
cttcagcct ttgcggcgcc ttggtacctg cagatcacac gcttctctct ctgttttcc 1020
aacatctctc ccatagtttt ggcggtggaac ctggaactgg gcaagatcgt gtacagctgg 1080
gtgattcgaa gggactcaaa aatcccccgg accgtgtgtt cctcccgccc gatctctgag 1140
cagctgggca ggaattctga ttaactcaca gacaaagac gacctcttac ccagaacgag 1200
atgattttca aaggcttcca tctcggaaca gtacgctacg gctcgaactc aatggacgaa 1260
gtcaaaagcc acattttcag catattacac cagcaatccc aggaacccacc ggtcagaag 1320
ggcccaagcc tcaacactaa ggtccggcgg accatgagca gccgcgtgca cgaagccgtg 1380
aaggccatcg cgtctgcca caacgtgact ccctgtgatg agtccaaagg tgtgactgat 1440
caggtctgag ccgagaagca gtacgaagaa tcttgccggg tatacgaagg atccagcccc 1500
gatgaagtgg cctctgtaca gtgacggaaa agtctggggt taacccctgt gggccgagac 1560
cagttctcca tgcagctgag gacccctggc gacagatccc tgaacttcac catcctacag 1620
atcttccctt tcaactatga aagcaaaagt atgggcatca tctgtcgagg tgaataaact 1680
ggggaatata cgttttacct gaaggagaca gatgggttca tggctgggat tgtcagtac 1740
aatgaatggt tggaggaaga gtgtggcaac atggtccgag aagggctggg ggtgctctg 1800
gtggcaaga agtctcttgc agaggagag tatcaggagc ttgaagcccg ctactgtcag 1860
gccaaagctga ggtgcaaga ccgctccttc aaagtggcca cgtgatcga gagcctggag 1920
atggagatgg aactgctgtg cctgacggcg gtggaggaac agctgacggc agatgtccgg 1980
cccgagctgg agacccctga gaatgtctgc atcaaggttt gcatgctgac aggggacag 2040
ctggagacag ctacgtgcac agcgagaat gacatctggt tgcacgaaa ccaagacatc 2100
cagtttttcc ggtgtgtgac caacccggcg gaggctcacc tggagctgaa cgccttcgcg 2160
aggaagcatg atgtgcccct ggtcatctcg ggagactccc tggaggtttg cctcaagtac 2220
tatgagtacg agttcatgga gctggcctgc cagtgcctgg ccgtagtctg ctgcagatgt 2280
gcccccaccc agaagggcca gatctgtcgc ctgcttcagg agcgccaggg caagctcacc 2340
tgtgcagtay gggacggagg caatgacgto agcatgatc aggaatctga ctgcggcgctg 2400
ggagtggagg gaaagggaag aaaaacggct tegtggctg cagaactctc catcactcaa 2460
tttaagcact ttggccgggtt gcttatggta catggccgga acagctaca ggggtcagcc 2520
gcccacagcc agtctgtgat tcaacagagc ctctglatca gaacctagaa ggtgtcttt 2580
tccctcgtgt ttactcttgc ctccgtcctc ctctatcaag gattcctcat catgggtac 2640
tcccgaattt acaccatggt tctctgtgtt tctctgttcc tggacaaaga tgtcaaatcg 2700
gaagtgtcca cgtgtatctc tgaagctaac aaggaccttc tcaaggagac gccgtgtctc 2760
tacaagacat ccttaaatg ggttttgatt agcatctatc aaggagacac catcatgtac 2820
ggggcgctgc tegtgtttga ctoggagttc gtgcacatcc tggccacttc ettcactcgt 2880
ctgatctcca ccagctgctc catggtggcg ctgacatccc agacctggca ctggctcatg 2940
acagtgccgg agctgctcag cctggcctgc tactatcgct cctgtgtgtt ettcacagag 3000
ttcatcgatg tgtacttcac cgcaccttg tcattcttkt ggaagttctc cgtcatcact 3060
ctggtcagct gctcctccct ctatgtcttc aagtactctc gaagacggtt ctctccccc 3120
agctactcaa agctcacatc ataggccgtg cgttcgtctg 3160

```

<310> 24

<211> 2848

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7472732CB1

<400> 24

WO 02/40541

PCT/US01/46055

```

cttaacactg aacccattac ttttcaaga ccagaaaaa atattacatg aacaggaaact 60
acttctccct cagataagaa ttcaagctt gacatgtta accacagag aattggagct 120
gggcatggaa aggaagtgtt ctgcaatgat ttttttctt gtttagagaa gttttctct 180
acaaagagaa atctgaaaaa tgcaggagc aaagaggaaa aaganaagca tgccttgag 240
caagatgcac acccccagt gtgagagcat tatacagtgq tctagaggc gactgcccac 300
tttggatttg gccaccatt acaatctgaa agaaaaactg ctccagaca ctgtgtcttg 360
gataatgttg gcagttcaac aggtgacca aggattggcc tttgctgttc tctcatctgt 420
gcccccagtg tttggtttat atgggtctct gtttcttgcc ataatttat ccatatttg 480
aatgggaact catgttgcca caggcaactt tgccttgaca tcttaatat cagccaaagc 540
cgtggaaagg atgttccctc agaaactgca gaattctcac acacagagta acacaagcgt 600
gctgggctta tccgacttgg aaatgcaaa gatcaactt gctgagcag tttctctct 660
gggaggtgtg attcaggttg ccatgtttgt gctgcaactt ggcagtgcca catttgtgt 720
cacagagcct gtgactagcg caatgacaac tggggctgcc accatgttg tgacttaca 780
agtcacatat ctcttgggaa tgaaatgccc atatatatcc ggacacactg gattcttla 840
tatttatgca tatgtttttg aaaaactcaa gtctgtgaga ctggagcat tgcctttatc 900
ctgtctgagc atgtgtgttc ttgttctgt taagagctg aatgaacag ttaaaagaa 960
aattcaagtt gttctctctg tagatttagt ttggctcca aacacatcgc cactccatca 1020
ccactagac tgtctcttgg ccaactttct tgagccccc tgggaggatg gacttccaga 1080
aggtgctctc aacaggcag aaggacattt gcgcaggaac ataattctc cacttagagc 1140
tcccacpatg aacatctctc ctgctgtgat cctgaagct ttcggagtgg caettgtagg 1200
ctatgtggcc tcactgctct ttgctcaagg atctgcaaaa aaattcaaat attcaattga 1260
tgcaaacacg gaatttttgg cccatggcct cagcaatata gttttctcat ttttctctg 1320
catcaaaagt gctgctgcca tgggaaggac ggtctgctg tacagcacag gagcgaagac 1380
acaggtggct tgtctaatat ctgtcatttt cgtctctata gtcatctatg caataggacc 1440
tttctcttac tggctgcccc tgtgtgtctc tgcagcatt atgtgtgtg gactgaaggg 1500
aatgctaatc cagttccgag atttaaaaa atattggaat gtggataaaa toagttggg 1560
aatatgggtc agtacaatg tattacaat atgtttgtt gccaattgtg gactgtgtgt 1620
tgggtgttgt tgaacatag ctatagtgat aggaacgttc ccaagagcaa tgactgtag 1680
tataaaaact atgaagaaa tgggaattta agtgaagaca gaalggaca gtaaacctt 1740
gcagcaggtg aaactatct caataaaca ccaggttgtt tctctgaatg caaaagatt 1800
ttatctgat tttaagaca tgatccaaa ggaatctgct tgaatcagc cacttgatga 1860
tataagcagc tgaacaaaa acacattgct taattctcta tccatggca actgcaatga 1920
agaaagctta gactcctgcc ctaatgagaa gtgttattta atcctggatt gcagtggatt 1980
tacttttttt gactattctg gagtctccat gcttgttggc gtttacatg actgtaaag 2040
caggagtgtg gatgtattgt tagccattg taagctctc ttgataaaag caatgacga 2100
ttatggaaac ctatagctcg agaaacaaat tttttttgaa tgggtatctg ctgcaataag 2160
tcatatccat tcaataaaga atttgagcaa actcagtga cagagtgaag tctgagaccc 2220
tttltgtaca gtacagctct tgtctttacc aactgcctga agaggccata tctgagcatt 2280
ttgcaacaact ttttggttgt tttagatcta cagatgaact ctgctacaat aagtacagtg 2340
tgacttagta actgcatagc agtggaaag aactgcaac tttttttct catctttgtt 2400
agtaagaaga ttcgcttagt tattttatgt aaaaatcagt atgtgttttg ttttagtgta 2460
ctgaagggtg aacatgggtt tattttatgt taactatata ttttgtgtg ttttatttct 2520
atttgtgtgt aagttgatgt ttaaaattga gaattatga ttctatagt aatttggaac 2580
atttaacagc cactgttaa atttcaagt aactgtaac taatacata aaacaccta 2640
gcaaatgtgc cacttttaca caactctct ctgtatagc ctctgaata rcaataaggg 2700
taaatattac ttacacagc aagatgtgaa attcaaaaa agtaacaaa actaaagaa 2760
tgaaaaactg gaataaatc gtttccatat ctttccatc gtccatttct gaagtattca 2820
ggaaatgttt caataogaa agaaacgg 2848

```

<210> 25
 <211> 3727
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7476938CB1

<400> 25
 atggttatgg agcctgggga gtccaaagggc atagtctgtt catctggcna gggccttcat 60
 gctgcacatc tcatggtgga agtgaaacac gtaagagaag ggattggctc agaaatgggc 120
 acctgcccc aatggaccaa tctttctcat tgcacaaatg gaataatgac agttttgtgt 180
 aaggggtctg tgcctgaggg aagcgggaag caaaagcggg tctgtctagc cccgaggctc 240

WO 02/40541

PCT/US01/46055

```

cgaactcggt ggtcctggaa gctccgcagg atgggggaga agatggcgga agaggagagg 300
ttccccaaat caactccatga gggtctcaat gtcacccctc acacccacct ggttgcacg 360
acgaaactgg tgcctccgac ccttgggaag cccatccctc ccgtggagac agggagagag 420
gcccagacag aggagcgatc cagcgccatg accatcttct tcagcctcct tctctcagct 480
atctgcatac tatctgtgca ttactctgac cgatacagat tacatttctt gccagagagt 540
gttctgtgtt ttctcttagg tttctctcat ggagcagtta taanaattat agagttttaa 600
aaactggcga attggaagga agaagaaatg ttctgtccaa acatgttttt cctctcctg 660
cttcccccta ttatctttga gttctgatat tcattacaca aggttaacct ctttcaaaat 720
attggttcca tcacctgttt tctgtgtttt ggagcggcaa tctcgccttt tctagttagt 780
ggaggaattt atttctggg tcaggctgat gtaactctta aactcaaat gacagacagt 840
tttgcgtttt gtcctcaat atctgctgtc gatccagtgg cactattgc catttcaat 900
gcacttcagt tggacccctg gtcacacatg ctggtctttg gagaagatat tctcaacagt 960
gcacttcca ttgttctgac caacacagct gaaggtttta caagaaaaaa tatgtcagat 1020
gtcagtggtt ggcacacatt tttaacagcc ctgactact tctcaaaaat gttctttggc 1080
tctgcagcgc tgggcactct cactggctta atttctgcat tagtgctgaa gcatattgac 1140
ttgaggaaaa cgcctctctt ggagtttggc atgatgatca tttttgctta tctgcttat 1200
gggtctgcag aagaaactct actctcaggg atcagtgcca tctgtttctc agggacgag 1260
atgtccact acagcagaca taacctctcc ccagtcaacc agatctcact gtagcagacc 1320
ctccgcaccc tggcctctct atgtgaaaca tgtgtgtttg catttctgg cctgtccatt 1380
tttagtttca ctacaaagt ttgaatttcc ttgtctact ggtgcatagt ccttgtacta 1440
tttggcagag cgttaaacat ttctctctt tctactctcc tgaatttctt ccgggcatat 1500
aaaaacacac cgaagatgat gttcatcatg tggtttagtg cctgcgggg agccatcccc 1560
tatgccttga gctcactct ggacctggag cccatggaga agcggaagt catcggaacc 1620
accacatctg tcatctgtct ctccaccatc ctgctgtctg cgggcagcac catgcccctc 1680
atttgcctca tggacatcga ggaacccaa gcauaccgca ggaacaagaa ggaactcaac 1740
ctcagcaaga ctgagacagat gggcaacact gtgagatcgg agcactgttc ggaactcaag 1800
gagggagagt agggagccca ctacatcagg cggcagggac ttaagggtct cgtctggtct 1860
gacgcacagt acctgaacco ctctctcact cggaggtctg cgcaggagga cctgcacacc 1920
gggcgcctcc agatgaaaaa tctcaccaac aagtggtaag agggagtaag ccaggggccc 1980
tcggctcccg agagcagcga gcaggagctg ctctgagccc aggtgccaag gcttcaagca 2040
gggcggccca ggttggcgtt caggaatata gaggcgagg gagaagagct cgcagagagt 2100
cgttcactca gtagcccttt cagaacctga caggcctctg gacccaggcg acttcttggg aaactgtcat 2160
ccctttagta ccaaacctga caggcctctg gacccaggcg acttcttggg aaactgtcat 2220
ctccgcactc ctccctgagc caggctcccg tcaagtgtgg tctcagccc acagagggga 2280
gggagcatgg ggcacagctg cagtcactct tgaagctagg gcgcctaccc cccaccccg 2340
aggacccctg cggcccccct cctagaggag caccatctac agttgtgcca ttcaccagcc 2400
actgccttca tctgtccccc gccgagctgg cagagccagg ggtcagccac ctgcctttga 2460
gtcatcaaga tgcctctgca gccacaatlc tgaacctaat ggcaggggcc agaaalctgt 2520
aaaaactccc gctgcctttt gtgatacttc ctgtgtctcc tcagagagaa acggagtgc 2580
cttttgtctt ttacctgatt ggcactctgc agtctatctc cctgggtagc agcgggtgc 2640
tgcctctctc tgggcatggt ctgaatgttt acaatgttac ctctgtgtat ctcttttga 2700
gccccctgca agctgcactc ctaggctttt atcttggggg gtcagagcgc cctctagagg 2760
gaaaagttag aggcacaggg ttctgcggg cccacaactg ctgtcttcat ttgcatttta 2820
cagcaagctg ctcaagacct ctagtgcct cctgcactct gactctcttc cccacaactc 2880
ccgtactcag ttctcttttt gctatactg aggcactgt gctgagggcc tgcagtgtct 2940
gctcactgct gccatcttct ctgtcagtca gggttccatc ctcttctccc tctccagtt 3000
ccctacacag ttggtatcca ttgtctccct atgtctaggg ccccaaaaga ctggggcagg 3060
ggccagacga gccgcaccca gtgtctccct ctctactctg acctggggcc ccagcatcct 3120
ggagcaccag ctccaccgac acacacccca gccctgtccc agggggcctg cccctcagc 3180
catctcaggg tgaaggagctg ccagtcatgt ccagtggaaa tgaatccat cctctctca 3240
tctccccctt gaggagcctc aaactgtcta gctcatcaaa gaggcaattg caactccgt 3300
atgtgttctt ggttccagg gagccttga accctggccc ctgggggtgt ttaattcatc 3360
attaagaagc altctctgtt ctcaagggac acagtgtgct gcatgggcca gcatggacc 3420
tgggtctgat atgtcatlc ctgtctctct ggggacacag tgggcccaca tgggcccaga 3480
tggacccctg gctagagcaa gcaactctcc atctcttcca cctcaggcag tgggtctca 3540
gatgtcagga gggactgacc tcaaggactt caaggttctc ctgtgcccag aatgagaggc 3600
caggcccgat cctacacact cgccttgacc ctgaagtcat agcaggccag ccaagcagga 3660
agcacactgt ctactttttg cctgaagaat aaatgtgtac ttgatagagc taanaataga 3720
tctctttt 3727

```

<210> 26
 <211> 2571
 <212> DNA

WO 02/40541

PCT/US01/46055

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 8128531CB1

<400> 26

```

ctaaagctgg acagaatttt taaagcaat gaagccagtt ccttgatat atccacgggc 60
tttgctttga gaaggaactg agtaggcagt gagaagagtc ggtgaagcc tggcccgta 120
gtgcttcaac aactgagatg aactgcgact cgttgcagg caagttgtca gggaggaagc 180
cacagcccaa gtaggtctgc acttgcgggg aaggtggctc gggccaggct gcaactcaaa 240
ccgtgtctct gtccacactg ctacggggcc agagccaaag aagcttccac ttcttccccc 300
agacagcccc aacagcggct accccaaggc gccagcagcc ttgtgtctg ggaaccacag 360
ccctgcaga atgacccacc aggatctgag catcacagcc aaactcatca atggagggtg 420
agcagggtct gtgggggtga cctgggtgtt ccccatcgac ttggccaagg ctggcctgaa 480
gaactcagcat gggaaagcca tgtaccaagg aatgatcgac tgcctgatga agcgggtcga 540
ggaggagggt ttctcaggga tgtaccagg ggtgagagtg aaactcactc tggcactcc 600
agagaaggcc atcaagctgg cggccaaagg ctttttccgg cggctgctca tggaaagatgg 660
gaacagcagg aactcgaaga tggagatgct tgcgggtgt ggggtggga tgtgccaggt 720
cgtggtgacc tgtcccatgg aatgctcaa gattcagctg caggatgctg gacgctggc 780
cgtccatcat cagggtctgg cctcagcacc ctccacctcc aggtcctaca caactggttc 840
ggcttccacc caagggcgcc cctctgcacc cctcattgcc tggagagatg tccgactca 900
gggctctggt gggctctaca gggcctcggg tgcactctc ctcagagaca ttctttctc 960
cactatctac ttccactagt ttgccaacct taacaacctg gggttcaacg agctcggcgg 1020
taaggcgtcc ttgacacatt ccttctgtgc aggtgtgtg gccgggttcc tagctggcgg 1080
cgaagtgcag cctctagatg ttctgaaan tcgaatccaa accctcaaga aagcctggg 1140
cgagagcatg taactggga taacgcagct tgcaggaaac ctctggatc aggagggacc 1200
atctgccttc atgaaggcg ctggctgccc ggcactgtc atagcaacct tctttgggat 1260
tgcctcaggc gtctatttta ttgggatggg agacgcctac ttaaaagtgt ttgactagac 1320
agagctggag gtcaagttcc ttgccttggc gccctctccc tagctgttc acttagctca 1380
gagggggcaa gggcaggtgg ggcactctg gctgcccgtg tctctgctg tgtagtgtca 1440
cctcaacttc gggagaaaca gccctatatt cttaaaggtt gaggacagcc ttcttccct 1500
tcgtgtctac actcgttttc ctttgtgggc acagctacca gggcttttg gaagccctca 1560
accactactc ttccacaaa aatggtactt tctgttactt aattgcaggc ccttaacagg 1620
tagtcaaac agaaaggttg ttctgtatt ttaacatttc tatttccag tcaactcgg 1680
cattcttcag tccagctgag gatttagcat tgttaacttc ggaactcata acttatgag 1740
cctagcaactg attttgaggg aaaggaggat cagaagtcca agggaccgtg aaagccctca 1800
gagtcagcac ctatgtttgag acaagacacc ctttggatcc cctggatggc tgaaggggct 1860
gaggccggct ctgactgggc agctcagccc ctcccacaga gccagggtc ttgcacacc 1920
ctcctgttaa ccaaggaaaca ctctgaata aaggtgaatg gctaaaatc catctgttc 1980
tcagtggtta cagcagatag gctgcagtga atgtatcac catctacttt tctactcca 2040
tttcaaaccc aancattaaa aaggccatag aagcagacc ccgtcactct tcaactggt 2100
acttgtgggg gtggaggaac acagccatag ggaalatct ccttgttagt gacactgggt 2160
tttaagcctt gattctatcc ctccataagt gaatcgtctt gaggagctga gtttgtgtg 2220
agagccctcc tcaagcctcc cgaatccccc cccaaggctt gctacaggag agataatgtc 2280
acagcagcag ggcagggtcc taagaaatc agcactcgtc caggagctg gtgtttaca 2340
tagtccctat tactgtgaaa cctgggttaa caaggagag gatggtgcta acatggtcag 2400
ccctgggggc ctcactctct gttatgagaa ctgcatttga gtatgggccc tggagacaga 2460
cctcagttca agtcccagct cccactatga ctactgcaca ggcctgggc actcttagt 2520
cgtaacctac ggaanantaa aacatgggac agggaaaggaa gaacagggcc t 2571

```

<210> 27

<211> 1660

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7476757CB1

<400> 27

```

ggagttctgg gctgtagtgc gctalgcaga tggggtgtcc gcactaagtt cggcatcaat 60
atggtgaact cccgggagcg ggggccacc aggtcgtctc tcaattctga caaaggtgtg 120

```

WO 02/0541

PCT/US01/46055

```

ctatggcatl ggtaggggtcc ccacacagab agcctccagc gccacagcct tttactctga 180
gcttttctgt ctgtatatac ccacagatcc tgccgccagc gtgtccattg ttctgttttg 240
gggaaagggtg ttggggggcg ctgtgtaccc tgtgctggg ttcttcatca acaggagcca 300
gaggacagga tctggacggc tcatgctctg ggtgctgggc tgacacccct tcatcgccct 360
ggcctacttc ttctctgggt tctgtccccc cttaaccagc ctggagggcc ttgtgtacac 420
gaatttctac tgcctgtttcc aggcctgggc cactgttttc caggtgcccc atcacagcgt 480
caccatctgt ctgactccct gcccaaggga gcgggactcg gccacgccca taccggatga 540
ctgtggagat ggccgggaaca ctgatggggc cactgttcca cgggctcabc gtgtcggagg 600
ccacagagac ccacaggtgc gaggccactg cgaacccggg gccagtcact gtctcccgca 660
atgcagccat ctctactgca ttgcggctgc cgtgggttga gtgacttacc cagtgtgcat 720
cagtttactg tgcctagggg tgaaggagcg gccaggtttt gottttgaac ttgtcggaag 780
caagtgaca cgtctctggc ttgcagaccc ctctgcccga gcttcaggcc caggttgag 840
tttcttggtc gggtcgagcc tcactaccgc gcccccaccc taactgaagc tgggtgatcc 900
cttctgttcc atctctgtgt ctgttcaggt ggagcagagc taactgggtc tgttctgtac 960
acatgcttcc cagctacacg accacgtcca gggcctggtc tccggcgtgc tgaagacccc 1020
gctctggagg tgggtttctcc agcgttttgg gaagaaagcg tcagcctttg ggtatcttgc 1080
gatgtgccc ttggcgatct tgcctggctg tctgcacaca gacactggcg catatgtgct 1140
ggcctttgta tctggcggtga gcaatgctgt gctcctgtgt ctacactggc caatgtgctc 1200
agcgtgggtg gatgacttcc agctgcagca cgttcacggg ccaggccttg agccatctt 1260
craactctcc taagtcttct tcaccaagct gtctggcgca tglgcccctg gcatctccac 1320
cctcagctct gactttctgt ggtataagcg aggggtctcg aagcaagcag aggggttgt 1380
ggtaacccct aaagtctcca ttggcgccgt gccacactgc atgatctctg ctgggctctg 1440
cactctcatg gtccgttcca ctccaaagac accagtcgga gacgcttcca gccggtgag 1500
ccttcaggag cgtgcacacg caccacaagt tcacacaagt aaggttccag agcatgcaca 1560
tatcatgcag gccacggcgg gacagccagt ggggtggcct gtcatcagcc actcctgct 1620
gaggttgacg gccctggggt ctgacagaga gagatactga 1680

```

<210> 28

<211> 2743

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 266243CB1

<400> 28

```

atggcgggcg cggcggtggg cggcgggcgc ggcggggggg gccggggggc agcgagcagc 60
agtgttgggg cggcgagagg cggcggggtg gggcgctgtg gctctgtgtg gtaacgctg 120
agcgcgggcg gcaacgtggc caacaagggt atcctgagcg ccttcccgtt ccgggtgacc 180
gtgtgctgtg gccacatcct ggctctgtgc gctgggctcc cgcgctgtct gcggcctgg 240
cgggtgcccc cggcgccgcc cgtctgtggc ccgggaccca gtccgcatcc gtctcgccg 300
ccgctgtctg cggcgccgctt ctaccgcgcc taactgtctc cgtctgctct cggcaggtac 360
ttcgcgctcg tgcagcgcca cgtcagcctc tggaggttgc cgtgttctca tgcacacacc 420
gtcaaggcca ccattgcccc ctgggtgggt cctctgtccc ggtatcttat gaaggagaag 480
caggagacca aggtatactt gtaactatc ccacatcaca ggggtgtcct gctggccacc 540
gtcacccagc tgccttttga catgtgggga ctgttccgcg cctctgcccc caagctgtgc 600
ttctcgcttc agaacatttt ctccaaaag gtcttgcgag attcacggat ccaactctc 660
cggctgtctc acatcctggg ctgcacgcgc gtcttcttta tcatccccac ctgggttctg 720
gtggacctct cggctttctt ggtcagcagc gacttgacct acgtctacca gtggccctgg 780
acgtctctgc tcttggtgtg cagcggtctc tgaactttg ccagaaagt tatcgcttc 840
agcatctcca acctcgctag ccgctgagc tactcggtcg ccaatgccc caaaagaaac 900
atggtcatca cgggttctct gatcatgctg cgaacccagc tcacacgac caacgtctg 960
ggcatgatga ccgcatcctt gggggtcttc ctctataaca agacaaagta cgtatgcaaa 1020
cagcaagcca ggaagcactt cctccccttc accacagcag acctgagcag caaggagcgt 1080
cacggagccc cactggagaa gcccccacac ggctctctct tccccagca cggggactat 1140
cagtaaggcc gcaacaacat cttaacagac caottccact acagcggcca gactatcca 1200
aaactcgtaaa gtttgaaccg ctatgatgtg tagagtcaca aggaacaggac cagactgttg 1260
gtgactcctt ccccgggccc ccacgagta tcagaaactt ctgacaatca gtyaatgtac 1320
aaacacagcg aggggaaggt gcaatactct ccatcagaag cctcggggtt cctggcccc 1380
cgtgagcccg agggagatgc gttgtctgca gtgcagagc ccgtgagctc tgggcaaac 1440
taaacagaga ccagtgcttc atgtctctc ttctggagc ctgtcatctg agggcgtgt 1500
ccttgaggag atcttgacca cgttgtacct ttccatgtg aattattccc caagcaggtg 1560

```


WO 02/40541

PCT/US01/46055

```

agctcagagg acttggtgtt gcatccaga taacattcag gacctgtgtg aaaagctggg 1620
gtcactgttg ctgtagaca tgaactggca gtgggggtgt ccaggagcgt gcttgagac 1680
gtcagactgt ctagttaa tccctgggca agatacgtat aggaacaca gggtaacca 1740
gcagacaggg agcccgagag aatcatcaaa aacatcccca gccacagaga tggatccagt 1800
ttcctgttca tcccttagc agttcacaa ttcttggcaa atgttccaa gcaaaaagcg 1860
atttgaatta gcatccagtt cctgcagcct ggtgtctctg cctgcacgtc aggggttgca 1920
tccaccaga tccagatgga agggaaactt cctctctctc ctttgctctc tcttccctca 1980
ccagagcagg ggccttctct tgggtgtgtg agaaggatct tcagagaatc gtttccagta 2040
ttccagctc tattctctgt gcacatgtct tttagagagc atcttccact cttctgtgat 2100
gacttggtat gttgttttgt agagagatct tgaattttcg aggatcttgc attttcttag 2160
ggaaattttt gtatgttgtt gtgtgtgttt ttgccttggt cccattatg ggtatgatta 2220
ggactggcct atgcattgaa aatctttttt ttgtaaaag tttaaaaca aagtccccc 2280
gcaaggcaca gtgctcaca cctgtagtcc cggcactttg ggaggccaa gttggcggat 2340
cacagggtca ggagttcgag accagcctgg ccaacatggt gagggccctg ctctactagg 2400
agtaacaaaa ttacccaggc atgggtctgc gtgctgtgtg tccagctcc tggggctgct 2460
caggcaggcg aatgtcttga acctggaggt gcatgtatgc gacctgggt gactgcaccc 2520
tctactccc gggttcaaac aattgtcttg tctcagctc ccaagtatct gggatccag 2580
gtgtgcacca ccatgctggg ctaattttta gttagagatg ggtttcaca tgttgccctg 2640
gatggctctg aactctgac ctcatgatg ccaactgctc tggctccca aagtgtctgg 2700
attacaggca tgagccaca cacaaccgac ctggccagc aca 2743

```

<210> 29
 <211> 3239
 <212> DNA
 <213> Homo sapiens

 <220>
 <221> misc_feature
 <223> Incyte ID No: 6585710CB1

```

<400> 29
ggcttaagt agggcgggc agccatggt ccttatbtt accgcatcg gctacccctg 60
tcttcgtac agcaatcctg ctgacttcta tgtggacctg accagcattg acaggcgag 120
cagagagcag gaattggcca ccaggggaaa ggtccagcta ctcgagccc tgtttctaga 180
aaaagtgcgt gacttagatg aotttctatg gaaagcagag acgaaggatc ttgacgagga 240
caactgtgtg gaaagcagcg tgaccccatc agacacccaa tgcctccaga gtctacgaa 300
gatgcctggg gcggtgcagc agtttaagac gctgatccgt cgtcagattt ccaacgactt 360
ccgagacctg cccacccctc tcactcatgg ggcggagggc tgcctgatgt caatgacct 420
cggcttctc tatttllggc atgggagcat ccagctctcc ttcattggata cagccgccct 480
cttgttcatg ateggtctc tcactccttt caactctatt ctggatgta tctccaaatg 540
ttactcagag agggcaatgc tttactatga actggaagac gggctgtaca ccactggctc 600
atatctcttt gcaagagatc tggggaggtc tccggagcac tgtgctaca tcactatcta 660
cgggagctcc acctactggc tggcaacct gaggccaggc ctccagcctc tctgctgca 720
cttctgtgtg gtgtgctgtg tggctctctg ttgcagatct atggccttg ccgcccgggc 780
cctgtctccc accttcaca tggactcctt ctccagcaat gccctcaca actccttcta 840
cctgcgcggg ggccttatga taaacttgag cagcctgtgg aagtgcccg cgtggatttc 900
caaatgtctc ttcctcggtt ggtgttttga agggctgatg aagattcagt tcaggcaga 960
aactataaaa atgcctctcg ggaacctcac catcgcggtc tcaggagata aaatcctcag 1020
tgccatggag ctggactcgt accctctcta cgcactctac ctcatcgtca ttggcctcag 1080
cgtgtgtctc atgttctgt actacgtgto cttaaggttc atcaaacaga aaccaagta 1140
agactgttga ttaacccag agtgtgtccc gctggtgggg gacctgagca gaccttcaa 1200
ctgactccc tctcaggag ccccttctcg gggacagtga ggacaatgac cctcagatg 1260
ctcagctaca tccggccag ggtgtctcag tggcacagac cagccacagg atggcagtag 1320
aataaagaca gtccaaaggg atttctgctc actggcagga gactgcgatg actggagaaa 1380
aacctgcact cgttggcacc tacaacgttg ctaattttat tctttttgat atgcatttat 1440
ataggtaact ccatatagga tggagcaaaa ctaggaaatga attggttagc tagactgtgc 1500
aggaattgtt ggaacctgga ggaacacaaa atagtaagta atgagtttag ctcaactc 1560
caggggcccc acctccgtg gttagcctac atcaatacag aaagtgaact aagatgtacc 1620
agcaaatgac catcccttct ttttgtgtgg ggtcatgggc tccaaaagcc aactgtgaac 1680
attaaaatgt tatctgagat ctactctgta gtaggtctcg tgaaaacct ttaggtggac 1740
aatcccttga ggtcagtggt atcccatctt ataggtgtga aaactgaagc aaaaattcat 1800
tttctcaagg gcacatggat acttgtgtgt gactcatatg gggatcagaa aagccttga 1860
ggccttgag tttaggggca gaaggcaagg cctgagccgc tgaagccct taggagttta 1920

```

WO 02/40541

PCT/US01/46055

```

ggagagctcc agaaacacaa tggggctctg agaggctgtt aactcagcca ggtctcttag 1980
agtctgcatl cactactctg tatggcttgg ctctggctcc tcaaccaak ctcaacttga 2040
atgttaataa tcccaagtg tcaagggggg gaccagatgg agataahtga atcaatagggg 2100
tggtttctct gatctgttcc tccctgagct gactgagttc tgatgagatc cgcaggtttt 2160
ataaggggct tccctcttgg ctgggtcttc attctctctc ccgtaccct gtgaagagga 2220
gccttccacc acgctctcaa gtttctctgag gctgcccacc ccgtgctgaa ctgtgagtaa 2280
gttaaacctc tttcttttat aaattaacca gctctgggta tttcttosta gcaagtgtgag 2340
agcagatgaa lacactggcc ctgcctgggt ttcagaacca gocttgaacc tttcacagtg 2400
gccagaggat ggggagggcag aggcaccaagt tgcacacttc ttgcttgagt gttggggact 2460
attctgacca aaacaggtgc acagagggcca ggagcgggat ttcccaagg aaaaattagc 2520
tttgaataca aaagaagggg aaggtgogtg tttgggaggt aaatagcaaa tactcttcac 2580
aggcttcaata gctcttggct acaccaagta cgtaccctgg gccagcagaa tgggcacago 2640
tggagctgat tggataggtc ccattagcct caggccccac ccaggctcaa tgagtccag 2700
tctggctctt aagaagaacc cctgggtgac tgtgcacatt caagcatgct gcccttltcc 2760
aaagactctg caacactcag gatgcttgca oggtcactgt gccacatcc aacctgcaga 2820
cccaactctt ggaattgact gggagttctt abcatgtctc ccatagcaag ggggtctaga 2880
ccagactcaa gacttgagtc tagttctcaa gtctcttctg ctcttctagt ttaggcaacg 2940
tttctcaact tctctcaact ttgtgacctt gatacttgag ttgaaggct gtctctcaat 3000
ttgtgtctct ccagtgcat cctatcagaa ggcattgtat ttcaacttct cccatccaa 3060
caagttctac ttgtatcact tgattaaagg ggtgtctgct aggtctctcc acagccagat 3120
tactatttct cctcccttta taattaataa gcatkttgta agtgggtact ttgaactat 3180
gtaaatgtta aacttccat ttatgcattb taaaattttg attgatgtaa aaaaaaaa 3239

```

<210> 30
 <211> 1615
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> Incyte ID No: 7493599C81

```

<400> 30
atggcacaat tctagacac atacaactta ccaagattga accaggaaga aatccaaaac 60
ctgaagagac caataacaag taatgagatc aaagcaatan taaaaagtc ccagatgtca 120
ttgcttggaa gggactacaa cagtgcagtg aactctcttg acaacggacc tcaagtccac 180
tcagagagca gaagtagcat tacttcagag aatgtccacc ctgctggaga agctggacta 240
tcgatgatgc aaactttgat ccaacttggt aaatgcacca ttggcaccgg gctcctgggg 300
cttccctctg ccataaagaa tgcgggcttg ttggctggtc ctgtcagcct tctggccatc 360
ggggctctca ccgtgcactg catgttcact ctgttgaaat gtgtcacaac cctcagccag 420
cctagactgc agaagacttt tgtgaactat ggagaggcca cgtatgacgg ccttgaaacc 480
tgcccgaaca cctggctgag ggcccactga gtgtggggaa ggtacactgt cagctcttta 540
ttagtcacca ccaagctggg ctctgcagat gtttatatta ttttctcggc agacatttta 600
cacacagatg tggaaaagg ccaagtgacc tccaaactat gccagccag ggaattctg 660
acgttgaccc caactctgga cacttgtttc taactgtctga taactctgce ctctctgac 720
ctgttgggtg ttatccagaa cctcaaggtg ctgtccgtct tctgacatb ggccaacato 780
accacccctg gtagcactgg tctgactctt gactatatac tggaggggat tccatctccc 840
agcaaccctac ccttgatggc aaactgggaag acctctctgc tgttctttgg taacgcaatc 900
ttcacatttg aaggcgtcgg katgttctct cctc:caaaa accagatgaa gcatccacag 960
cagttttctt ttgttctgta cttggggatg tccattgtca tcaactctta tatcttactg 1020
gggacaactg gctacatgaa gtttgggtca gacacccagg ccagcatcac cctcaacttg 1080
cccaattgct ggtatgtcct gccaccccca ggtgagatag ggagagacac tggnaactgt 1140
ctggttgtca tagcagagag caccagaaag ctgagccatg aagctggtaa tccatccactg 1200
gaagtgcact atgtctctcc tgcctcaact gcatcagtea aagcaagcca catggccgca 1260
cctcaactcca agggggcagg gaagtgcact tctgcatgtt gcttggaagt atttgggtga 1320
cagcacaact aactgtgtgt ctactgagtg aagtgcaggt ggggattaat tggatttaata 1380
catagataat gcttagaaaa gtgtctagca ctgtgtgaag attcatgagc gttagctata 1440
atcaattgta tgaatcccca cagccttcat ttttccaaag taagtggat gatgtgcat 1500
ttatctccca caastocaga gctgtagaat gagaaaaaat taaccatccc ccccacctt 1560
gctgtgttat gataatgact agatgagaca ataaatgtgg agtttctctg aaaaaa 1615

```

<210> 31
 <211> 1245

WO 02/40541

PCT/US01/46055

<212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2507246CB1

<400> 31
 atggcgacgg ggggcacgca gaaggagaa acgtgtcttc acctcttcgc gggcggtgt 60
 ggggcacgag ttggtgctat tttaacttgt cactagaag tcaatlaagac acggttgag 120
 tobtcaagat tagctctccg gacgtctac tatctctagg ttcatctggg gaccattagt 180
 gggagctggaa tggtagagac aacatccgtg acacctggac tctttcaggt tctgaagtcg 240
 atcttggaga aagaggagac aaagtcaatt tttagaggc tgggtccaaa ttgggttggg 300
 gttgcacat caagggtgt atactttgca tttatctcca aagccaaaga gcaatttaat 360
 ggcattttcg tgcctaacgg caatattgtg catattttct cagctggctc tgcagctttt 420
 atoccaaat ccttaatgaa tctataggg atggtaaaa cccgaatgca gctagaacag 480
 aaagtgaagg gcttaagca gatgaataga ctccagtgtg ctccgttaag ttacagagac 540
 gaaggcattc gtggtctta tagaggatta actgctcgt atgttgaat ttccgaaact 600
 acaactcgtc ttgctattta tgaagttta aagaagtatc tgaagaagc tccattagcc 660
 totttctgca atgggactga gaaaaattcc acaagtttt ttgacttat ggcagctgct 720
 gctcttctta agggctgtgc ctctgcaat gcttatccac acgaagtcac aaggacgagg 780
 ctccgggaag agggcaccna gtacaagtct ttgtccaga cggcgcgct ggtgtccgg 840
 gaagaaggct accttgctt ttatagagg ctgtttgcc agcttatccg gcagatccca 900
 aatctgcca ttgtttgtc tactatag ttaattgtgt acctgttag agacgctact 960
 cagtaacagg ccggaattt gtgctcaga agaataaac tgaataactc tagagaattt 1020
 tttttccca ttgatgttta gaagtttga gactgaaca gaaaggcca taaaatctc 1080
 ggttatcat acctgttga catctcttt tggattcat cttcttgaa ggtttaaatt 1140
 catcaagctt aatggtta tataacttt ttttaactt aagaggatc aggggttaagc 1200
 accaactaaa ttaaatcatg ctatttaatt taagtataa aaaaa 1245

<210> 32
 <211> 4169
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3033505CB1

<400> 32
 gcaaggctca ctgacacgca gctttggtta aagagcgggc gcacaggagg gaggagagac 60
 gcgcgcggga cggggaggaa tggcctgtcc gcgttaaac atcccaagcc atggttgcgg 120
 aaggggccagc cgtcccccag taggagaatg actccgattc gtgacccca cggcgggtgc 180
 atgtcgatat atttatggag tctctactgt gtgcccagca ctatattat gtgcacagaa 240
 aaaccttga agccataga acatataata tagagtgatc gtctttgctt gctgaagtaa 300
 cagggtgtgt aagcttccat ttgggtatct ctttctaact acactcaga caggagaaat 360
 ttgactaat ttccaacta cagacacttt ctatcatga tgcatttcaa aagtggactc 420
 gaaltacttg agttgcaaaa catgacagt cccgaggatg ataacattag caatgactcc 480
 aatgatattc ccgaagtga aatgggtcag ataatagca agtttatctc tgatcgtgaa 540
 agtagaagaa gtctcacaaa cagccatttg gaaaaaaga agtgtgatga gtatattcca 600
 ggtaacaact ccttaggcat gtctgttttt aaactaagca acgcaattat gggcagtggt 660
 attttgggac tgccttttgc cctggcaaac actggaatcc tactttttct ggtacttttg 720
 acttcagtga cttgtgtgtc tatatatcca ataaacctcc taltgatctg ttcaaaaga 780
 acaggctgca tgggtgatga aaagctgggg gaacaagtct ttggaaccac agggaaagttc 840
 gtaacttttg gagccacctc tctacagaa actggagcaa tctgageta cctcttctc 900
 gtanaaaatg aactacctc tgcataaag ttctaatyg gaaaggaa gaacatttca 960
 gcctggtaag tggatggcgg cgtttctgtg gtgatgtta cctttggcat aattctccct 1020
 ctgtctctct tgaagaactt aggttatctt ggtataacta gtgatttttc ctgtgctgt 1080
 atggtttttt tctcaattgt ggttatctac aagaatttcc aatttccctg catgtttcca 1140
 gagctaaatt caacaataag tgcatactca acaaatgctg acacgtgtac gccaaaatat 1200
 gttaccttca atccaaagac cgtgtatgct ttacccacca ttgcatttgc attgtttgc 1260
 caaccgtcag tcttgcaaat ttacagtgag cttaagagac gatccagaa aaaaatgcag 1320
 atggtttcaa acatctctt ttgcgcagc ttgttatgt acttctgac tgcattttt 1380

WO 02/40541

PCT/US01/46055

```

ggctacttgc cattctatga caacgtgcag tccgacctcc ttccaaata ccagagtaaa 1440
ggtgacattc tcatctcgac agtcggctg gctgtcattg ttgctgtgat cctcacagtg 1500
ccggtgttat tttccacggt tggttcatct ttakttgac ttgctataga aacaaagttt 1560
aatattatgt gtcataccgt gggtacccgt atactcttgg ttgttatcaa ctgtttgggt 1620
atcttcatac cctccatgaa ggatattttt ggagtcgtag gagttacac ttgctaactg 1680
cttatttttca ttcttccttc atctctttat ttacaaatca cagaccagga tggagatana 1740
ggaactcaaa gaatttgggc tgcccttttc ttgggcttgg gggtgttgtt ctcttgggtt 1800
agcattccct tgggtatcta tgaactgggc tgcctatcga gtatggacca aggcacttga 1860
aacccgcgga gaaaaagaaa catccctgtt gtotgtccag tcaagtcccc aacatccagc 1920
aatctctcac cacttctttt gcaagtttac agaagcaaac agaatgttac aggtacttca 1980
aaatggataa acttttttgt tgcacaaacg agacatgggt ctataatgct tcatgtccct 2040
ccagatattg agatcaattt agggatgtgt aaattttttt ttccaattt catataatca 2100
tattttccag taactttcac aatcaatttt taccatcta actotatgtt ttgtggcttc 2160
ccggtctctt agaactttga aaacatgata tacaataatg ttatattatt ctacatccag 2220
atcttgaatc aaatttccca ctgaggttca gctcacaca tctgtacct tttagagaga 2280
aaagaaact tgaattgcat atatttattt tgccttaccg aaaaaaattg ttctgttaat 2340
aatttgctca ttctggttaa catagacatc ggaatatac atctgaagt tataggcac 2400
tgccactgct gaalccagac atgcccaata tttaggtgtg ctctgatttc ctggcagctg 2460
aactcgggtg gtccagtggc ctactgtgta cccatctat tccatccag agacattctc 2520
tgccaaagtgt tctcagctga aaagtgttg gggatgattc ttacttgggt aattaaatga 2580
agctacacat ttgggtaatc tagcaaatga agtattttt cctcttggc aacttgtctc 2640
aggttactc ttgtctgagt caactttgc tggggaacac ctatggaacc tactgcaaaa 2700
agattgtcca aaatgctcaa gaaaatactc ctctgatgca tttagccttc aacctactc 2760
gtttgtctga aggggaaaa atgttttagt acatttatag cccagcagct ttattccatg 2820
tccacagctg agtgcacag agaatcatgt gtacctaac aaggatgac taggataagt 2880
aactcctgtt ttatattgag tattttaggg aagtctttaa aagacttgtt ttatatctar 2940
aaatctaggc tattcaaatc aaaggaattt tgcaccttaa atagcctca ttctatttc 3000
ttcttcatta atctctcact tagtcttctg aaaaaaaa aaaaaaac ctccagagata 3060
gtcttctgta agagcttctg acagatccag tgaactcct ctctcccca ctgagagag 3120
acaaaggggt ctacgtgtct gtctgtctc ctctctctc ccccaaccaa ggaactgtcc 3180
attactccc gtctcaactg tccatgcagg aggcacagct tgcctgttac tcttacctt 3240
gtctctctcc taaaggagac acaaggaaac tgaagagact gaaaaagag agagtttgta 3300
gctgaaaaag aatagggata gaaaggaaac ccagaaatgc attccctaa gtggggccat 3360
ccatctgtat tgaattgtcc atagcttgc ttatgttga aatgtgcat ctccgtgagc 3420
tggctctctt aaacaggact tatgtctctc ctatatctg gttaaattt ccaaacacat 3480
aagtctactg agcacagatt tcttatccag agacaagtat aatctaacg cagactgttg 3540
gcagagtttc caggccacta gccatgttcc ctctctgact caatcccca aaggccttca 3600
ctctcactga gaatcacact actgtcccat agataaggca ggcattgaag cactgttgt 3660
gatctcttag gggggagaat gaaaggttat ttctgtcatt gcatcatcat agcttttaat 3720
ataatgctac gaatcatat ccaatttag ttatagttca gatatttga tatgaatac 3780
tazctagccc atatccatgg caactctgtt tcttttcagc aatgttttcc atattatatt 3840
agcaatgaca gaaacagaac aagccagat ccagtcagtt ctgtggagct tgtctagagc 3900
accaagtat gaaatagcca ggtagtggga tgaactgac tttaaaata catatcttag 3960
tttgaagct atattatgt actttctatt ttctctgtta ctttataga attcatttta 4020
ccctccaaa gtcaatttag aaccttatca ttaactggga tgtgtatgca tatttttggg 4080
cctctgggtt tcatgtgtca ataccaggca tctctcttcc aaatagattt atttagaggg 4140
ggccagtgtt gttgactgtg tggaaacccc 4169

```

<210> 33
 <211> 3440
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4027693CB1

```

<400> 33
gatccccacc adaccaccag cccggccgca cggggcactg agccgggtgc tgaaccacgg 60
aggcccccgc gaggccggga ctacggacct gcagagaaac gctccttgat ttgtcttac 120
aatggaaactt aaaaagtgc ctgaaggtgg atggggctgg gtgatttgtt tgtctccctt 180
ccttactcag tttttgtgtt accgatcccc actagctgtt ggagtcctat acatagaagt 240
gttgatgtgc ttgtgtgaag gaaaaggaaa aacagccctg gttggatccc tggcaagtgg 300

```

WO 02/40541

PCT/US01/46055

```

agttgggttg  cttacaagtc  ctgtctgcag  tctctgtgtc  tcatcttttg  gagcaagacc  360
tgcccaaatc  ttcaatgggt  tcaatggggc  tggagggctg  atgttgagca  gttttgtccc  420
caaatatcac  ttctgttttt  ttctctatgg  catgttggtg  ggtcttggat  gtggtttatt  480
atacaatgca  acagtggaca  ttacgtggca  gtattttgac  gatgcggag  gcttagcgct  540
tggcctgatt  tcaacagggt  caagcgttgg  ccttttcata  tatgctgctc  tgcagaggat  600
gctggttgag  ttctatggag  tggatggatg  ctgtgctgatt  gtggttgctt  tagctttaaa  660
tatattagcc  tgtggcagtc  tgatgagacc  cctccaatct  tctgattgtc  ctttgccctaa  720
aaaaatagct  ccagaagatc  taccagataa  atactccatt  tacaatgaaa  aaggaaagaa  780
tctggagaga  aacataaaca  ttcttgacaa  gagctacagt  agtgaggaaa  aatgcaggat  840
cacgttagcc  aatggtgact  ggaacacaga  cagcctactt  cataaaaaac  ccacagtgac  900
acacacaaaa  gagcctgaaa  cgtacaaaaa  gaaagttgca  gaacagacat  atttttgcaa  960
acagcttgcc  aagaggaagt  ggcagttata  taaaaactac  tgtggtgaaa  ctgtggtctt  1020
ttttaaaaac  aaagtatttt  cagccctttt  cattgctatc  ttactctttg  acatcggagg  1080
gtttccacct  tcaattactta  tggagagatg  agaagaagt  tcaaacgtga  aagaagaaga  1140
gtttattatg  ccaacttttt  ccattatagg  cattatgaca  gcaatttgta  aactgctttt  1200
agggatcttg  gctgacttca  agtggattaa  taccctgtat  ctttatgttg  ctacccctaa  1260
catcatgggc  ctacgctctg  gtggcaatcc  atttgccaaa  agctatgtca  catcggcgtt  1320
gctttctagg  atcctagggt  ttctactagg  taattggtcc  atcttccct  atgtgaccac  1380
gaagactgtg  ggaattgaaa  aattagccca  tgcctatggg  atattaatgt  tctttgtggt  1440
acttggaaat  agcctaggac  caccatcgtt  tggttggttt  tatgactgga  cccagacctt  1500
tgatattgca  ttttatttta  gtggctctgt  cgtctctgtg  ggaagtttta  ttctgctgct  1560
ggagagcctt  cctcttgggt  atecatgcaa  caagcaactc  cccagaccag  ctccaacaa  1620
tttcttgtac  aaagtgtcct  ctactgttta  gaagaatttt  ggaagacctt  atttttgtca  1680
ttttatacca  tatagcaaac  atattttaac  agattctcaa  gcaaatlttc  tagagtcagg  1740
actattttct  catagcaaaa  ttccacaagt  actgactctg  aatgaattat  ttttttttat  1800
atatcctatt  ttttatgtag  tgtatccgta  gccttatctt  cgtatttttt  tctattttct  1860
ctccccacac  catcaatggg  actattctgt  ttgtctgtta  tcaactagtt  cttaacattg  1920
tcaaaaattt  gccacgcttc  agaaggtttt  ctctgtgtta  agaagtataa  ttctctgccc  1980
gaotcatttt  aatccactgc  aaggcaccta  gagagactgc  tcttatttta  aaagtgtgct  2040
aagcatcatt  atagacatgt  tctgaagccc  actgggaatt  aatctattct  ctctctatgt  2100
tttgacttgc  tagtaaacag  aagacttcaa  gccagcgg  aaattaaagt  ggcgaactaa  2160
acagccttaa  gaattgcagt  ggaacaaatt  ggtcattttt  taaaaaata  tattttaacc  2220
taocgttaac  agtttttaatt  attctattta  cctcactgaa  gtactcgcat  gttgtttggt  2280
accocctgag  caactgtttc  agttcttaag  gtatttgtgt  agatgtgggt  gaactccaaa  2340
tggagaagta  gtcaactgtg  actttcttca  tggttgacca  ctccacatt  gctcactttt  2400
gcttctgggc  catccactca  gctgagtgtt  cctgggaagt  gctaakttta  cctgtttcca  2460
aattggaac  acatttctca  atcattctgt  tctggcaatt  gggaaactc  catttgtctt  2520
ggggaacagt  gggatggggt  gcaagtcttt  gcatatcttc  ccagtgaagc  atttatttgc  2580
tactatcaga  tttttaccct  atcaaatata  attcaagggc  agaattaaac  gtgagtgtgt  2640
gtgtgtgtgt  gtgtgtgtgt  gtgtgtctat  catgctctaa  gtctgcattg  gatatggaaa  2700
tggaaaaggg  caataagaaa  ttaataccct  tatgcagttg  catttaacct  taagaaaaat  2760
gtccttggga  taaactccaa  tctttaatac  attgattttt  tttctaaaga  aatggtgttt  2820
aaactttggt  atgcatacga  atccctcata  gatctttttg  aaaatatagg  taacttggta  2880
tcaacataag  aacttttaac  tctgtgtgtg  tagctgtgtg  cccaacac  taacttttta  2940
ctgagctctt  caagtgtatt  tgataacaaa  gcttgagattg  agaattttta  taagatttgg  3000
aatggaaaaa  catttattct  tttaataaat  aattttttta  aaacccaaga  ggtcagggga  3060
ttttataaac  caatagccaa  gtgtctctta  aatagggggc  acccttccca  ttgtgccaaa  3120
atcatctttt  catttatttt  gaaatttgta  tgattatttt  atacttggat  gttgcttttc  3180
ttcgaaggcg  cctgaagcac  tttaaaaaa  caaatctcca  caatacctct  gtgaggtagg  3240
taaatagtac  tttttatgt  agtaaacctg  gaatatggag  aatttcataa  cagttcattc  3300
tacttaataa  tgcaataaat  gagtccaaag  ttgtcttggg  cttctacacc  acactcagac  3360
ttctggaaag  ttctctgtac  ctcaattctt  agtccctgtc  aaggttagta  aataaaaaaa  3420
gtgcataaaa  aaaaaaaaaa  3440

```

<210> 34
 <211> 3699
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7472030CB1

WO 02/40541

PCT/US01/46055

```

<400> 34
atggtttatt ccggaatg agagatgttt aacattcaaa aatcaactgc tctaaact 60
gcggaagaa acgcaaaact gagaagga gacgttggat ctattgagat attccgttt 120
gctgsktgac tggacatcac actcatgato ctgggtatad tgacatcaact gttcaatgga 180
gcctgccttc ctttaatgcc actgtgtala ggagaatga gtgataaact tattagtga 240
tgctgtgtcc acactaacac aacaaattat cagaactgta ctgagtctca agagaagctg 300
aatgaagata tgactctgtt gacccgttat tatgttggaa taggtgttgc tgccctgatt 360
tttggttaca tacagatttc ctgtgtgatt ataactgcag cagcagacag caagagagatt 420
cgaaaacagt tttttcattc agttttggca caggacatcg gctggtttga tagctgtgac 480
atcggtgaac ttaacactcg catgacagat gacattgaca aaatcaagtga tggatgtgga 540
gataagattg ctctgtgtgt tcaaaacatg tctacttttt cgattggcct ggacgttgg 600
ttggtgaagg gctggaactc cactctagt actctatcca cgtctcctct tataatggct 660
tcagcgccag catgttctag gatgtctatc tcatgacca gtaaggaaatt aagtgcctat 720
tccaaagctg gggctgtggc agaaagagtc ttgtctcaaa tccgaacagt catagccttt 780
agggccaaag agaaagaact tcaaaaggtat acacagaatc tcaagatgc aaaggatttt 840
ggcctaaaaa ggcctatagc tcaaaagtg tctcttggtg ctgtgtactt cttatgaa 900
ggaaaccttg gacttgcttt ttgtgtatga acctccttga ttctaatgg agaaacctga 960
tatacaatcg ggaactgtct tgcgtgtttc ttaagtgtaa tccatagtag ttattgcatt 1020
ggagcagcag tccctcaact tgaacacttc gcaatgccc gaggagctgc ctttcatatt 1080
ttccaggtta ttgataagaa acccagttata ggttaactttt cccagctgg atataaact 1140
gaatccatag aaggacatgt ggaatttaaa aatgtttctt tcaattatcc atcaaggcca 1200
tctatcaaga ttctgaagg tctgaatctc ggaattaaat ctggagagac agtgccttg 1260
ctggtctcca atggcagtg gaaagatcag gtatgcacg ttctgcagag gttatattgat 1320
ccgcatgatg gctttatcat ggtggatgag aatgacatca gagctttaa tgtggcga 1380
tctcgagacc atattggagt ggttagtcaa gacgctgttt tglcgggac caccatcagt 1440
aacaatatac agtatggagc agatgatgtg actgatgaag agatggagag agcagcaagg 1500
gaagcaaatg cgtatgtttt tcatatggag ttctcaata aatttaatat attggtagg 1560
gaaaagggag ctcaaatgag tggagggcag aacacagaga tggcaattgc tctgtcctta 1620
gttcgaacac ccaagattct gatcttagat gaggtacagt ctgacctga tccagaagc 1680
aagtcagctg tccaagctg actggaag actggacaaag gtcggactac aatcgtgga 1740
gcaacacgag ttctactat tccaagtgca gatttgattg tgacctaaa ggaatggaatg 1800
ctggcggaag aaggagcgca tgcgtacta atggcacaac gaggctata ttattcaact 1860
ctgatgtcac aggalatlaa aaaagctgat gaacagatgg agtcaatgac atatttcaact 1920
gaaagaaaga ccaactcact tctctgcac tctgtgaaga gcatcaagtc agacttcaat 1980
gacaaggctg aggaatccac ccaatctaaa gagataagtc ttctgaagt ctctctatta 2040
aaaattttaa agttaaacaa gctgaatgg ccttttgttg ttctgggagc atttggtct 2100
gttctaaatg gaactgttca tccagtattt tccatctatc ttgcacaaat tataacctg 2160
tttggaata atgataaaac cactataaag catgatcgag aaatttatto catgatattc 2220
gtcatttttg gtgtattttg ctttgtcagt tatttcatgc aggalattgc ctggtttgat 2280
gaaaaggaaa acagcacagg aggtctgaca acaattatag ccatagatat agcaacaatt 2340
caaggagcaa caggttccag gattggcgtc ttaacacana atgcaactaa catgggactt 2400
tcagttatca tttcctttat atcggatgg gagatgacat tctgatctct gattattgct 2460
ccagtaactg ccttgacagg aatgatgaa accgacgaa tgaatggatt tgcacacaa 2520
gataagcag aacttaagca tgcgtggaag atagcaactg aagctttgga gaatacagt 2580
actatagtg cattaaacag ggaaaaagcc ttccgagaaa tglatgaag gatgcttcag 2640
actcaaaoca gaaatctctc gaagaagca oagattattg gaagctgta tgcattcagc 2700
catgocctta tatattttgc ctatggcgca gggtttogat ttggagccta tttaattcaa 2760
gctggacgaa tgaccccgag gggcattgtt atagttttta ctgcaattgc atatggagct 2820
atggccatcg gagaaacgct cgtttttggc cctgaattat ccaagcccaa atcgggggct 2880
ggcactctgt ctgcttgtt ggaagaagaa ccaaatatag acagccgag tcaagaggg 2940
aaaaagccag acacatgtga agggaaattt gagtttcag aagtctctt cttctatcca 3000
tgtgcgccag atgttttcat cctcctgggc ttatccctca gtattgagc aggaagaca 3060
gtagcatttg tggggagcag cggctgtggg aaaagcactt ctgttcaact tctgcagaga 3120
ctttatgacc cctgcaagg acaagtcgt ttgatggg tggatgcaaa agaattgaat 3180
gtcagtggtc tccgttccca aatagcaatc gttcctcaag agcctgtgct cttcaactgc 3240
agcattgctg agaacatcgc ctatgtgac aacagccgtg tgggtccatt agatgagatc 3300
aagaagccg caatgagag aatcatcaat ttttttttga aaggtctccc tgagaataa 3360
aacaacacag ttgactgaa aggaacacag cttttgtggc gccagaaaca aagactagct 3420
altgcaaggg cctttctoca aaaaaccaaa attttattgt tgaalagggc cacttcagcc 3480
ctcgataatg acagtgaaga ggtgttctag catgcodttg ataaagccag gacgggagg 3540
acatgcctag tgtcatoca caggctctct gcaattcaga aogcagatt gatagtgtt 3600
ctgcacactg gaaagataaa ggaacaggga actcatcaga agctcctgag aaatcgagac 3660
atatatttta agttagtga tgcacagtca gtgcagtga 3699

```

WO 02/40541 PCT/US01/46055

<210> 35
 <211> 2428
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7476089CE1

<400> 35
 atagatccctg aaaaggaac gactgatctc accatcaagc tagtgatcat ccatatggct 60
 tgcctgcagtt ctccacaggc ctgcctctctc agcctaagga cgcctgaccc tgcctgcag 120
 cagcgttcca catatcactg tggaggagga agatggagaa atcagggttat ggtcatccgt 180
 gcaacaggga cttctgggaa gggtagctgc ggaatttaga acagtgctct tgcacgatt 240
 cagtcctgag gattaccaga atgttgcctgg cacttagaa ttccaccag gageaagata 300
 taatcacatt ttataaaca tcaactgata tctattctc gaactggaaa aaattcttaa 360
 agttggagtg ttaaccltgg aagggagggc aggtotagga ttggtctcc aaattctagt 420
 gaaattgga gacctgac acgtctatgg cgtatttga tttagccctg agtcaacttt 480
 ttgcagtgga aotgaaccag aagatgggta tagcaactgt acattaatg ttataagaca 540
 tcatggaaat ctgtctccag tgactttgca ttggaacata gactctgac ctgattggtg 600
 tctgcctctc acctctggca acatccatt tgagattggg cagcagagcg ccaatatcac 660
 tgggagata ttgcctgagc aagaccaga actggataag gcaattctct tgcagttct 720
 cagttcttcc agtggttctt tgggagctca tattaatgcc acgttaacag ttttggctag 780
 tggatgatcca tatgggatat tcattttttc tgagaaaaac agacctgtta aagttgagga 840
 agcaacccag aacatcacac tatcataat aaggttgaaa ggcctctg gaaaagtctt 900
 tgtctctat gcaacactag atgatatga aaaaccaact talttctcac ctaatttagc 960
 gaggagcaat caaggaagag actatatacc agcttctgga tttgctctt ttgagctaa 1020
 tcaagtgtag gcaacaatag ctatttcaat ttggatgat gatgagccag aaaggtccga 1080
 atctgtcttt atagaactac tcaactctac tttagttagc aagtagcaga gtcgttcaat 1140
 tccaaattct ccaactctg ggcctaaggt agaaactatt ggcgaactaa ttactcttgc 1200
 caatgatgat gcaatttga cttctcagct ctccagacca attgtccag tggcagaaa 1260
 tcaatttga ccaattatca atgtgactag aacaggagga gcaattgtag atgtctctg 1320
 gaagttttaa cctgtgcaaa taactgcaat agctgttga gattataga tagctctc 1380
 agatgttgct ttgttagaag gggcaaccag tgaagccgt ccaatatag tcaattatg 1440
 tatctatctc gaactggaag aatctttctc tctgcaactg atgaatgaaa caacaggagg 1500
 agccagacta ggggttttaa caggagcagc cactattatt gaggctctg atgacctata 1560
 tggattatct ggttttcaga ttactaaact tattgtagag gaacctgagt ttaactcagt 1620
 gaagtagaac ctgccaataa ttgaaactc tgggacactc ggcattgtta ctgttcagt 1680
 ggttgcaacc ataatggac agcttgctac tggcgaactc caggttgtct caggtatagt 1740
 gacctttgcc cctgggaaa ccattcaaac ctgtttgta gaggctctg ctgacgact 1800
 tccgagatt gaagaggtta tccaagtga actaactgat gctctgtg tgggtactat 1860
 tgggttagat cgaattgcaa atattattat tctgccaact gatgatctct atggtacag 1920
 agcctttgct cagatgggtt atcgtgttca agacctctc gaagaagtt cctgtctaa 1980
 tacaactgtc agcagagcg gaggcactt tggtagctg ttgttgtct acagtactc 2040
 cgaactgat ctatgtgctc tggcaatgga gaaaggtcaa gatttactgt ccaactatga 2100
 atctccactt caagggtgct ctgaccactc ttgagaactc tggatgaatg tctctgacct 2160
 ggggagacc ctgtataact gtgccaactt gtgccttaag gaacaagctt gctcagcgtt 2220
 ttcaattttc agtgcctctg aggttcccaa gtgtttctgg atgacatcat ggaatagacc 2280
 agctgtcaac aattcagact tctggacctc caggaaaaac atgaccaggg tagcatctct 2340
 tttagtggtc aggtctgtgc tgggagtgac tatgagcttg tgacaaggca ntgggacata 2400
 atgacaggaat gtgatgaatt cgcacaaa 2428

<210> 36
 <211> 2243
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 6428177CB1

<400> 36
 gtaactccag gaagagaccg gagcagaccg cgcagagcat agggcggcaa ctgogccgg 60

WO 02/4041

PCT/US01/46055

```

ggccggagac cggcagctgc gtcgggggggg ggcctggccc gagcccgatc tgcgggctcc 120
gagtggtctc ggaagagggg tgggtggtcc gacgggatgc gcttggtggg agaaaccttg 180
gagattccac gcaaggcgta aagcctgggg ctcccaacga tactctgggg agggatggaa 240
gcctcgatgc ctacgcgaa ggaagcgccc agcgggtcct cgtctcagac acacggcgagc 300
acggcgacgc cggagggagg aaaaatgtcc cggctgtctc tcaccccggtc gctgtgtctc 360
ccctcggtcg ccagcgcat cccttcgca gcccauctca ccaagttcaa tgcacctgtg 420
cacctcgatg tggcgggcca calgtacacc agcagcctgg caagctcac caagtacct 480
gaactccagg taagcgccct ctccaatggc actgaaccca tgcctctgga cagtttgagg 540
caacattatt taattgaccg ggaatggggg attttcggct acgtctcgag ctctctggg 600
acgtccaaag tgcctgtccc ggaatgactt aaggacttca gtctgtctga cgaaggaggc 660
cgctactatc agctccagcc catggtgcgc gagctggagc gctggcagca ggaacaggag 720
cagcggcgcc gacgcggggc ctgtgaactg ctggtgggtg gctgcagccc cgaattgggc 780
gagcggatcg cactcagcgg cagaagggcc ctcatcgagg aggtcttccc cgaacaggaa 840
gacgtcatgt gcaactccgt caacgcggcg tggaaacagg accccacgca cgtcatccgc 900
ttcccgctca atggctactg ccggctcaac tgggtacagg tctggaggcg gctgttccag 960
aggggtttca cgtggtcgtc gtcctgtggg ggggtgtgtg actcttccca gttcagcgag 1020
tatgtgcttt gccagagaga cggcgggccc cagctccccc ccactgctgt tcgaatcaag 1080
caggaaaccc tgaactaggc cctggttcag tgcacactg ggcaccccca ggaactggga 1140
aacagtgtcg ggaagttctg cctgtgtata cttggcggcg ggcattgagag ggaagggtgg 1200
gctgaagggg ccaagcgtgg ccagcgagc acaagggtcc caggtgtcat ggcnaaagaa 1260
cgtggagatg tggagggcat cctgcagaa gactgttgat ggcacccaaa gatcacagg 1320
tcggatctct gctgccagct ctcccaagcc ctacgttctg cagcctggcg cagcaltctc 1380
tggggccccc gggcctgttg gggcggggtt ggaagagcgg tctgcagcta cttcagagga 1440
cgtgtttatc cctctcagc cggggcagac tctgagcggt cctctagcgt ccgagagatg 1500
gcttatttcc kacagtattt aaaaatggat cagccctaac tgcaaaagtc agagagggct 1560
acaagagcca atgcttcttt atctggtgct cagttctcag tcagacgtgc agcatggctg 1620
cagggtagac cagctgcctg gaattcagcg ccagatgcct gcaaggctgg ggtctctggg 1680
acagatcgag ggaatgtgtc tgcagggtcg ctggagaggg agtgggtggg gctcagggcg 1740
tgagtgatcc tgtaaccacc tgagacccctc ccgtttggcg ctcaggctgc 1800
actcccccgg taccctgaac tgcctgcagc gggcttccag tccgtgtctg gtcgactggc 1860
acctgggtcg ctggagagat ctctcccgct tggagccagc cccagggtcg caggttactc 1920
caggatggg ccacccatg agggggccca tctgtcagca gctcttcta ggtcccccag 1980
tcaggggagc atccccagct ccagttttct catgcgaata tgcacagttt taattccagt 2040
tgttaactta cctgcccgat gagaacccga caccaggaga cctggcgctc tgcacccctg 2100
attccagtga ggaactggcc tgaaggatcc ttgcagacct gctgcctgcc ccaagacagg 2160
cccaagatg gaccccccgt ggccttgtga cagctcccca agtgttctcc ggtggagaaa 2220
ctgcagagga ctggtggggg ggg 2243

```

<210> 37
 <211> 3711
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Inocyte ID No: 7477243CB1

```

<400> 37
gagcgggtcg ccggcgccgc cgcctcgtgc ctccgcttgg cggccgcggc ccacgcgcca 60
gtgtgttttg tggacgggdc ntcccaagac agccgggtag agccacagctc agcgcgggc 120
agcctctcag cgcagtgttc ggcggagctt gaatcgttlt tgtcctggag aagagaacag 180
agttggcaca cgcacagtgt ttgttggcaa tcatccagtt tcggaaacag aagcttacct 240
tcacaaagaa ttttgtgata atagaatagt ctcatctaa gatacacttt ggaattttct 300
cccaagaagt ctgtttgaac agtttagaag aattgcaaat ttttattttc tcaaatctt 360
ccttgtacag gtcacagtac acacacaaac tagccagtt accagtgga cttcaacttt 420
ctttgttata actgttacag ccacaaagca gggatattga gattgtctga gacacagaga 480
tgacaatgaa gtcaacaaaa gcactgttta cattattgaa eatgcacagc gactgagaaa 540
agaaagtcaa aaaaacaggg ttggtgatgt agtagaagta caggcagatg aacacttccc 600
ctgtgatctt attctctat catcttgccc cactgatgga actgttatg tcaactacag 660
cagctctgat ggggaatbcca attgcagac acatttatga gtacgtgata ccaattgact 720
gtgtacagca caatccatg ataccctccg agcagcaatt gaatgtgac agcctcaacc 780
tgactctcac caatttgttg ggcgaatcaa tatctacagt aatgtcttg agctgtgtg 840
caggtctttg ggaactgaaa atctcttgct gaaaggagct ccgtcaaaa ataccgagaa 900

```


WO 02/40541

PCT/US01/46055

```

galatatgga gttgctgttt acactggaat ggaacccaaa atggccttga actaocaaagg 960
gaactctcag aaagcttctg ctgttgaaaa atctattcaat gtttctctga ttgtatattt 1020
atttatctta ctgacccaaag ctgcagtatg caactactca aagtatgttt ggcaaatgac 1080
cccatacaat gatgaacchtt ggtataacca aaagactcag aaagagcgag agcctctgaa 1140
ggttttaaaa atgttccacc acttctctat atttatggtt ctattcaact ttatctatcc 1200
tgctctccatg taactccacg tagaaatgca gaaattcttt ggctcccttt tcatctcatg 1260
ggataaaggac ttttatgatg aagaatctaa tgaaggagcc ctggttaaca catcacagct 1320
taataagagaa ctgtgtcagg tggttatgtt atttaccagt aagactgaaa cactcactga 1380
aaacagcatg gaattcattg aactgtcat agatggccac aaatataaag gtgtaactca 1440
agaggttgat ggattatctc aaactgatgg aactttaaca tatittgaca aagtagataa 1500
gaatcgagaa gagctgttct taactgctct gtgtttatgt calactgtag aaatcaaaac 1560
aaacgatgct gttgatggag ctacagaatc agctgaatta acctatatct cctcttaacc 1620
agatgaataa gctttggtga aaggagctaa aaggtacggg ttcacatttt taggaatctg 1680
aaatggatat atgagagtag agaaccaaa gaaagaaata gaagaatctg aactcttca 1740
caacttaaac ttgatgctg ccggcgagcg tatgagtgta attgtgaa gaacagagag 1800
agcaatactt ctctcttga aaggagcaga ctggcgagt ttctccagag tgcaaatca 1860
tgaatctgag ttaactaaag tccatgtgga acgtaatgca atggatgggt atcggaact 1920
ctgtgtagcc ttcaaaagaa ttgtctcaga tgattatgaa agaattaca gacagctcat 1980
agggcgaaaa atggccttac aagacagaga agaaaaatg gaaaaagttt tcatgatat 2040
tgagcaaaac atgaatttaa ttggagccac tgcagttgaa gacagctac aagatcaagc 2100
tcagagagcc attgaagctc tgcatgcagc aggcctgaaa gtctgggtgc tcaactggga 2160
cgatgtggag aagactaaa ccacatgcta tgcctgcgcg ctittccaga ccaactgga 2220
gctcttagaa ctacccacaa aaacacttga agaaagtga aggaagagag atcgattaca 2280
tgaaatctat atagaatctc gcaagaaatt gctgcatgag ttctctaaaa gtaactgaag 2340
ctttaaanaa gcatggagag aacatcagga atatggatta atcatagct gctccact 2400
gtcactcata ctaaattcta gtcaagactc tagttcaaac aattacaaaa gcattttct 2460
acaaatatgt atgaagtgta ctgcagctgc ctgcctgcgg atggcaaat taccagaagc 2520
ccagattgtc agaaatgga agaaattaaa aggcagccca ataactctgt ccataggtga 2580
tggtgcaaat gatgttagta tgatcttga atccatgtg ggaatagyla tcaagggaga 2640
agaaagctgc caagcagcta ggaatagaga ttattctgtt ccaaggtta acaactaaa 2700
gaaactgtgt ttggctcatg gacatctata ttatgtgaga atagccacc tcttacagta 2760
ctttctctat aagcaacttt gtttctttt gccacggttt ttgtaccagt tctctgtg 2820
attctccaaa cagccactgt atgatgctgc ttaccttaaa atgtacaata tctgtctc 2880
atcttgcac atcttgctc atagctact ggaacagcac atcaacattg acctctgac 2940
ctcagatccc cgattgtata tgaaaatttc tggcaatgcc atgctacagt tgggcccctt 3000
cttatattgg acatttctgg ctgccttga aggcacagt ttcttctttt ggacttactt 3060
tctttttcag actgcacccc tagaagaaaa tggaaaggtg tacggaaact ggacttttgg 3120
aacattgtt ttacagctc tagtatccac tgaacccctg aagcttgctt tggatacccg 3180
attctggagc tggataaaac aatttgtgat ttggggttct ttacgttct atgtattttt 3240
ctcattcttc tggggagaaa ttatttggcc ttctctcaag caacagagaa tctattttt 3300
atttgcccaa atgctgtctt ctgtatccac atggttggtt ataattcttc taattattat 3360
cagctgttct cctgagatct ttctgtagt attaaagaat gtaagagaaa gaagtgcac 3420
aggaactgtg agcttagaa gggctatbga ctcatctatc ggcagacct cagtcagac 3480
ttttttttt cgaaacttct cagcagatc taatgtattg taacagaatc cgaatctga 3540
actgctcatg ttattgtctc acaagcatac tgacagtgt tatagctaaa aaagaaagca 3600
tgaagaaaca actcaaaaaa gttacatctc oaggtactct gatatgcaac acactaaacc 3660
actctcatgt ctgaatcac aataaatctt attaatgag ggtagaggtt a 3720

```

<210> 38

<211> 2717

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7473042CB1

<400> 38

```

ccgcgcggcg tccaaactcc cagcgtcgga ggcgcggcggg cagcaacttt ctccccggag 60
cggcgcgtgg ggccgtgctt gcctgtgcag ccggagcgga agccggggag aagaagcgg 120
cggcgcggcg gggtgtctcc gcgcgtcggg gctgtctagc tgcgcgagac tgcgcgcgg 180
cggcgcggcg taccgccggg cagcccccgg cccctgcctt atgtcccgca aggcagcga 240
gaatgtggag tacacgctgc ggagcctgag cagcctgatg ggcgagcgcc gcaggagca 300

```

WO 02/40541

PCT/US01/46055

```

gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 360
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 420
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 480
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 540
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 600
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 660
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 720
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 780
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 840
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 900
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 960
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1020
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1080
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1140
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1200
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1260
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1320
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1380
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1440
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1500
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1560
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1620
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1680
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1740
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1800
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1860
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1920
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 1980
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2040
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2100
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2160
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2220
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2280
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2340
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2400
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2460
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2520
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2580
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2640
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2700
gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg gagcgagcgg 2717

```

<210> 39

<211> 2235

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7482060CB1

<400> 39

```

aggagagcgg gagagagcgg agctattcgg occcgggggc tccattcggc gagcgagcgg 60
ctcagggggg cggcccccgg gcttgggaga gggcccccgg gcttcgggtg gagcgagcgg 120
cggggcggag ggttcggggc gagagacagg cggcgttccc agcgggtggg gctcagggct 180
ggcgccggcg agagctcccc ggccaccccc agttaccggc gggcgagggg cgcgttctta 240
ggcgcccgag agatactcaa gcccagggag ccccgaggca cggagagcca gagagcggag 300
cccgagggag cggcgccccc gagcgccggc tcccggta ggcaggggag ggcagggccc 360
gctgagcggc ttgggggtgg gggtggagc gggcccgctg ggcgaaagt ccccggggc 420
ggcagcggag accttcgggg ggcggggggg ggcctcggtg gtgctgaacg tggcgggcg 480
ccgggtattg ctgtccgggg agctgtgaa ggaattcccg ctgcccggcg tgagcgggct 540
ggcagggctg cgtcccgagg ggcaggtgct cggaggtgag gacgactacg acccgaggcg 600
caacgagtac ttcttcgacc ggcactcgga ggccttcggc tctactcgtc tctactcgtc 660

```

WO 02/40541

PCT/US01/46055

```

cggccacggc aagctgggct tggcgccggc gatgtggcgc ctctccttct acaaccgagat 720
gacttactgg ggctctggct ggcggccgct cgggtactgc tggcagggcc gctccgagga 740
cggcatgtcc gacacgtaca ccttctactc ggccgacggc cggggcgtgc tgggcggcga 840
cggggcggcc cccggccggc gaggcggtcc cctccaggcg ctggctggag cgcctggccc 900
ggaccttcga ggagcccaac tcttgcttag cgtgtcggtg gtgttcgtga togtgtccat 960
gggtggtgctg tggccagcca cgttgcccca ctggcgcaac gcagccggcg acaaccggag 1020
cctggatgac cggagcagga taattgaagc tatctgcata ggttggttca ctgccagatg 1080
catcgtgagg ttcatgtgtc ocaaaaacaa gtgtgagttt gtcaagagac ccttgaacat 1140
cattgattta ctggcaatca cgcgtatta catctctgtg ttgatgacag gttttacagg 1200
cggagactct caactccaga gggctggagt cactctgagg gtacttagaa tgaatgagat 1260
tttttgggtg attaagcctg ccogtcaact cattggtctt cagacactcg gtttgaactc 1320
caaacgttgc taacagagaa tggttatgtt acttgtcttc atttgtgttg ccatggcaat 1380
cttttagtga ctttctcagc ttcttgaaca tgggctggac ctggaacac ccaacaagga 1440
ctttacagc attcctgctg cctgctgggt ggtgattatc tctatgacta cagtgggcta 1500
tggagatatg tatcctatca cagtgcctgg aagaattctt ggaggagttc gtgtgtctag 1560
tggagatgtt ctatcggcat taactatcac ttctatctac ctatgcttg tgaagtgta 1620
tcattgagctc aagtttagat ctgtcaggag catctgccta caaagtgtca cttctgtgtc 1680
gggcaactgt gggatatacag agatgacat caacgggctc tgcctgacg ccttgaagaa 1740
tccttgtatg tgcacaagac ccttgaagac ccatctggg gtcccttaca agggcattgc 1800
tgatttgttg cagtctctag aaggtaggcc accggtggag cagctgcgcc cagaccctt 1860
gacccgttgg tgcctccacc ctgccgggag cactctgtgt ggcccggcca acagcattgc 1920
ggttgcattc ccaggagcca ggcgcggcgc gcccggagg ggtttcctga ggacagaggg 1980
ccttgtcctg attgtgcgag caggccctgt cgtatgactt aactgtgaaa atcaccctt 2040
cagggttggg tgcagagact tctgaggcgc gagaagtaga taacctctcg gatagctgt 2100
gagccggggt cctgcatctc cctctgggct ctctgtgac tgagatgtga agcagtgcc 2160
ccatgtcccg aagagaggtt ggccacaact ctgtgcacaa tgctcttcat tttagaatcc 2220
aggatgaagg atatg 2235

```

<210> 40
 <211> 2563
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1578772CB1
 <220>
 <221> unsure
 <222> 2466
 <223> a, t, c, g, or other

```

<400> 40
gagatggag ataacatgt tccagtgaag cttttgtagg ctctaagcgc tgaanaagtt 60
tatggtgccc caggtgtggc tcaaaatcaa aacacaaaac tataaataaa aettaagaga 120
taaaattctt atttttttca oacaaatgaca ttgtotitta aaaggaatat oacatataaa 180
agtctccccc taattttttt ggtttccctt actttgctgt tgcctaaac acatataccc 240
tgtttatagg gtaacccaac sggtgctctt aaggtggagg tgggctacaa aactggggaa 300
cttgaacaat ggaactctac aaggagactg gtgaataatg ggcatttaac taattggggg 360
gagcgggagg gttagatgtg aatggcagaa tattaagaag ggggttgtgt ggaaggagat 420
ttgggagagg ggagacttcc gaggagagat aggcagagtg ggcaggaaga ccagctctca 480
tgttgggggt ggaggtcttc ttcctttttt gctcctgttc ctccctttct ctatgctgca 540
ggcctcttct cctgcacagt ggtcctgtgc gtctgtgtgt ggttggggcc cttcttttac 600
tatctgccca aggtgtctct ggcttgcatc aacatctcca gcattgcgca ggtcttctgc 660
cagatgcagg aacttcacca actatggcac atcagccgag tggactttgc tgtgtggatg 720
gtacactggg tggcaagtgt gacccagagt gtgatttgg gccctggctgt ggggtgtggtc 780
ttctccatga tgaactgagc ctggccgacc cgaagctctt ccaagctccc gggctctaca 840
tcctgagcta tccaaaccca ctgtactctg ggaacggtag gaagtttgcg tggcaactgg 900
agtgcacact ggggctcgga gaaggagaaa agagacttc aaagccagat gggccaatgg 960
ttgcagtgtc tgagcctgtc aggtgtgtgt tcctagactt cngtgggtgt aacttttgcg 1020
atgtctgtcg gccacagaaa gtgtgtgagc tggccagccg atgtcgagat gctaggatcc 1080
gctcctctct ggctcagttt aatgccttgg tgcaggggac autgaccogg gtatgagctc 1140
tggacagggt gaetccagat cagctgcttg tgaagtgtga ggaatgcagc gcttatgccc 1200

```

WO 02/40541

PCT/US01/46055

```

tggggagcct ggttaaggggc agtagaccac gggagcggag ccaggaggcc ctgggctgcy 1260
gcaagtgaag caggggagct cactgaccca aagatttgca ccgltgtggg ctgacctcat 1320
catgtggagt gcagggggoc ctgatgacat gtgtgtgagt aggaacctga cctttgaac 1380
cccttaocta acgttaactaa taaaatgaag ctggagagctt tgaatccat gaagtgaatc 1440
taggtgtttt cacaaggact ctggcgcccc ttcttttctg cccacagcat tgcagagaca 1500
caactaagaa tggctttcac caacocccag ccttaccccc agccccaagc ccaacagttc 1560
tctctgtggg ggtggggctg gacaggtac acagagtact ggaatctgaag atgcagatga 1620
ggggcacagt cttgggatta tgtgttggg aacttcccca cccctctggt cccaagatga 1680
ggggacagta ttccacatt aggtttctag agtccctctg ggctctttgg cacttggaaa 1740
gtatccccc catttctctg cccatgaaa tgggcagggg gaggacttgg cactggctgt 1800
gggagaggtt atggctccac caggcctctg ggaactggaa aaaggagggg tgtcaccagg 1860
adaaccccta ctgggcacca ggttctgaaa aggaagagt aggaactaga ggtcaggga 1920
cagccagtag agtgcacaca ggtgtgtgg gtttgttgg aattctctgg agggacaagg 1980
aggcaggccc agctcgacag tcagaatccc ccaggggccc atcactggga ggtcatgac 2040
tgcagccggt gtttgaagaa gagcagcggg tgtttggcca tctggctctc gtccagtcat 2100
cttgggtaac ggtacccggc gtaagtctct gttccggcat cctttagatg caaggtcagc 2160
tcaagtttgg atgcatgac cccacgagc tgggagaggg agccaacccg aaggttaaca 2220
agcccatccc agaagaattc cagatgagcc acccggctga agcgggggtc gaaacgaac 2280
tcgcgcacct tgtcagtcgc cgcocggag aagttaacca cgcctctgat gaccacgaag 2340
ccttgggaagc cgaagagctc gtgtgtggaag cagcgccctt gccggaggag ttcccagagg 2400
ctggggcgcc gggctccacy ctacagagct gccgataagt ggtggcaag ccggagatct 2460
cgcgcnacgg cccccacca ggtccagcgg cgtccgtcca gcacgtcaca agcttctcag 2520
cgcgtccggc cgtgaagacg aatcgttctg caccacagca cga 2563

```

【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
23 May 2002 (23.05.2002)

PCT

(10) International Publication Number
WO 02/040541 A3(51) International Patent Classification: C07K 14/705,
C12N 15/12, C07K 16/18, C12Q 1/68, A01K 67/027,
A51K 38/17, 39/295, C01N 33/68

(21) International Application Number: PCT/US01/46055

(22) International Filing Date: 28 October 2001 (28.10.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/345,989 17 October 2000 (17.10.2000) US
60/245,964 3 November 2000 (03.11.2000) US
60/247,673 3 November 2000 (03.11.2000) US
60/249,661 17 November 2000 (17.11.2000) US
60/257,232 30 November 2000 (30.11.2000) US
60/250,790 1 December 2000 (01.12.2000) 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): TANG, Y., Tom
[US/US], 1240 Rarick Court, San Jose, CA 95118 (US);
YUE, Henry [US/US], 826 Lois Avenue, Sunnyvale,
CA 94087 (US); NGUYEN, Donniel, B. [US/US], 1403
Ridgewood Drive, San Jose, CA 95118 (US); HAFALIA,
April, J., A. [US/US], 2237 Calle de Primavera, Santa
Clara, CA 95051 (US); ELLIOTT, Vicki, S. [US/US],
3770 Polon Place Way, San Jose, CA 95121 (US); LU,
Yan [US/US], 3880 Cortina Way, Palo Alto, CA 94303
(US); WALIA, Narinder, K. [US/US], 800 Davis Street,
#205, San Leandro, CA 94577 (US); YAO, Monique,
G. [US/US], 1189 Woodgate Drive, Carmel, IN 46033
(US); BAUGHN, Mariah, R. [US/US], 1424 Santiago
Road, San Leandro, CA 94577 (US); GANDHI, Ameena,
R. [US/US], 205 5th Avenue, San Francisco, CA 94118
(US); DING, Li [CN/US], 3353 Alma Street, #146, Palo
Alto, CA 94306 (US); SANJANWALA, Madhusudan
[US/US], 210 Silvia Court, Los Altos, CA 94024 (US);
RAMKUMAR, Jayashanti [IN/US], 34359 Maybird
Circle, Fremont, CA 94555 (US); ARVIZU, Chandra
[US/US], 490 Sherwood Way, #1, Menlo Park, CA
94025 (US); GIEZERN, Kimberly, J. [US/US], 691
Los Ilucos Drive, San Jose, CA 95123 (US); LAL,
Preethi, G. [IN/US], PO, Box 5112, Santa Clara, CA95056 (US); AZIMZAI, Yalda [US/US], 5518 Boulder
Canyon Drive, Castro Valley, CA 94552 (US); KHAN,
Farrah A. [IN/US], 9445 Hamilton Street, Des Plaines,
IL 60016 (US); THANGAVELU, Kavitha [IN/US],
950 Montecito Avenue, #73, Mountain View, CA 94035
(US); THORNTON, Michael [US/US], 9 Medway
Road, Woodside, CA 94062 (US); LU, Dyoung, Aina,
M. [US/US], 233 Coy Drive, San Jose, CA 95124 (US);
TRIBOULEY, Catherine, M. [FR/US], 1121 Tennessee
Street, #5, San Francisco, CA 94107 (US); WARREN,
Bridget, A. [US/US], 10130 Parkwood Drive, #2, Cupertino,
CA 95014 (US); ISON, Craig, H. [US/US], 1242
Weathersfield Way, San Jose, CA 95118 (US); DAS,
Debojyoti [IN/US], 2267 Parkington Avenue, Sunnyvale,
CA 94087 (US); KAUMANN, Brigitte, E. [US/US], 2801
South Duncanson Avenue #3B, Chicago, IL 60627 (US);
POLICKY, Jennifer, L. [US/US], 1511 Court Lewis
Court, San Jose, CA 95118 (US); KEARNEY, Liam
[IE/US], 50 Woodside Avenue, San Francisco, CA 94127
(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, AI, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, GR, GU, HD, 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, NO, NZ, 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); European
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, SI, SK, UA); OAPI patent (BF, BJ, CG,
CI, CM, GN, GQ, GW, ML, MR, NI, SN, TD,
TG).Published:
— with international search report(88) Date of publication of the international search report:
10 July 2003For two-letter codes and other abbreviations, refer to the "Instructions
to Applicants" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/040541 A3

(54) Title: TRANSPORTERS AND ION CHANNELS

(57) Abstract: The invention provides human transporters and ion channels (TRICH) and polynucleotides which identify and encode TRICH. 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 TRICH.

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/46055
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C07K14/705 C12N16/18 A01K67/027 A61K38/17 A61K39/395 G01N33/68		
According to International Patent Classification (IPC) by both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C07K C12N C12Q A01K A61K G01N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, PAJ, MEDLINE, SEQUENCE SEARCH		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99 24836 A (ENDRESS GREGORY A ; HUMAN GENOME SCIENCES INC (US); FENG PING (US);) 20 May 1999 (1999-05-20) see whole doc. esp. seq ids 25,150,277	1-56,76
X	WO 00 08157 A (HU PING ; LAUBERT BORIS (US); AKYS PHARM INC (US); BUCKLER ALAN J) 17 February 2000 (2000-02-17) see whole doc. esp. claims and p.23. ff.	1-56,76
X	WO 99 47670 A (TATE SIMON NICHOLAS ; GLAXO GROUP LTD (GB); GROSE DAVID THOMAS (GB)) 23 September 1999 (1999-09-23) see whole doc. esp. claims and e.g. p.16 ff.	1-56,76
--- -/--		
☒ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "F" earlier document published on or after the international filing date "L" document which may throw doubts on the (y) claims or which is cited to establish the publication date of another citation or other special category (see 6.02) (y) "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 (see 6.02)		
"T" later document published after the international filing date or priority date and not in conflict with the application but used to articulate 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 considered with one or more other such documents, such combination being obvious to a person skilled in the art. "S" document member of the same patent family		
Date of the actual completion of the international search		Date of making of the international search report
27 November 2002		12.03.03
Name and mailing address of the ISA European Patent Office, P.O. 5818 Patentstein 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040; Tx: 31 651 apo nl; Fax: (+31-70) 340-5016		Authorized officer: Mueller, F

Form PCT/ISA/210 (second sheet) July 2002

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/46055
C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Designation of document, with indication where appropriate, of the relevant passages	Relevant to claim No.
X	MARICQ A V ET AL: "Primary structure and functional expression of the 5HT3 receptor, a serotonin-gated ion channel" SCIENCE. AAAS. LANCASTER, PA, US, vol. 254, 18 October 1991 (1991-10-18), pages 432-437, XP002157173 ISSN: 0036-8075 the whole document	1-56,76
A	COOPER E C ET AL: "Ion channel genes and human neurological disease: Recent progress, prospects, and challenges" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 96, no. 9, 27 April 1999 (1999-04-27), pages 4759-4766, XP002174532 ISSN: 0027-8424 the whole document	1-56,76

Form PCT/ISA/210 (continued from sheet 1) (July 1997)

INTERNATIONAL SEARCH REPORT	International Application No. PCT/US 01/46055
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: Although claims 19, 33, 35 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
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.: 1-20, 23, 26-55 (partially); 56, 76 (complete)
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 01/46055

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 21,22,24,25

Present claims 21,22,24,25 relate to a compound defined by reference to a desirable characteristic or property, namely agonists or antagonists to seq id 1-20.

The claims cover all compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, no search has been carried out.**

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 A6055

FURTHER INFORMATION CONTINUED FROM PCT/SA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-20,23,26-55(partially); 56,76 (complete)

protein and nucleic acid sequence seq id: 1/21 and dependent subject-matter therefore as defined by claims 1-20,23,26-56, and 76.

inventions 2-20:
protein and nucleic acid sequences 2/22; 3/23 etc. and dependent subject-matter thereof as defined in the dependent claims (1-20, 23,26-55,77-95 (as far as they relate to the claimed sequences 2/22, 3/23 etc)).

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/US 01/46055

Patent document cited in search report		Publisher date	Patent family member(s)	Publication date
WO 9924836	A	20-05-1999	AU 1303799 A	31-05-1999
			CA 2308768 A1	20-05-1999
			EP 1032838 A1	06-09-2000
			WO 9924836 A1	20-05-1999
			US 2003028003 A1	06-02-2003
			AU 1303199 A	31-05-1999
WO 0008157	A	17-02-2000	AU 5340699 A	28-02-2000
			EP 1141282 A2	10-10-2001
			WO 0008157 A2	17-02-2000
WO 9947670	A	23-09-1999	AU 2946399 A	11-10-1999
			EP 1062334 A1	27-12-2000
			WO 9947670 A1	23-09-1999
			JP 2002508941 T	26-03-2002

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード (参考)
C 0 7 K 14/705	C 0 7 K 16/18	4 B 0 6 5
C 0 7 K 16/18	C 0 7 K 16/28	4 C 0 8 4
C 0 7 K 16/28	C 0 7 K 16/46	4 C 0 8 5
C 0 7 K 16/46	C 1 2 M 1/00	4 H 0 4 5
C 1 2 M 1/00	C 1 2 N 1/15	
C 1 2 N 1/15	C 1 2 N 1/19	
C 1 2 N 1/19	C 1 2 N 1/21	
C 1 2 N 1/21	C 1 2 P 21/02	C
C 1 2 N 5/10	C 1 2 P 21/08	
C 1 2 N 15/02	C 1 2 Q 1/68	A
C 1 2 P 21/02	G 0 1 N 33/15	Z
C 1 2 P 21/08	G 0 1 N 33/50	Z
C 1 2 Q 1/68	G 0 1 N 33/53	D
G 0 1 N 33/15	G 0 1 N 33/53	M
G 0 1 N 33/50	G 0 1 N 37/00	1 0 2
G 0 1 N 33/53	C 1 2 N 15/00	C
G 0 1 N 37/00	C 1 2 N 15/00	F
	C 1 2 N 5/00	A
	A 6 1 K 37/02	

- (31)優先権主張番号 60/249,661
 (32)優先日 平成12年11月17日(2000.11.17)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/252,232
 (32)優先日 平成12年11月20日(2000.11.20)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/250,790
 (32)優先日 平成12年12月1日(2000.12.1)
 (33)優先権主張国 米国(US)

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,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 ユエ、ヘンリー
 アメリカ合衆国カリフォルニア州9 4 0 8 7・サニーバイル・ルイスアベニュー 8 2 6
 (72)発明者 ニュエン、ダニエル・ビー
 アメリカ合衆国カリフォルニア州9 5 1 1 8・サンノゼ・リッジウッドドライブ 1 4 0 3
 (72)発明者 ハファリア、エープリル・ジェイ・エイ
 アメリカ合衆国カリフォルニア州9 5 0 5 4・サンタクララ・コーレデブリマベール 2 2 2 7
 (72)発明者 エリオット、ビッキー・エス
 アメリカ合衆国カリフォルニア州9 5 1 2 1・サンノゼ・ポルトンプレイスウェイ 3 7 7 0
 (72)発明者 リュ、ヤン
 アメリカ合衆国カリフォルニア州9 4 3 0 3・パロアルト・コリーナウェイ 3 8 8 5
 (72)発明者 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 3
3
- (72)発明者 ヤオ、モニーク・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 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 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 ディング、リー
アメリカ合衆国カリフォルニア州94306・パロアルト・#146・アルマストリート 335
3
- (72)発明者 サンジャンワラ、マデュスダン
アメリカ合衆国カリフォルニア州94024・ロスアルトス・シルビアコート 210
- (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 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 ギーツェン、キンバリー・ジェイ
アメリカ合衆国カリフォルニア州95123・サンノゼ・ロスウエコスドライブ 691
- (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 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 カーン、ファラ・エイ
アメリカ合衆国イリノイ州60016・デブレイン・ハリソンストリート 9445
- (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 ソーントン、マイケル
アメリカ合衆国カリフォルニア州94062・ウッドサイド・メッドウェイロード 9
- (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・コイドライブ 233
- (72)発明者 トリボレー、キャサリン・エム
アメリカ合衆国カリフォルニア州94107・サンフランシスコ・#5・テネシーストリート 1
121
- (72)発明者 ワレン、ブリジット、エイ
アメリカ合衆国カリフォルニア州95014・クーペルティノー・#2・パークウッドドライブ
10130
- (72)発明者 アイソン、クレイグ・エイチ
アメリカ合衆国カリフォルニア州95118・サンノゼ・ウェザーズフィールドウェイ 1242
- (72)発明者 ダス、デボプリヤ
アメリカ合衆国カリフォルニア州94087・サニーバイル・パーキントンアベニュー 1267
- (72)発明者 ラーマン、ブリジット・イー
アメリカ合衆国イリノイ州60637・シカゴ・#3ビー・サウスドーチェスターアベニュー 5
801
- (72)発明者 ポリッキー、ジェニファー・エル
アメリカ合衆国カリフォルニア州95118・サンノゼ・コートジャービスコート 1511
- (72)発明者 キーニー、ライアム
アメリカ合衆国カリフォルニア州94127・サンフランシスコ・ウッドサイドアベニュー 50

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 BA80 CA04 CA05 CA06 CA07 CA09 CA10
CA11 DA02 DA06 EA02 EA04 FA02 FA10 GA05 GA11 HA03
HA08 HA12 HA14 HA19
4B029 AA07 AA21 AA23 BB20 CC03 CC08 FA12
4B063 QA01 QA18 QQ20 QQ42 QR08 QR13 QR15 QR32 QR41 QR42
QR43 QR55 QR59 QR66 QR69 QR77 QR82 QS12 QS24 QS25
QS28 QS34 QS36 QS39 QX01 QX02 QX07
4B064 AG01 AG27 CA02 CA10 CA19 CA20 CC24 CE13 DA01 DA13
4B065 AA26X AA90X AA93Y AB01 AC14 BA02 BA25 CA24 CA44 CA46
4C084 AA01 AA17 BA02 ZA022 ZA942 ZB072
4C085 AA14 EE01
4H045 AA11 AA20 AA30 BA10 BA41 CA40 DA50 DA76 DA86 EA20
EA50 FA72 FA74 GA26

专利名称(译)	转运蛋白和离子通道		
公开(公告)号	JP2004531213A	公开(公告)日	2004-10-14
申请号	JP2002543549	申请日	2001-10-25
[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	タングワイトム ユエヘンリー ニュエンダニエルビー ハファリアエープリルジェイエイ エリオットビッキーエス リュヤン チョーラナリンダーケイ ヤオモニークジー ボーグンマライアアール ガンディーアミーナアール デイングリー サンジャンワラマデウスダン ランクマールジャヤラクシミ アービズチャンドラ ギーツエンキンバリージェイ ラルプリーティジー アジムザイヤルダ カーンファラエイ サンガベルカピサ ソーントンマイケル リュデュングアイナエム トリボレーキャサリーンエム ワレンブリジットエイ アイソククレイグエイチ ダスデボプリヤ ラーマンブリジットイー ポリッキージェニファーエル キーニーライアム		
发明人	タング、ワイトム ユエ、ヘンリー ニュエン、ダニエル・ビー ハファリア、エープリル・ジェイ・エイ エリオット、ビッキー・エス リュ、ヤン チョーラ、ナリンダー・ケイ ヤオ、モニーク・ジー ボーグン、マライア・アール ガンディー、アミーナ・アール デイング、リー サンジャンワラ、マデウスダン ランクマール、ジャヤラクシミ アービズ、チャンドラ ギーツエン、キンバリー・ジェイ ラル、プリーティ・ジー		

アジムザイ、ヤルダ
カーン、ファラ・エイ
サンガベル、カピサ
ソーントン、マイケル
リュ、デュング・アイナ・エム
トリボレー、キャサリーン・エム
フレン、ブリジット、エイ
アイソン、クレイグ・エイチ
ダス、デボプリヤ
ラーマン・ブリジット・イー
ポリッキー、ジェニファー・エル
キーニー、ライアム

IPC分类号	G01N33/50 A61K38/00 A61K39/395 A61K45/00 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 C12Q1/68 G01N33/15 G01N33/53 G01N37/00
CPC分类号	C07K14/705
FI分类号	C12N15/00.ZNA.A A61K39/395.N A61K45/00 C07K1/22 C07K14/705 C07K16/18 C07K16/28 C07K16/46 C12M1/00.A C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/68.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
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/BA80 4B024/CA04 4B024/CA05 4B024/CA06 4B024/CA07 4B024/CA09 4B024/CA10 4B024/CA11 4B024/DA02 4B024/DA06 4B024/EA02 4B024/EA04 4B024/FA02 4B024/FA10 4B024/GA05 4B024/GA11 4B024/HA03 4B024/HA08 4B024/HA12 4B024/HA14 4B024/HA19 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/QR13 4B063/QR15 4B063/QR32 4B063/QR41 4B063/QR42 4B063/QR43 4B063/QR55 4B063/QR59 4B063/QR66 4B063/QR69 4B063/QR77 4B063/QR82 4B063/QS12 4B063/QS24 4B063/QS25 4B063/QS28 4B063/QS34 4B063/QS36 4B063/QS39 4B063/QX01 4B063/QX02 4B063/QX07 4B064/AG01 4B064/AG27 4B064/CA02 4B064/CA10 4B064/CA19 4B064/CA20 4B064/CC24 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/AA01 4C084/AA17 4C084/BA02 4C084/ZA022 4C084/ZA942 4C084/ZB072 4C085/AA14 4C085/EE01 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
優先权	60/243989 2000-10-27 US 60/245904 2000-11-03 US 60/247673 2000-11-09 US 60/249661 2000-11-17 US 60/252232 2000-11-20 US 60/250790 2000-12-01 US
外部链接	Espacenet

摘要(译)

本发明提供了人转运蛋白和离子通道 (TRICH) , 以及鉴定和编码TRICH的多核苷酸。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TRICH异常表达相关的疾病的方法。

.1) 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	39/395	N	4 B O 2 4
A 6 1 K	39/395	A 6 1 K	45/00		4 B O 2 9
A 6 1 K	45/00	C O 7 K	1/22		4 B O 6 3
C O 7 K	1/22	C O 7 K	14/705		4 B O 6 4
		審査請求	未請求	予備審査請求	有 (全 302 頁) 最終頁に
.1) 出願番号		特願2002-543549 (P2002-543549)		(71) 出願人	
.6) (22) 出願日		平成13年10月25日 (2001. 10. 25)		インサイト・ゲノミックス・インコーポ	
.5) 翻訳文提出日		平成15年4月24日 (2003. 4. 24)		イテッド	
.6) 国際出願番号		PCT/US2001/046055		アメリカ合衆国カリフォルニア州94:	
.7) 国際公開番号		W02002/040541		4・パロアルト・ボータードライブ : 60	
.7) 国際公開日		平成14年5月23日 (2002. 5. 23)		(74) 代理人	
.1) 優先権主張番号		60/243, 989		100089266	
.2) 優先日		平成12年10月27日 (2000. 10. 27)		弁理士 大島 陽一	
.3) 優先権主張国		米国 (US)		(72) 発明者	
.1) 優先権主張番号		60/245, 904		タング、ワイ・トム	
.2) 優先日		平成12年11月3日 (2000. 11. 3)		アメリカ合衆国カリフォルニア州95	
.3) 優先権主張国		米国 (US)		8・サンノゼ・ランウィックコート・30	
.1) 優先権主張番号		60/247, 673			
.2) 優先日		平成12年11月9日 (2000. 11. 9)			
.3) 優先権主張国		米国 (US)			