

(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2004-507218

(P2004-507218A)

(43) 公表日 平成16年3月11日(2004.3.11)

(51) Int.Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A 2 G O 4 5
A 6 1 K 38/00	A 6 1 K 45/00	4 B O 2 4
A 6 1 K 45/00	A 6 1 K 49/00	A 4 B O 6 3
A 6 1 K 49/00	A 6 1 P 1/04	4 B O 6 4
A 6 1 P 1/04	A 6 1 P 1/16	4 B O 6 5
審査請求 未請求 予備審査請求 有 (全 262 頁) 最終頁に続く		

(21) 出願番号	特願2001-575644 (P2001-575644)	(71) 出願人	301005050
(86) (22) 出願日	平成13年4月6日 (2001.4.6)		インサイト・ゲノミックス・インコーポレイテッド
(85) 翻訳文提出日	平成14年10月4日 (2002.10.4)		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/US2001/011206		
(87) 国際公開番号	W02001/077174	(74) 代理人	100089266
(87) 国際公開日	平成13年10月18日 (2001.10.18)		弁理士 大島 陽一
(31) 優先権主張番号	60/195,595	(72) 発明者	レディ、ルーパ
(32) 優先日	平成12年4月6日 (2000.4.6)		アメリカ合衆国カリフォルニア州94086・サニーベイル・#3・ウェストマッキンレーアベニュー 1233
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/196,872		
(32) 優先日	平成12年4月12日 (2000.4.12)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/199,020		
(32) 優先日	平成12年4月20日 (2000.4.20)		
(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、ヒト輸送体及びイオンチャネル (T R I C H) と、T R I C Hを同定及びコードするポリヌクレオチドとを提供する。本発明はまた、発現ベクター、宿主細胞、抗体、アゴニスト、およびアンタゴニストを提供する。更に、本発明は、T R I C Hの異常発現に関連する疾患の診断・治療・予防方法を提供する。

【特許請求の範囲】

【請求項 1】

単離されたポリペプチドであって、

(a) SEQ ID NO: 1 乃至 SEQ ID NO: 15 (SEQ ID NO: 1-15) からなる群から選択されたアミノ酸を含むポリペプチドと、

(b) SEQ ID NO: 1、SEQ ID NO: 4、及び SEQ ID NO: 6-15 からなる群から選択されたアミノ酸配列と少なくとも 90% の同一性を有するアミノ酸配列を含む天然のポリペプチドと、

(c) SEQ ID NO: 2 のアミノ酸配列に 92% の同一性を有するアミノ酸配列を含む天然のポリペプチドと、

(d) SEQ ID NO: 3 のアミノ酸配列に 96% の同一性を有するアミノ酸配列を含む天然のポリペプチドと、

(e) SEQ ID NO: 5 のアミノ酸配列に 99% の同一性を有するアミノ酸配列を含む天然のポリペプチドと、

(f) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、

(g) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されることを特徴とする単離されたポリペプチド。

10

【請求項 2】

SEQ ID NO: 1-15 からなる群から選択された請求項 1 の単離されたポリペプチド。

20

【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: 16-30 からなる群から選択された請求項 4 の単離されたポリヌクレオチド。

【請求項 6】

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

30

【請求項 7】

請求項 6 の組換えポリヌクレオチドで形質転換された細胞。

【請求項 8】

請求項 6 の組換えポリヌクレオチドを含む遺伝子組換え生物。

【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、

(a) 前記ポリペプチドの発現に好適な条件下で、請求項 1 のポリペプチドをコードするポリヌクレオチドに機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養するステップと、

(b) そのように発現したポリペプチドを回収するステップとを含むことを特徴とする請求項 1 のポリペプチドの生産方法。

40

【請求項 10】

請求項 1 のポリペプチドに特異的に結合する単離された抗体。

【請求項 11】

単離されたポリヌクレオチドであって、

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

(b) SEQ ID NO: 16-30 からなる群から選択されたポリヌクレオチド配列と

50

少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、

(c) 前記(a)のポリヌクレオチドに相補的なポリヌクレオチドと、

(d) 前記(b)のポリヌクレオチドに相補的なポリヌクレオチドと、

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

【請求項12】

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

【請求項13】

サンプルにおける請求項11に記載のポリヌクレオチド配列を有する標的ポリヌクレオチドを検出する方法であって、

(a) 前記サンプルをプローブでハイブリダイズするステップであって、前記プローブが、前記サンプル内の前記標的ポリヌクレオチドに相補的な配列を含む少なくとも20個の連続するヌクレオチドを含み、前記プローブと前記標的ポリヌクレオチドまたはその断片との間でハイブリダイゼーション複合体が形成される条件下で、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズする、該ステップと、

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

【請求項14】

前記プローブが少なくとも60個の連続するヌクレオチドを含むことを特徴とする請求項13に記載の方法。

【請求項15】

サンプルにおける請求項11のポリヌクレオチド配列を有する標的ポリヌクレオチドを検出する方法であって、

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

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

【請求項16】

請求項1のポリペプチド及び医薬的に容認できる賦形剤を含む組成物。

【請求項17】

前記ポリペプチドが、SEQ ID NO: 1-15からなる群から選択されたアミノ酸配列を含むことを特徴とする請求項16の組成物。

【請求項18】

機能的なTRICH（新規の輸送体及びイオンチャネル）の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項16の組成物を投与することを含むことを特徴とする治療方法。

【請求項19】

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

(a) 請求項1のポリペプチドを含むサンプルを化合物に曝露するステップと、

(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20】

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

【請求項21】

10

20

30

40

50

機能的な T R I C H の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 20 の組成物を投与することを含むことを特徴とする治療方法。

【請求項 22】

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

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

【請求項 23】

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

【請求項 24】

機能的な 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 23 の組成物を投与することを含むことを特徴とする治療方法。

【請求項 25】

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

- (a) 請求項 1 のポリペプチドを好適な条件下で少なくとも 1 つの試験化合物と結合させるステップと、
- (b) 請求項 1 のポリペプチドと前記試験化合物との結合を検出して、請求項 1 のポリペプチドと特異的に結合する化合物を同定するステップとを含むことを特徴とするスクリーニング方法。

【請求項 26】

請求項 1 のポリペプチドの活性を変化させる化合物をスクリーニングする方法であって、

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

【請求項 27】

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

- (a) 前記標的ポリヌクレオチドの発現に好適な条件下で、前記標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、
- (b) 前記標的ポリヌクレオチドの発現の変化を検出するステップと、
- (c) 様々な量の前記化合物の存在下での前記標的ポリヌクレオチドの発現と、前記化合物の非存在下での前記標的ポリヌクレオチドの発現とを比較するステップとを含むことを特徴とするスクリーニング方法。

【請求項 28】

試験化合物の毒性を評価する方法であって、

- (a) 核酸を含む生体サンプルを前記試験化合物で処理するステップと、
- (b) 処理した前記生体サンプルの核酸を、請求項 11 のポリヌクレオチドの少なくとも 20 個の連続するヌクレオチドを含むプローブとハイブリダイズさせるステップであって、このハイブリダイゼーションが、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われ、前記

10

20

30

40

50

標的ポリヌクレオチドが、請求項 11 のポリヌクレオチドのポリヌクレオチド配列またはその断片を含むポリヌクレオチドである、前記ステップと、

(c) ハイブリダイゼーション複合体の収量を定量するステップと、

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

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

【請求項 29】

生体サンプルにおける T R I C H の発現に関連する病態または疾患の診断テストであって 10

(a) 請求項 10 の抗体と前記ポリペプチドとの結合に好適な条件下で、前記生体サンプルを請求項 10 の抗体と結合させて、抗体／ポリペプチド複合体を形成するステップと、

(b) 前記複合体を検出するステップとを含み、

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

【請求項 30】

前記抗体が、

(a) キメラ抗体、

(b) 一本鎖抗体、

(c) F a b 断片、

(d) F (a b ')₂ 断片、または

(e) ヒト化抗体であることを特徴とする請求項 10 に記載の抗体。 20

【請求項 31】

請求項 10 の抗体及び許容される賦形剤を含む組成物。

【請求項 32】

患者において、T R I C H の発現に関連する病態または疾患を診断する方法であって、請求項 31 の組成物を前記患者に有効量投与することを含むことを特徴とする方法。

【請求項 33】

前記抗体が標識されていることを特徴とする請求項 31 に記載の組成物。 30

【請求項 34】

患者において、T R I C H の発現に関連する病態または疾患を診断する方法であって、請求項 33 の組成物を前記患者に有効量投与することを含むことを特徴とする方法。

【請求項 35】

請求項 10 の抗体の特異性を有するポリクローナル抗体を作製する方法であって、

(a) 抗体反応が起こる条件下で、S E Q I D N O : 1 - 15 からなる群から選択されたアミノ酸配列またはその免疫原性断片を有するポリペプチドで動物を免疫化するステップと、

(b) 前記動物から抗体を単離するステップと、

(c) 前記単離された抗体を前記ポリペプチドでスクリーニングして、S E Q I D N O : 1 - 15 からなる群から選択されたアミノ酸配列を有するポリペプチドに特異的に結合するポリクローナル抗体を同定するステップとを含むことを特徴とするポリクローナル抗体作製方法。 40

【請求項 36】

請求項 35 の方法により作製された抗体。

【請求項 37】

請求項 36 の抗体及び適当な担体を含む組成物。

【請求項 38】

請求項 10 の抗体の特異性を有するモノクローナル抗体を作製する方法であって、

(a) 抗体反応が起こる条件下で、S E Q I D N O : 1 - 15 からなる群から選択され 50

たアミノ酸配列またはその免疫原性断片を有するポリペプチドで動物を免疫化するステップと、

(b) 前記動物から抗体を産生する細胞を単離するステップと、

(c) 前記抗体産生細胞を不死化細胞と融合し、モノクローナル抗体産生ハイブリドーマ細胞を形成するステップと、

(d) 前記ハイブリドーマ細胞を培養するステップと、

(e) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドに特異的に結合する前記培養モノクローナル抗体から単離するステップとを含むことを特徴とするモノクローナル抗体作製方法。

【請求項 39】

10

請求項 38 の方法で作製したモノクローナル抗体。

【請求項 40】

請求項 39 の抗体及び適当な担体を含む組成物。

【請求項 41】

Fab 発現ライブラリのスクリーニングにより得られた請求項 10 の抗体。

【請求項 42】

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

【請求項 43】

サンプルにおける SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドを検出する方法であって、

20

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

(b) 特異的な結合を検出するステップとを含み、

前記特異的な結合が、前記サンプルに SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドが存在することを示唆することを特徴とする方法。

【請求項 44】

SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドをサンプルから精製する方法であって、

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

30

(b) 前記サンプルから前記抗体を分離して、SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有する精製されたポリペプチドを得るステップとを含むことを特徴とする方法。

【請求項 45】

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 46】

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 47】

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。

40

【請求項 48】

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 49】

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 50】

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 51】

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 52】

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。

50

【請求項 53】

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 54】

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 55】

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 56】

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 57】

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。

【請求項 58】

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

【請求項 59】

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

【請求項 60】

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

【請求項 61】

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

【請求項 62】

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

【請求項 63】

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

【請求項 64】

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

【請求項 65】

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

【請求項 66】

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

【請求項 67】

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

【請求項 68】

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

【請求項 69】

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

【請求項 70】

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

【請求項 71】

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

10

20

30

40

50

【請求項 7 2】

SEQ ID NO: 28 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 7 3】

SEQ ID NO: 29 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【請求項 7 4】

SEQ ID NO: 30 のポリヌクレオチド配列を含む請求項 1 1 に記載のポリヌクレオチド。

【発明の詳細な説明】

10

【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、免疫疾患、及び細胞増殖異常の診断・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分されている。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代謝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. and C. Sansom (1998) The Transporter Facts Book. Academic Press, San Diego CA, pp. 3-29)。輸送は、受動濃度依存メカニズムによって発生させることができ、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リンクすることができる。輸送時に機能するタンパク質には、担体タンパク質及びチャネルタンパク質がある。担体タンパク質は、特定の溶質に結合して膜を通過して結合溶質を転位させるような高次構造上の変化を受ける。チャネルタンパク質は、特定の溶質が電気化学溶質勾配下向きに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。

20

30

【0003】

膜の一方の側から他方の側へ単一溶質を輸送する担体タンパク質をユニポーターと呼ぶ。対照的に、結合輸送体は、1つの溶質の移動を第2溶質の同時または連続移動に、同一方向 (シンポート) または逆方向 (アンチポート) のいずれかでリンクする。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、電気化学勾配下向きに細胞移動し、それによって溶質を細胞に運ぶ。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アーゼ系によって維持される。ナトリウム結合輸送体には、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、並びに細胞質的に方向付けられた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. ら (1997) Proc. Natl. Acad. Sci. USA 94 : 5568-5573)。SMVT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塩などの水溶性ビタミンの輸送に結び付けられる (Prasad, P. D. ら (1998) J. Biol. Chem. 273 : 7501-7506)。

40

50

【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担体である。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、リン酸塩、硝酸塩、ヌクレオチド、モノカルボン酸塩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 (Paol, S. S. ら (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、基礎グルコース要件を有する多くの細胞型を提供して上皮及び内皮の障壁組織全体で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、インシュリン調整グルコース処理に責任があり、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、糖原病、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. and 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結合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、解糖中に乳酸で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を維持する際にクリティカルな役割を果たすと推定されている。最も特徴付けられた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⁺ 連鎖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 (L-乳酸からD-乳酸に対する立体選択性を含む) 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. and A. P. Halestrap (1993) Am. J. Physiol. 264 : C761-C782, Price, N. T. ら (1998) Biochem. J. 329 : 321-328、及びMartinelli, K. and I. Hagstrom (1993) J. Biotechnol. 30 : 339-350)。

【0006】

ATP結合カセット(ABC)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、大きなタンパク質、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。ABC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(NBD)及び2つの膜貫通ドメイン(MSD)を含む。NBDは、ATPを加水分解して輸送に必要なエネルギーを供給し、各MSD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

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、嚢胞性繊維症膜制御因子（CFTR）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NBD及びMSDが含まれる。これらの「半分子」は、小胞体ベースの主要組織適合複合体（MHC）ペプチド輸送系であるTap1及びTap2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、ABC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（CFTR、イオンチャネル）、副腎白質萎縮症（副腎白質萎縮症タンパク質、ALDP）、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、PMP70）及び高インシュリン性低血糖症（スルホニル尿素受容体、SUR）がある。別のABC輸送体である多剤耐性（MDR）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（Taglicht, D. and S. Michaelis (1998) Meth. Enzymol. 292 : 130-162)。

10

【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コファクターとして重要である。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コファクター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（Danks, D. M. (1986) J. Med. Genet. 23 : 99-106）。

20

【0008】

原形質膜を貫通する脂肪酸の輸送は、拡散、高容量、低親和性プロセスにより発生し得る。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発生するように見える。脂肪酸輸送タンパク質（FATP）は、4つの膜貫通セグメントを有する複合膜タンパク質であり、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。FATPの発現は、脂肪変換中に3T3-L1細胞において上方制御され、COS7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（Hui, T. Y. ら (1998) J. Biol. Chem. 273 : 27420-27429）。

30

【0009】

ミトコンドリア担体タンパク質は、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担体タンパク質、2-オキソグルタル酸／リンゴ酸担体、リン酸担体タンパク質、ピルビン酸担体、ジカルボン酸担体（リンゴ酸塩、コハク酸塩、フマル酸塩及びリン酸塩を輸送する）、トリカルボン酸担体（クエン酸塩及びリンゴ酸塩を輸送する）、及びグレーブス病担体タンパク質、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ブ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）。

40

【0010】

このクラスの輸送体には、複合ミトコンドリア膜を通過してプロトンを通洩せしめるよう

50

なミトコンドリア脱共役タンパク質も含まれ、従ってATP合成からの非結合酸化的リン酸化も含まれる。結果として、熱の形態でエネルギー散逸が行われる。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して意味づけられ、肥満症などの代謝病に対する薬剤の電位ターゲットとして提案されてきた (Ricquier, D. ら (1999) J. Int. Med. 245 : 637-642)。

【0011】

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、生成及び維持される。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依存性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活発に輸送する。依存性イオンチャネルは、制限された条件下でイオンの電気化学勾配下向きにイオンの受動流れを許容する。同時に、これらのタイプのイオンチャネルは、1) 神経細胞の軸索へ下向きの電気インパルス伝導、2) 濃度勾配に逆らった細胞への分子の輸送、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生成、維持及び利用する。

10

【0012】

イオン輸送体

イオン輸送体は、細胞の静止電位を生成及び維持する。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アーゼは、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アーゼ、 Ca^{2+} ATPアーゼ及び H^+ ATPアーゼがあり、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+} のサイトゾル濃度が低く、 K^+ のサイトゾル濃度が高くなるように静止電位分布を維持する責任がある。空胞(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。Vクラスイオン輸送体は、機能に必要なこれらのオルガネラのルーメン内で低pHを生成する責任がある。結合因子(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。

20

【0013】

P-ATPアーゼは、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細胞質領域を有する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アーゼは、2つの機能的ドメインから構成される。 V_1 ドメインは、ATP加水分解の責任を果たす周辺複合体であり、 V_2 ドメインは、膜を通過してのプロトンの転位置に責任を果たす複合体である。F-ATPアーゼは、構造的及び進化的にV-ATPアーゼに関連する。F-ATPアーゼ F_0 ドメインは、2つの膜貫通ドメインを含み、プロトン輸送に不可欠なTM2に単一埋没カルボキシル基を含むような高度に疎水性のタンパク質であるような、12個のコピーのcサブユニットを有する。V-ATPアーゼ V_0 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aサブユニットも含む (Forgacs, M. (1999) J. Biol. Chem. 274 : 12951-12954)。

30

40

【0014】

静止細胞の電位は、担体タンパク質及び依存性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。担体タンパク質は、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にリンクされ(シンポート)、それによって Na^+ の電気化学勾配下向きの動きが他の分子を濃度勾配に上向きに輸送するように駆動する。同様にして、心筋は、細胞からの Ca^{2+} の移動を

50

細胞への Na^+ の輸送にリンクさせる（アンチポート）。

【0015】

依存性イオンチャネル

依存性イオンチャネルは、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。依存性イオンチャネルは、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^+ 、 Ca^{2+} 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耐性）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の関連によって時折調節される。

10

【0016】

機械依存性または機械感覚性イオンチャネルは、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。伸長不活性化チャネル（SIC）は、最近ラットの肝臓からクローニングされたものである。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、 Ca^{2+} 及び Na^+ を共に伝導するような 1 群のチャネルに属している（スズキ， M. ら（1999） J. Biol. Chem. 274 : 6330-6335）。

20

【0017】

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、細胞内アミノ末端及びカルボキシ末端がある。 Na^+ 及び Ca^{2+} サブファミリーでは、このドメインは4回繰り返され、Fizze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、GYGトリペプチドはこの選択性に関与する（イシイ， T. M. ら（1997） Proc. Natl. Acad. Sci. USA 94 : 11651-11656）。

30

【0018】

電位依存性 Na^+ 及び K^+ チャネルは、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過渡変化から生じる。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^+ チャネルを開き、これによって脱分極を細胞の長さの下に成長させる。脱分極はまた、電位依存性カリウムチャネルを開く。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、この約1ミリ秒間に、膜電位と無関係に開き得ない不活性化状態にチャネルが自発的に変換される。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。

40

【0019】

電位依存性 Na^+ チャネルは、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、チャネルの機能的発現の増加、そのゲーティン

50

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. ら (1995) Cell 83: 433-442)。

【0020】

非電位依存性 Na^+ チャネルには、アミロライド感受性 Na^+ チャネル / *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ネルサブユニットは、細胞内に位置するアミノ末端及びカルボキシル末端を有するような、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、肝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^+ 再吸収に関与するような上皮性 Na^+ チャネル (*EnaC*) がある。*EnaC* における突然変異は、偽低アルドステロン症1型及びリドル症候群 (偽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ネルまたは酸感受性イオンチャネル (*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^+ 浸透性チャネルを形成する。これらのチャネルは、活性化のための酸性 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. and M. Lazdunski (1998) Curr. Opin. Neurobiol. 8: 418-424; Eglén, R. M. ら (1999) Trends Pharmacol. Sci. 20: 337-342)。

【0021】

K^+ チャ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^+ チャネルを制御し得る。非興奮性組織では、 K チャネルが、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^+ チャネルが静止膜電位の設定の責任を有する。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ネルが含まれる。ポンプは、活発に Na^+ を細胞へ流出させて K^+ を細胞から流入する ($\text{Na}^+:\text{K}^+$ の割合は3:2である)。原形質膜のイオンチャネルは、 K^+ 及び Cl^- が受動拡散により流れることを許容する。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 (Salkoff, L. and T. Jegla (1995) Neuron 15: 489-492)。

【0022】

Shaker 様スーパーファミリーのカリウムチャネルサブユニットはすべて6回膜貫通/1孔の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 チャネルを形成する。これらのポア形成サブユニットは、チャ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 *hether-a-go-go* 関連遺伝子 (*HERG*) などの遅延整流型チャネルや、電圧依存性 K チャネルがある (Curran, M. E. (1998) Curr. Opin. Biotechnol. 9: 565-572; Kaczarowski, G. J. and M. L. Garcia (1999) Curr. Opin. Chem. Biol. 3: 448-458)。

【0023】

K^+ チャネルの第2のスーパーファミリーは、内向き整流チャネル (*Kir*) で構成される。*Kir* チャネルは、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、四量体としても関連している。Kirファミリーには、腎尿細管疾患のバー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 (Doupnik, C. A. ら (1995) Curr. Opin. Neurobiol. 5 : 268-277、前出のCurran)。

【0024】

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. ら (1997) EMBO J 16 : 5464-5471)。

【0025】

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、 α_2 δ 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 α_2 δ 、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. ら (1998) J. Biol. Chem. 273 : 2361-2367、McCleskey, E. W. (1994) Curr. Opin. Neurobiol. 4 : 304-312)。

【0026】

カルシウムイオンチャンネルの一過性受容体ファミリー (Trp: transient receptor family) は、容量性カルシウム流入 (CCE) を仲介すると考えられている。CCEは細胞内への Ca^{2+} の流入であって、これにより、様々なホルモンや成長因子に応答したイノシトール三リン酸 (IP3) 及び他の物質の作用により枯渇した Ca^{2+} が再供給される。Trp及びTrp類似体は、ショウジョウバエから初めに同定されたものであり、S3からS6領域において電位依存性 Ca^{2+} チャンネルに類似している。このことは、Trp及び/または関連タンパク質が、哺乳動物CCC流入チャンネルを形成することを示唆している (Zhu, X. ら (1996) Cell 85 : 661-671 ; Boulay, G. ら (1997) J. Biol. Chem. 272 : 29672-29680)。melastatinは、マウス及びヒトの両方から単離されたものであり、黒色腫細胞におけるその発現は in vivo での黒色腫の攻撃性と逆の相互関係にある。そのヒトcDNA転写物は、Trpファミリーのメンバーにホモロジーを有する1533個のアミノ酸からなるタンパク質に対応する。melastatinのmRNAの発現状態と腫瘍の厚さを組み合わせて用いて、転移性疾患を発症するリスクの高いグループと低いグループの2つに患者を分けることが可能であると報告された (Duncan, L. M. et al (2001) J. Clin. Oncol. 19 : 568-576)。

【0027】

クロライドチャンネルは、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、電気化学平衡濃度以上で細胞内に蓄積する。尖端表面からの Cl^- を分泌することにより、ホルモンの刺激に応じて Na^+ 及び水が分泌性ルーメンに流入す

る。嚢胞性線維症膜コンダクタンス制御因子 (CFTR) は、ヒトにとって普通の致命的な遺伝病である嚢胞性繊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、ヌクレオチド結合部位が後に続く 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、取り除くのが困難である。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」及び荒廃的な「慢性閉塞性肺疾患」を生じさせる (Al-Awqati, Q. ら (1992) J. Exp. Biol. 172 : 245-266)。

【0028】

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。骨格筋で優性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、肝臓チャンネル CLC-5 における突然変異は、肝臓結石を引き起こす (Jentsch, T. J. (1996) Curr. Opin. Neurobiol. 6 : 303-310)。

【0029】

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. ら (1994) Curr. Opin. Neurobiol. 4 : 313-323)。

【0030】

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、カルシウムの流入は、脱分極中に K^+ チャンネルを開いて活動電位の大きさを調節する (前出のイシイら)。大コンダクタンス (BK) チャンネルは、頭脳及びその決定されたサブユニット組成から精製されてきた。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 A 28 アミノ酸伸長は、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、細胞内 N 末端及び C 末端を有する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む (前出の Kaczorowski, Vergara, C. ら (1998) Curr. Opin. Neurobiol. 8 : 321-329)。

【0031】

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

10

20

30

40

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化を与える (Zufall, F. ら (1997) *Curr. Opin. Neurobiol.* 7 : 404-412)。

10

【0032】

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. and 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. and D. S. Bredt (1998) *Cell* 93 : 495-498)。

20

【0033】

疾病の相関

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、Talent, G. M. ら (1994) *Ann. Intern. Med.* 120 : 218-226、及びChillon, M. ら (1995) *New Engl. J. Med.* 332 : 1475-1480)。

30

【0034】

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、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. and 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. and I. E. Scheffer (1999) *Curr. Opin. Neurology* 12 : 177182)。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に起因し

40

50

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

【0035】

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. and 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所麻酔薬は、神経障害性の痛みの治療において有用であった (前出のEnglen)。

10

【0036】

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のペプチドアンタゴニストKv1.3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チャネル遮断薬の考えを有効にすることがわかっている (Calahan, M. D. and K. G. Chandry (1997) Curr. Opin. Biotechnol. 8 : 749-756)。

20

【0037】

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、神経の疾患、筋疾患、免疫疾患、及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化合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

【0038】

(発明の要約)

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」、及び「TRICH-15」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。本発明の一実施態様では、(a) SEQ ID NO : 1-15からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO : 1-15からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO : 1-15からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO : 1-15とからなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択された単離されたポリペプチドを提供する。別法では、SEQ ID NO : 1-15のアミノ酸配列を含む単離されたポリペプチドを提供する。

40

【0039】

更に本発明は、(a) SEQ ID NO : 1-15からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO : 1-15からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと

50

、(c) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドをコードする単離されたポリヌクレオチドを提供する。別法では、このポリヌクレオチドは、SEQ ID NO: 1-15 からなる群から選択されたポリペプチドをコードする。更なる別法では、このポリヌクレオチドは、SEQ ID NO: 16-30 からなる群から選択される。

【0040】

更に、本発明は、(a) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドをコードするポリヌクレオチドと機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドを提供する。別法では、本発明は、この組換えポリヌクレオチドで形質転換された細胞を提供する。更なる別法では、本発明は、この組換えポリヌクレオチドを含む遺伝子組換え生物を提供する。

【0041】

また、本発明は、(a) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドの生産方法を提供する。この方法は、(a) このポリペプチドの発現に好適な条件下で、このポリペプチドをコードするポリヌクレオチドと機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養するステップと、(b) このように発現したポリペプチドを回収するステップとを含む。

【0042】

更に、本発明は(a) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドに特異的に結合する単離された抗体を提供する。

【0043】

更に、本発明は、(a) SEQ ID NO: 16-30 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(b) SEQ ID NO: 16-30 からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(c) 前記(a)のポリヌクレオチドに相補的なポリヌクレオチドと、(d) 前記(b)のポリヌクレオチドに相補的なポリヌクレオチドと、(e) 前記(a)-(d)のRNA等価物とで構成される群から選択された単離されたポリヌクレオチドを提供する。別法では、このポリヌクレオチドは、少なくとも60個の連続するヌクレオチドを含む。

【0044】

更に本発明は、(a) SEQ ID NO: 16-30 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(b) SEQ ID NO: 16-30 からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレ

オチド配列を含む天然のポリヌクレオチドと、(c)前記(a)のポリヌクレオチドに相補的なポリヌクレオチドと、(d)前記(b)のポリヌクレオチドに相補的なポリヌクレオチドと、(e)前記(a) - (d)のRNA等価物とで構成される群から選択されたポリヌクレオチド配列を有するサンプルにおける標的ポリヌクレオチドを検出する方法を提供する。この方法は、(a)前記サンプル内の標的ポリヌクレオチドと相補的な配列を含む少なくとも20個の連続するヌクレオチドを含むプローブと前記サンプルをハイブリダイズさせるステップであって、前記プローブと前記標的ポリヌクレオチドまたはその断片とでハイブリダイゼーション複合体が形成される条件下で、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズする、該ステップと、(b)前記ハイブリダイゼーション複合体の存在するか否かを検出し、存在する場合には必要に応じてその収量を測定するステップとを含む。別法では、前記プローブは、少なくとも60個の連続するヌクレオチドを含む。

10

【0045】

更に本発明は、(a)SEQ ID NO: 16-30からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(b)SEQ ID NO: 16-30からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(c)前記(a)のポリヌクレオチドに相補的なポリヌクレオチドと、(d)前記(b)のポリヌクレオチドに相補的なポリヌクレオチドと、(e)前記(a) - (d)のRNA等価物とで構成される群から選択されたポリヌクレオチド配列を有するサンプルにおける標的ポリヌクレオチドを検出する方法を提供する。この方法は、(a)ポリメラーゼ連鎖反応増幅法を用いて、前記標的ポリヌクレオチドまたはその断片を増幅するステップと、(b)増幅された前記標的ポリヌクレオチドまたはその断片が存在するか否かを検出し、存在する場合には必要に応じてその収量を測定するステップとを含む。

20

【0046】

更に本発明は、(a)SEQ ID NO: 1-15からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b)SEQ ID NO: 1-15からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c)SEQ ID NO: 1-15からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d)SEQ ID NO: 1-15からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択された有効量のポリペプチド及び好適な医薬用賦形剤を含む組成物を提供する。一実施例では、SEQ ID NO: 1-15からなる群から選択されたアミノ酸配列を含む組成物を提供する。本発明は更に、患者にこの組成物を投与することを含む、機能的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

【0047】

更に本発明は、(a)SEQ ID NO: 1-15からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b)SEQ ID NO: 1-15からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c)SEQ ID NO: 1-15からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d)SEQ ID NO: 1-15からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドのアゴニストとして有効な化合物をスクリーニングする方法を提供する。この方法は、(a)このポリペプチドを含むサンプルを化合物に曝露するステップと、(b)このサンプルのアゴニスト活性を検出するステップとを含む。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的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

【0048】

更に、本発明は、(a)SEQ ID NO: 1-15からなる群から選択されたアミノ酸

50

配列を含むポリペプチドと、(b) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドのアンタゴニストとして有効な化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを含むサンプルを化合物に曝露するステップと、(b) このサンプルのアンタゴニスト活性を検出するステップとを含む。別法では、本発明は、この方法によって同定されたアンタゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的 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

【0049】

更に本発明は、(a) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドに特異的に結合する化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを好適な条件下で少なくとも1つの試験化合物と結合させるステップと、(b) このポリペプチドとこの試験化合物との結合を検出して、このポリペプチドと特異的に結合する化合物を同定するステップとを含む。

20

【0050】

更に本発明は、(a) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸と少なくとも90%の同一性を有するアミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-15 からなる群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドの活性を調節する化合物をスクリーニングする方法を提供する。このスクリーニング方法は、(a) このポリペプチドを、その活性が許容される条件下で少なくとも1つの化合物と結合させるステップと、(b) この試験化合物の存在下でのこのポリペプチドの活性を評価するステップと、(c) この試験化合物の存在下でのこのポリペプチドの活性と、この試験化合物の非存在下でのこのポリペプチドの活性とを比較するステップとを含み、この試験化合物の存在下でのこのポリペプチドの活性の変化が、このポリペプチドの活性を調節する化合物の存在を示唆する。

30

【0051】

更に本発明は、SEQ ID NO: 16-30 からなる群から選択された配列を含む標的ポリヌクレオチドの発現を変化させるのに有効な化合物をスクリーニングする方法であって、(a) この標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、(b) この標的ポリヌクレオチドの発現の変化を検出するステップとを含む、該スクリーニング方法を提供する。

40

【0052】

本発明はさらに、試験化合物の毒性を評価する方法を提供する。この方法は、(a) 核酸を含む生体サンプルを前記試験化合物で処理するステップと、(b) 処理した前記生体サンプルの核酸をプローブとハイブリダイズするステップと、(c) ハイブリダイゼーション複合体の収量を測定するステップと、(d) 前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処理の生体サンプルにおけるハイブリダイゼーション複合体の収量とを比較するステップとを含み、前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量の差異が試験化合物の毒性を示唆する。この方法におけ

50

る前記プローブが、(1) SEQ ID NO: 16-30 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(2) SEQ ID NO: 16-30 からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(3) 前記(1)に相補的な配列を有するポリヌクレオチドと、(4) 前記(2)のポリヌクレオチドに相補的なポリヌクレオチドと、(5) 前記(1)-(4)のRNA等価物とで構成される群から選択されたポリヌクレオチドの少なくとも20個の連続するヌクレオチドを含む。また、前記ハイブリダイゼーションは、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われる。また、前記標的ポリヌクレオチドが、(1) SEQ ID NO: 16-30 からなる群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(2) SEQ ID NO: 16-30 からなる群から選択されたポリヌクレオチド配列と少なくとも90%の同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(3) 前記(1)のポリヌクレオチドに相補的なポリヌクレオチドと、(4) 前記(2)のポリヌクレオチドに相補的なポリヌクレオチドと、(5) 前記(1)-(4)のRNA等価物とで構成される群から選択される。別法では、標的ポリヌクレオチドは、その標的ポリヌクレオチドの断片である。

10

【0053】

(本発明の記載について)

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

20

【0054】

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

【0055】

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

30

【0056】

(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。

40

【0057】

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

【0058】

用語「アレル変異配列」は、TRICHをコードする遺伝子の別の形を指す。アレル変異配列は、核酸配列における少なくとも1つの変異によって生じ、変異mRNA若しくは変異ポリペプチドになり、これらの構造や機能は変わる場合もあれば変わらない場合もある

50

。ある遺伝子は、天然型のアレル変異配列が存在しないもの、1つ或いは多数存在するものがある。一般にアレル変異配列を生じる変異は、ヌクレオチドの自然な欠失、付加、或いは置換による。これらの各変異は、単独或いは他の変異と同時に起こり、所定の配列内で一回或いはそれ以上生じる。

【0059】

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アレル変異配列との不相当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形性を含む。コードされたタンパク質も変異され得り、サイレント変化を生じ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、負に荷電したアミノ酸にはアスパラギン酸及びグルタミン酸が含まれ、正に荷電したアミノ酸にはリシン及びアルギニンが含まれ得る。類似の親水性の値をもち極性非荷電側鎖を有するアミノ酸には、アスパラギン、グルタミン、セリン、トレオニンが含まれ得る。類似の親水性の値をもち非荷電側鎖を有するアミノ酸には、ロイシン、イソロイシン、バリン、グリシン、アラニン、フェニルアラニン及びチロシンが含まれ得る。

10

20

【0060】

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。

【0061】

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

30

【0062】

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。

【0063】

用語「抗体」は、抗原決定基と結合可能なF a b及びF（a b'）₂、及びそれらの断片、F v断片などの無傷の分子を指す。T R I C Hポリペプチドと結合する抗体は、抗体を免疫する小ペプチドを含む無傷の分子またはその断片を用いて作製可能である。動物（例えば、マウス、ラット、若しくはウサギ）を免疫化するのに使用されるポリペプチド或いはオリゴペプチドは、RNAの翻訳から、或いは化学的に合成可能であり、必要に応じて担体タンパク質と結合させることも可能である。ペプチドと化学的に結合した一般に用いられる担体は、ウシ血清アルブミン、チログロブリン、及びキーホールリンペットヘモニアン（KLH）を含む。次ぎに、この結合したペプチドを用いて動物を免疫化する。

40

【0064】

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

50

【0065】

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

【0066】

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。

【0067】

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

【0068】

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

【0069】

「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。伸長及び重複の両方によって構築されるコンセンサス配列もある。

【0070】

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

元の残基	保存的な置換
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu

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

10

一般に、保存されたアミノ酸置換の場合は、a) 置換された領域のポリペプチドの骨格構造、例えば、 β シートや α ヘリックス高次構造、b) 置換された部位の分子の電荷または疎水性、及び／または、c) 側鎖の大半が維持される。

【0071】

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。

20

【0072】

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

【0073】

30

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

【0074】

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

40

【0075】

SEQ ID NO: 16-30の断片は、例えば、この断片を得たゲノム内の他の配列とは異なる、SEQ ID NO: 16-30を明確に同定する固有のポリヌクレオチド配列の領域を含む。SEQ ID NO: 16-30のある断片は、例えば、ハイブリダイゼーションや増幅技術、またはSEQ ID NO: 16-30に関連ポリヌクレオチド配列か

50

ら区別する類似の方法に有用である。ある断片と一致する SEQ ID NO: 16-30 の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に測定できる。

【0076】

SEQ ID NO: 1-15 のある断片は、SEQ ID NO: 16-30 のある断片によってコードされる。SEQ ID NO: 1-15 のある断片は、SEQ ID NO: 1-15 を特異的に同定する固有のアミノ酸配列の領域を含む。例えば、SEQ ID NO: 1-15 のある断片は、SEQ ID NO: 1-15 を特異的に認識する抗体の開発における免疫原性ペプチドとして有用である。ある断片と一致する SEQ ID NO: 1-15 の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に決定できる。 10

【0077】

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

【0078】

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

【0079】

ポリヌクレオチド配列についての用語「パーセントの同一性」または「%の同一性」とは、標準化されたアルゴリズムを用いてアラインメントされる、2つ以上のポリヌクレオチド配列間の一致する残基の百分率のことである。このようなアルゴリズムは、標準化され再現できる方法で、2つの配列間のアラインメントを最適化するべく、配列にギャップを挿入して、より意味をもつ2つの配列間の比較を行うことができる。 20

【0080】

ポリヌクレオチド配列間の同一性のパーセントは、MEGALIGN version 3.12e 配列アラインメントプログラムに組込まれる CLUSTAL V アルゴリズムのデフォルトパラメータを用いて決定可能である。このプログラムは LASERGENE ソフトウェアパッケージの一部であり、分子生物学分析プログラム一式 (DNASTAR, Madison WI) である。この CLUSTAL V は、Higgins, D. G. 及び P. M. Sharp (1989) CABIOS 5:151-153、Higgins, D. G. 他 (1992) CABIOS 8:189-191 に記載されている。ポリヌクレオチド配列の対のアライメントの場合、デフォルトパラメータは、Ktuple=2、gap penalty=5、window=4、「diagonals saved」=4 と設定する。「重み付けされた」残基重み付け表が、デフォルトとして選択された。同一性のパーセントは、アラインメントされたポリヌクレオチド配列間の「類似性のパーセント」として CLUSTAL V によって報告される。 30

【0081】

別法では、National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S. F. 他 (1990) J. Mol. Biol. 215:403-410) が提供する、広く用いられている無料の配列比較アルゴリズム一式が、NCBI (Bethesda, MD) を含む幾つかのソース及びインターネット (<http://WWW.ncbi.nlm.nih.gov/BLAST/>) で入手可能である。この BLAST ソフトウェア一式には、既知のポリヌクレオチド配列と様々なデータベースの別のポリヌクレオチド配列とのアラインメントに用いられる「blastn」を含む、様々な配列分析プログラムが含まれる。「BLAST 2 Sequences」と呼ばれるツールが入手可能であり、2つのヌクレオチド配列の対を直接比較するために用いられる。「BLAST 2 Sequences」は、<http://WWW.ncbi.nlm.nih.gov/gorf/bl2.html> に 40 50

アクセスして、対話形式で利用できる。「BLAST 2 Sequences」ツールは、blastn 及び blastp (以下に記載)の両方に用いることができる。BLASTプログラムは、一般的には、デフォルトを設定するギャップ及び他のパラメータと共に用いられる。例えば、2つのヌクレオチド配列を比較する場合、ある者は、デフォルトパラメータに設定された「BLAST 2 Sequences」ツールVersion 2.0.12 (April-21-2000)でblastnを使用するであろう。そのようなデフォルトパラメータは、例えば、以下のようにする。

【0082】

Matrix: BLOSUM62

Reward for match: 1

10

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

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

20

【0083】

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

【0084】

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

30

【0085】

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

40

【0086】

別法では、NCBI BLASTソフトウェア一式が用いられる。例えば、2つのポリペプチド配列を対で比較をする場合、ある者は、デフォルトパラメータで設定された「BLAST 2 Sequences」ツールVersion 2.0.12 (Apr-21-2000)でblastpを使用するであろう。そのようなデフォルトパラメータは、例えば、以下のようにする。

【0087】

50

Matrix: BLOSUM62

Open Gap: 11 及び Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

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

10

【0088】

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

【0089】

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。

【0090】

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

20

30

【0091】

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

40

【0092】

本発明のポリヌクレオチド間の高いストリンジェンシーのハイブリダイゼーションでは、約0.2×SSC及び約1%のSDSの存在の下、約68℃で1時間の洗浄過程を含む。別法では、65℃、60℃、55℃、42℃の温度で行う。SSCの濃度は、約0.1%のSDSが存在の下、約0.1～2×SSCの範囲である。通常は、ブロッキング試薬を用いて非特異的なハイブリダイゼーションを阻止する。このようなブロッキング試薬には、例えば、約100～200 µg/mlの切断され変性したサケ精子DNAが含まれる。約35～50% v/vの濃度のホルムアミドなどの有機溶剤が、例えば、RNAとD

50

NAのハイブリダイゼーションなどの特定の場合に用いることができる。これらの洗浄条件の有用な改変は、当業者には周知である。特に高いストリンジェントな条件でのハイブリダイゼーションは、ヌクレオチド間の進化における類似性を示唆し得る。このような類似性は、それらのヌクレオチド及びコードされたポリペプチドが類似の役割を果たしていることを強く示唆する。

【0093】

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

10

【0094】

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。

【0095】

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

【0096】

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TRICH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TRICHのポリペプチド断片またはオリゴペプチド断片を含む。

20

【0097】

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

【0098】

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。

30

【0099】

用語「調節」は、TRICHの活性の変化を指す。例えば、調節によって、TRICH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。

【0100】

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

【0101】

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

40

【0102】

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

50

チレングリコール化して細胞における寿命を延ばし得る。

【0103】

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。

【0104】

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

【0105】

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

【0106】

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

【0107】

プライマーとして用いるオリゴヌクレオチドは、当分野で周知のプライマー選択用のコンピュータプログラムで選択される。例えば、OLIGO 4.06ソフトウェアは、それぞれが最大100ヌクレオチドまでのPCR用のプライマーの対の選択、及び32,000塩基までの入力ポリヌクレオチド配列から最大5,000ヌクレオチドまでの大きなポリヌクレオチド及びオリゴヌクレオチドの分析に有用である。類似のプライマー選択用プログラムには、能力を拡大する追加の機能が含まれている。例えば、PrimOUプライマー選択プログラム(Genome Center at University of Texas South West Medical Center, Dallas TXより入手可能)は、メガベース配列から特定のプライマーを選択できるため、ゲノムワイドスコープ(genome-wide scope)におけるプライマーの設計に有用である。Primer3プライマー選択プログラム(Whitehead Institute

10

20

30

40

50

／MIT Center for Genome Research, Cambridge MA 1より入手可能)によって、ユーザーは、プライマー結合部位として避けたい配列を指定できる「非プライミングライブラリ (mispriming library)」を入力できる。また、Primer 3は、特にマイクロアレイのオリゴヌクレオチドの選択に有用である(後の方の2つのプライマー選択プログラムのソースコードは、それぞれのソースから得ることができ、ユーザーのニーズを満たすように変更することもできる)。Primer Genプログラム(UK Human Genome Mapping Project Resource Centre, Cambridge UKより入手可能)は、多数の配列アラインメントに基づいてプライマーを設計するため、アラインメントされた核酸配列の最も保存された領域或いは最も保存されていない領域のどちらかとハイブリダイズするプライマーを選択することができる。従って、このプログラムは、固有及び保存されたオリゴヌクレオチドやポリヌクレオチドの断片の同定に有用である。上記した任意の選択方法で同定されたオリゴヌクレオチドやポリヌクレオチドの断片は、例えば、PCR法やシーケンシングプライマー、マイクロアレイエレメント、或いはサンプルの核酸の完全或いは部分的に相補的なポリヌクレオチドを同定する特定のプローブなどの、ハイブリダイゼーション技術に有用である。オリゴヌクレオチドの選択方法は、上記した方法に制限されるものではない。

10

【0108】

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

20

【0109】

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

30

【0110】

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

【0111】

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

40

【0112】

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

【0113】

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。

50

【0114】

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

【0115】

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。

10

【0116】

「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。

【0117】

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

20

【0118】

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

【0119】

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

30

【0120】

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

40

【0121】

特定の核酸配列の「変異配列」とは、デフォルトパラメータ設定の「BLAST 2 Se

50

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

10

【0122】

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

20

【0123】

(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、及び細胞増殖異常の診断、治療、及び予防に関する。

【0124】

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

30

【0125】

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

40

【0126】

表3は、本発明の各ポリペプチドの様々な構造的特徴を示す。列1および列2はそれぞれ

50

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

【0127】

表 2 および表 3 は共に、本発明の各ポリペプチドの特性を一覧にしたものであって、これらの特性が請求するポリペプチドが輸送体及びイオンチャネルであることを立証する。SEQ ID NO : 5 が、残基 M1 から残基 I1009 においてドブネズミのグルタミン酸受容体 δ -1 サブユニット (GenBank ID g475542) と 98% 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 を参照)。BLAST の確率スコアは 0.0 である。これは、探しているポリペプチド配列アラインメントが偶然の一致により得られる確率を示す。SEQ ID NO : 5 はまた、受容体ファミリーリガンド結合領域を含む。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 を参照) 。BLIMPS で得られたデータにより、SEQ ID NO : 5 が、グルタミン酸受容体であるということが裏付けられた。SEQ ID NO : 1-4 及び SEQ ID NO : 6-15 も同様の方法で解析してアノテーションを付けた。SEQ ID NO : 1-15 の解析のためのアルゴリズムおよびパラメータを表 7 に記載する。

【0128】

表 4 に示されているように、本発明の完全長ポリヌクレオチド配列は、cDNA 配列、またはゲノム DNA 由来のコード (エキソン) 配列、或いはこれらの 2 種類の配列のあらゆる組み合わせを用いて組み立てた。列 1 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 (Polynucleotide SEQ ID NO :) およびそれに対応するインサイトポリヌクレオチドコンセンサス配列番号 (Incyte Polynucleotide ID) を示す。列 3 は、塩基対における各ポリヌクレオチド配列の長さを示す。列 4 は、例えば、SEQ ID NO : 16-30 を同定するため、或いは SEQ ID NO : 16-30 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 から推定されるコード配列 (エキソン) 、および／または cDNA およびゲノム DNA の両方からなる群に対応する識別番号を示す。これらの配列を用いて本発明の完全長ポリヌクレオチド配列を組み立てた。表 4 の列 6 および列 7 はそれぞれ、それぞれの完全長配列に対する列 5 の cDNA 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。

【0129】

表 4 の列 5 に示されている識別番号は、具体的には、例えばインサイト cDNA およびそれらに対応する cDNA ライブラリの識別番号を示す。例えば、6431853H1 はインサイト cDNA 配列の識別番号であり、LUNGNON07 はそれが由来する cDNA ライブラリの識別番号である。cDNA ライブラリが示されていないインサイト cDNA は、プールされている cDNA ライブラリ (例えば、71456748V1) に由来する。または、列 5 の識別番号は、完全長ポリヌクレオチド配列の組み立てに用いた GenBank の cDNA すなわち EST (例えば、g5396013) の識別番号の場合もある。または、列 5 の識別番号は、Genscan 分析によって推定されるゲノム DNA のコード領域の場合もある。例えば、GBI. g6742097_0000003 fasta. edit ha は、Genscan 推定コード配列の識別番号であり、g674209

10

20

30

40

50

7 は、G e n s c a n に適用した配列の G e n b a n k の識別番号である。この G e n s c a n 推定コード配列は、配列を組み立てる前に編集する場合がある（実施例 4を参照）。または、列 5 の識別番号は、“ e x o n - s t i t c h i n g ” アルゴリズムによって c D N A および G e n s c a n 推定エキソンの両方からなる群の場合もある（実施例 5を参照）。または、列 5 の識別番号は、“ e x o n - s t r e t c h i n g ” アルゴリズムによって c D N A および G e n s c a n 推定エキソンの両方からなる群の場合もある（実施例 5を参照）。場合によっては、列 5 に示されている配列の範囲と重複するインサイト c D N A の範囲が得られ、最終的なコンセンサス配列が決定されるが、それに相当するインサイト c D N A の識別番号は示されていない。

【0130】

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

【0131】

本発明はまた、T R I C H の変異体も含む。好適な T R I C H の変異体は、T R I C H 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 T R I C H 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。

【0132】

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を提供する。特定の実施例において、本発明は、T R I C H をコードする S E Q I D N O : 16 - 30 からなる群から選択された配列を含むポリヌクレオチド配列を提供する。配列表に示した S E Q I D N O : 16 - 30 のポリヌクレオチド配列は、窒素系塩基のチミンがウラシルに置換され、糖鎖の背骨がデオキシリボースではなくリボースからなる等価 R N A 配列を含む。

【0133】

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 70 % のポリヌクレオチド配列同一性、或いは少なくとも 85 % のポリヌクレオチド配列同一性、更には少なくとも 95 % ものポリヌクレオチド配列同一性を有する。本発明の特定の実施形態は、S E Q I D N O : 16 - 30 からなる群から選択された核酸配列と少なくとも 70 % のポリヌクレオチド配列同一性、或いは少なくとも 85 % のポリヌクレオチド配列同一性、更には少なくとも 95 % ものポリヌクレオチド配列同一性を有する S E Q I D N O : 16 - 30 からなる群から選択された配列を含むポリヌクレオチド配列の変異配列を提供する。上記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。

【0134】

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、可能なコドン選択に基づいた組み合わせの選択によって作り出され得る可能なポリヌクレオチド配列の変異の全てが含まれ得る。これらの組み合わせは、天然の 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考慮する。

【0135】

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

10

20

30

40

50

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は有利となり得る。特定のコドンが宿主によって利用される頻度に基づいてコドンを選択して、ペプチドの発現が特定の真核細胞または原核宿主に発生する割合を高めることが可能である。コードされたア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。

【0136】

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。

10

【0137】

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、S E Q I D N O : 16-30 及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる（例えば、Wahl, G. M. 及び S. L. Berger (1987) *Methods Enzymol.* 152:399-407; and Kimmel, A. R. (1987) *Methods Enzymol.* 152:507-511. を参照）。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。

20

【0138】

当分野で周知の D N A のシーケンシング方法を用いて、本発明の何れの実施例も実行可能である。この方法には、例えば D N A ポリメラーゼ I のクレノウ断片、S E Q U E N A S E (U S B i o c h e m i c a l , C l e v e l a n d O H) 、T a q ポリメラーゼ (A p p l i e d B i o s y s t e m s) 、熱安定性 T7 ポリメラーゼ (A m e r s h a m , P h a r m a c i a B i o t e c h P i s c a t a w a y N J) 、或いは E L O N G A S E 増幅システム (L i f e T e c h n o l o g i e s , G a i t h e r s b u r g M D) にみられるような校正エキソヌクレアーゼとポリメラーゼとの組み合わせなどの酵素が用いられる。好ましくは、M I C R O L A B 2200 液体転移システム (H a m i l t o n , R e n o , N V) 、P T C 200 T h e r m a l C y c l e r 200 (M J R e s e a r c h , W a t e r t o w n M A) 及び A B I C A T A L Y S T 800 (P E B i o s y s t e m s) などの装置を用いて配列の準備を自動化する。次に、A B I 373 或いは 377 D N A シーケンシングシステム (P E B i o s y s t e m s) 、M E G A B A C E 1000 D N A シーケンシングシステム (M o l e c u l a r D y n a m i c s , S u n n y v a l e C A) または当分野で周知の他の方法を用いてシーケンシングを行う。得られた配列を当分野で周知の様々なアルゴリズムを用いて分析する（例えば、Ausubel, F. M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R. A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853. を参照）。

30

40

【0139】

当分野で周知の P C R 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、T R I C H をコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば制限部位 P C R 法を利用する 1 つの方法では、一般的なプライマー及び入れ子プライマー (n e s t e d p r i m e r) を用いてクローニングベクター内のゲノム D N A から未知の配列を増幅する（例えば、Sarkar, G. (1993) *PCR Methods Applic* 2:318-322 を参照）。逆 P C R 法を用いる別法では、広範な方向に伸長して環状化した鋳型から未知の配列を増幅するプライマーを用いる。この鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制

50

限断片に由来する（例えば、Triglia, T. ら（1988）Nucleic Acids Res 16:8186を参照）。キャプチャPCR法を用いる第3の方法は、ヒト及び酵母菌人工染色体DNAの既知の配列に隣接するDNA断片のPCR増幅を含む（例えば、Lagerstrom, M. 他（1991）PCR Methods Applic 1:111-119を参照）。この方法では、多数の制限酵素による消化及びライゲーションを用いて、PCRを行う前に未知の配列の領域の中に組換え二本鎖配列を挿入することが可能である。また、当分野で周知の別の方法を用いて未知の配列を得ることも可能である。（例えば、Parker, J. D. 他（1991）Nucleic Acids Res. 19:3055-3060を参照）。更に、PCR、ネスト化プライマー、PROMOTERFINDERライブラリ（Clontech, Palo Alto CA）を用いれば、ゲノムDNA内の歩行が可能である。この方法はライブラリをスクリーニングする必要がなく、イントロン／エキソン接合部を探すのに有用である。全てのPCR法をベースにした方法では、プライマーは、市販のOLIGO 4.06 Primer Analysis software（National Biosciences, Plymouth MN）或いは別の好適なプログラムなどを用いて、長さが22～30ヌクレオチド、GC含量が50%以上、約68℃～72℃の温度で鋳型に対してアニーリングするよう設計される。

【0140】

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

【0141】

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

【0142】

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のクローン化及び発現に利用可能である。

【0143】

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ン化、プロセッシング及び／または発現の調節が含まれるが、これらに限定されるものではない。ランダムな断片によるDNAの混合や遺伝子断片と合成オリゴヌクレオチドのPCR再組み立てを用いて、ヌクレオチド配列の組換えが可能である。例えば、オリゴヌクレオチド媒介性定方向突然変異誘発を利用して、新しい制限部位を生成する突然変異の導入、グリコシル化パターンの変更、コドン選択の変更、スプライスバリエーションの作製等が可能である。

【0144】

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

10

【0145】

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) Nucl. Acids Res. Symp. Ser 7:215-223; 及びHorn, T. 他 (1980) Nucl. Acids Res. Symp. Ser. 225-232を参照）。別法として、化学的方法を用いて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、ペプチド合成は種々の固相技術を用いて実行可能である（例えば、Creighton, T. (1984) Proteins. Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; Roberge, J. Y. ら (1995) Science 269:202-204を参照）。また、合成の自動化は例えばABI 431Aペプチドシンセサイザー (PE Biosystems) を用いて達成し得る。更にTRICHのアミノ酸配列または任意のその一部は、直接的な合成の際の変更、及び／または化学的方法を用いた他のタンパク質または任意のその一部からの配列との組み合わせにより、天然のポリペプチド配列を有するポリペプチドまたは変異体ポリペプチドを作製することが可能である。

20

30

【0146】

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

【0147】

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。このようなエレメントは、その長さ及び特異性が様々である。特定の開始シグナルによって、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開始コドン及び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

40

50

ルは必要なくなるであろう。しかしながら、コーディング配列或いはその断片のみが挿入された場合は、インフレームのATG開始コドンを含む外来性の翻訳調節シグナルが発現ベクターに含まれなければならない。外来性の翻訳要素及び開始コドンは、自然及び合成の様々なものから得ることが可能である。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である。(例えば、Scharf, D. 他 (1994) *Results Probl. Cell Differ.* 201-18-162. を参照)。

【0148】

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, 及び16-17章; 及び Ausubel, F. M. 他. (1995) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York NY, ch. 9章及び13章1-4章を参照)。

10

【0149】

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

20

30

40

【0150】

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、PB

50

L U E S C R I P T (S t r a t a g e n e , L a J o l l a C A) または p S P O R T 1 プラスミド (G I B C O B R L) 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 T R I C H をコードする配列をライゲーションすると l a c Z 遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。更に、これらのベクターを用いて、クローニングされた配列の i n v i t r o での転写、ジデオキシンスクリーニング、ヘルパーファージによる一本鎖の救出、入れ子状態の欠失を作り出すことが可能である（例えば、V a n H e e k e , G . 及び S . M . S c h u s t e r (1 9 8 9) J . B i o l . C h e m . 2 6 4 : 5 5 0 3 - 5 5 0 9 . を参照）。例えば、抗体の産生のためなどに多量の T R I C H が必要な場合は、T R I C H の発現をハイレベルで誘導するベクターが使用できる。例えば、強力に発現を誘発する T 5 または T 7 バクテリオファージプロモーターを含むベクターを使用できる。

10

【0151】

T R I C H の発現に酵母の発現系の使用が可能である。α 因子やアルコールオキシダーゼや P G H プロモーターなどの構成型或いは誘導型のプロモーターを含む多種のベクターが、酵母菌サッカロミセス・セレビジエまたは P i c h i a p a s t o r i s 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、A u s u b e l , 1 9 9 5 , 前出、B i t t e r , G . A . ら (1 9 8 7) M e t h o d s E n z y m o l . 1 5 3 : 5 1 6 - 5 4 4 , 及び S c o r e r , C . A . ら (1 9 9 4) B i o / T e c h n o l o g y 1 2 1 - 1 8 1 - 1 8 4 . を参照）。

20

【0152】

植物系も T R I C H の発現に使用可能である。T R I C H をコードする配列の転写は、例えば、C a M V 由来の 3 5 S 及び 1 9 S プロモーターなどのウイルスプロモーターが単独で、或いは T M V (例えば、C o r u z z i , G . ら (1 9 8 4) E M B O J . 3 : 1 6 7 1 - 1 6 8 0 ; B r o g l i e , R . ら (1 9 8 4) S c i e n c e 2 2 4 : 8 3 8 - 8 4 3 ; および W i n t e r , J . ら (1 9 9 1) R e s u l t s P r o b l . C e l l D i f f e r . 1 7 : 8 5 - 1 0 5 を参照) 由来のオメガリーダー配列と組み合わせる。これらの作製物は、直接の DNA 形質転換或いは病原体を介したトランスフェクションによって、植物細胞の中に導入可能である。（例えば、The M c G r a w H i l l Y e a r b o o k o f S c i e n c e a n d T e c h n o l o g y (1 9 9 2) M c G r a w H i l l N Y , p p . 1 9 1 - 1 9 6 を参照）。

30

【0153】

哺乳動物細胞では、多種のウイルスベースの発現系が利用され得る。アデノウイルスが発現ベクターとして用いられる場合、後発プロモーター及び 3 連リーダー配列からなるアデノウイルス転写物／翻訳複合体に T R I C H をコードする配列を結合し得る。ウイルスのゲノムの非必須の E 1 または E 3 領域への挿入により、感染した宿主細胞に T R I C H を発現する生ウイルスを得ることが可能である (L o g a n , J . 及び S h e n k , T . (1 9 8 4) P r o c . N a t l . A c a d . S c i . 8 1 : 3 6 5 5 - 3 6 5 9 を参照)。さらに、ラウス肉腫ウイルス (R S V) エンハンサーなどの転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させることが可能である。タンパク質を高レベルで発現させるために、S V 4 0 または E B V を基にしたベクターを用いることが可能である。

40

【0154】

ヒト人工染色体 (H A C) を用いて、プラスミドで発現しそれに含まれているものより大きな DNA の断片を供給可能である。治療のために約 6 k b ~ 1 0 M b の H A C s 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、H a r r i n g t o n , J . J . 他 (1 9 9 7) N a t G e n e t . 1 5 : 3 4 5 - 3 5 5 . を参照)。

50

【0155】

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。ベクターの導入の後、細胞を選択培地に移す前に、強化培地で約1～2日の間増殖させる。選択マーカーの目的は選択的な媒介物に対する抵抗性を与えるとともに、その存在により導入された配列を確実に発現する細胞の増殖及び回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に好適な組織培養技術を用いて増殖可能である。

10

【0156】

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

20

30

【0157】

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

40

【0158】

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。

50

【0159】

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イムノアッセイ（two-site, monoclonal-based immunoassay）が好ましいが、競合の結合アッセイも用いることもできる。これらのアッセイ及びその他のアッセイは、当分野では十分に知られている。（例えば、Hampton, R. 他. (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Sect. IV; Coligan, J. E. 他 Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York, NY; 及び Pound, J. D. (1990) Immunochemical Protocols, Humana Press, Totowa NJ）。

10

【0160】

種々の標識技術及び結合技術が当業者には周知であり、様々な核酸アッセイおよびアミノ酸アッセイに用いられ得る。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。当分野では周知であり市販されているこのようなベクターを、T7, T3, またはSP6などの好適なRNAポリメラーゼ及び標識されたヌクレオチドの追加によって、in vitroでのRNAプローブの合成に用いることができる。これらの方法は、例えば、Amersham Pharmacia Biotech及びPromega (Madison WI)、U. S. Biochemical Corp (Cleveland OH) が市販する種々のキットを用いて行うことができる。容易な検出のために用い得る好適なレポーター分子或いは標識には、基質、コファクター、インヒビター、磁気粒子、及び放射性核種、酵素、蛍光剤、化学発光剤、色素産生剤などが含まれる。

20

30

【0161】

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。形質転換された細胞から産生されたタンパク質が分泌されるか細胞内に留まるかは、使用されるその配列及び／またはそのベクターによる。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。

【0162】

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

40

50

【0163】

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

10

20

【0164】

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

【0165】

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。

30

【0166】

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、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。

40

【0167】

あるアッセイは、単に試験化合物をポリペプチドに実験的に結合させ、結合を、フルオロフォア、放射性同位体、酵素抱合体またはその他の検出可能な標識により検出することができる。例えば、このアッセイは、少なくとも1つの試験化合物を、溶液中の或いは固体支持物に固定されたTRICHと結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

50

の結合の検出及び測定を行うことができる。更にこのアッセイでは、細胞遊離剤、化学ライブラリまたは天然の生成混合物を用いて実施することができ、試験化合物は、溶液中で遊離させるか固体支持体に固定させる。

【0168】

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、TRICHが少なくとも1つの試験化合物と結合する、TRICHの活性が許容される条件下でアッセイを実施し、試験化合物の存在下でのTRICHの活性が試験化合物非存在下でのTRICHの活性と比較する。試験化合物の存在下でのTRICHの活性の変化は、TRICH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TRICHの活性に適した条件下でTRICHを含む*in vitro*または細胞遊離系と結合させてアッセイを実施する。これらアッセイの何れかにおいて、TRICHの活性を調整する試験化合物は間接的に結合することが可能であり、試験化合物と直接接触する必要がない。少なくとも1つから複数の試験化合物をスクリーニングすることができる。

10

【0169】

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

20

30

【0170】

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

【0171】

TRICH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、TRICHをコードするポリヌクレオチドの或る領域を動物ES細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について研究し、潜在的な医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばTRICHを乳汁内に分泌するなどTRICHを過剰に発現する哺乳動物近交系は、便利なタンパク質源となり得る（Janne, J. 他 (1998) Biotechnol. Annu. Rev. 4:55-74）。

40

【0172】

50

(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字列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の発現は肺の疾患や、無酸素症を引き起こす生理的病態に密接に関連するだけでなく、増殖組織、腫瘍組織、神経組織、副腎組織、脳腫瘍組織、胎児結腸組織、成人結腸組織、前立腺上皮組織、リンパ節癌組織、卵巢組織、膵臓組織、及び胎児脾臓組織に密接に関連する。従って、T R I C Hは、輸送障害、神経の疾患、筋疾患、免疫疾患、及び細胞増殖異常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。

10

【0173】

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、このような疾患には輸送障害、神経の疾患、筋疾患、免疫疾患、及び細胞増殖異常が含まれ、輸送障害の中には運動不能症、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、口峽炎、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、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、シスチン尿症、イミノグリシン尿症、H a r t u p病、ファンコニ病が含まれ、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、G e r s t m a n n - S t r a u s s l e r - S c h e i n k e r症候群を含むプリオン病 (p r i o n d i s e a s e) 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 (c e r e b e l l o r e t i n a l h e m a n g i o b l a s t o m a t o s i s)、脳3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 (S A D)、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、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

20

30

40

50

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チー、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）、口峽炎、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、および酸性マルターゼ欠損症（AMD、ポンペ病としても知られる）を含む筋障害が含まれ、免疫疾患の中には、炎症及び日光性角化症、後天性免疫不全症候群（AIDS）及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー（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、瘧疾、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病（MCTD）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、子宮の癌などが含まれる。

10

20

【0174】

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。

【0175】

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担体と共に患者に投与することも可能である。

30

【0176】

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。

【0177】

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、神経の疾患、筋疾患、免疫疾患、及び細胞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。

40

【0178】

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。

【0179】

別の実施例では、本発明の任意のタンパク質、アンタゴニスト、抗体、アゴニスト、相補的な配列、ベクターを別の好適な治療薬と組み合わせて投与することもできる。当業者は

50

、従来の医薬原理にしたがって併用療法で用いる好適な治療薬を選択可能である。治療薬との組み合わせにより、上に列記した種々の疾患の治療または予防に相乗効果をもたらし得る。この方法を用いて少ない量の各薬剤で医薬効果をあげることが可能であり、広範囲な副作用の可能性を低減し得る。

【0180】

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。治療用には、中和抗体（即ち、二量体の形成を阻害するもの）が特に好ましい。

10

【0181】

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。このようなアジュバントにはフロイントアジュバント、水酸化アルミニウムなどのミネラルゲルアジュバント、リゾレシチン、プルロニックポリオール、ポリアニオン、ペプチド、油性乳剤、キーホールリンペットヘモシニアン、及びジニトロフェノールなどの界面活性剤が含まれるが、これらに限定されるものではない。ヒトに用いられるアジュバントの中では、BCG (b a c i l l i C a T R I C H t t e - G u e r i n) 及び C o r y n e b a c t e r i u m p a r v u m が特に好ましい。

20

【0182】

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或いはペプチド、またはそれらの断片は、天然のタンパク質のアミノ酸配列の一部と同一であることが望ましい。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。

30

【0183】

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。これらの技術には、ハイブリドーマ技術、ヒトB細胞ハイブリドーマ技術、及びEBV-ハイブリドーマ技術が含まれるが、これらに限定されるものではない（例えば、K o h l e r, G. ら. (1975) N a t u r e 256:495-497; K o z b o r, D. ら. (1985) J. I m m u n o l. M e t h o d s 81-8-42; C o t e, R. J. ら. (1983) P r o c. N a t l. A c a d. S c i. 80:2026-2030; C o l e, S. P. ら. (1984) M o l. C e l l B i o l. 62:109-120を参照）。

40

【0184】

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、M o r r i s o n, S. L. 他. (1984) P r o c. N a t l. A c a d. S c i. 81-4851-4855; N e u b e r g e r, M. S. 他. (1984) N a t u r e 312:604-608; T a k e d a, S. ら. (1985) N a t u r e 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。関連する特異性を備えるが別のイディオタイプの組成の抗体は、ランダムな組み合わせの免疫グロブリンライブラリから鎖混合によって生成す

50

ることもできる（例えば、Burton D. R. (1991) Proc. Natl. Acad. Sci. 88:11120-3を参照）。

【0185】

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

【0186】

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、抗体分子のペプシン消化によって生成される $F(ab')_2$ 断片、及び $F(ab')$ 断片のジスルフィド架橋を減少させて生成される Fab断片が含まれる。別法では、所望の特異性を有するモノクローナル Fab断片の迅速且つ容易な同定が可能となるように、Fab発現ライブラリを作製することもできる（例えば、Huse, W. D. 他. (1989) Science 254:1275-1281を参照）。

【0187】

種々のイムノアッセイを用いてスクリーニングし、所望の特異性を有する抗体を同定する。確立された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる競合的な結合アッセイ、または免疫放射線アッセイの様々なプロトコルが、当分野では周知である。通常このようなイムノアッセイでは、TRICHとその特異的抗体との間で形成された複合体の測定が含まれる。二つの非干渉性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）。

【0188】

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 $10^9 \sim 10^{12} \text{ L/mol}$ 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 $10^6 \sim 10^7 \text{ L/mol}$ 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. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY)。

【0189】

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

【0190】

10

20

30

40

50

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、治療目的で使用する事ができる。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。

【0191】

治療に用いる場合、アンチセンス配列を好適な標的細胞に導入するのに好適な任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を発現する発現プラスミドの形で細胞内に送達することができる（例えば、Slater, J. E. 他（1998）*J. Allergy Clin. Immunol.* 102（3）：469-475；and Scanlon, K. J. 他（1995）9（13）：1288-1296.を参照）。また、アンチセンス配列は、例えばレトロウイルスやアデノ関連ウイルスベクター等のウイルスベクターを用いて細胞内に導入することもできる（例えば、Miller, A. D.（1990）*Blood* 76：271；Ausubel, 前出；Ucker, W. and W. Walther（1994）*Pharmacol. Ther.* 63（3）：323-347を参照）。その他の遺伝子送達機構には、リポソーム系、人工的なウイルスエンベロップ、及び当分野で周知のその他の系が含まれる（Rossi, J. J.（1995）*Br. Med. Bull.* 51（1）：217-225；Boado, R. J. 他（1998）*J. Pharm. Sci.* 87（11）：1308-1315；and Morris, M. C. 他（1997）*Nucleic Acids Res.* 25（14）：2730-2736.を参照）。

【0192】

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

【0193】

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な導入技術には、(i) 個々の細胞内へのDNAのマイクロインジェクション、(ii) 金粒子の打ち込み、(iii) リポソーム仲介性トランスフェクション、(iv) 受容体仲介性遺伝子導入、及び(v) DNAトランスポソン(Morgan, R. A. and W. F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Recipon (1998) Curr. Opin. Biotechnol. 9:445-450)の使用が含まれる。 10

【0194】

TRICHの発現に影響を及ぼし得る発現ベクターには、限定するものではないが、PCDNA 3.1、EPITAG、PRCCMV2、PREP、PVAXベクター(Invitrogen, Carlsbad CA)、PCMV-SCRIPT、PCMV-TAG、PEGSH/PERV(Stratagene, La Jolla CA)、PTET-OFF、PTET-ON、PTRE2、PTRE2-LUC、PTK-HYG(Clontech, Palo Alto CA)が含まれる。TRICH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-R-EXプラスミド(Invitrogen)に含まれている、テトラサイクリン調節性プロモーター(Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. U. S. A. 89:5547-5551; Gossen, M. 他 (1995) Science 268:1766-1769; Rossi, F. M. V. and H. M. Blau (1998) Curr. Opin. 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 30)、FK506/ラパマイシン誘導性プロモーター、またはRU486/ミフェプリストン誘導性プロモーター(Rossi, F. M. V. and H. M. Blau, 前出)、または(iii) 正常な個体に由来するTRICH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。

【0195】

市販のリポソーム形質転換キット(例えば、Invitrogenが販売しているPERFECT LIPID及びTRANSFECTION KIT)を用いれば、当業者は経験にそれほど頼らないでもポリヌクレオチドを培養中の標的細胞に導入することが可能である。別法では、リン酸カルシウム法(Graham, F. L. and A. J. Eb (1973) Virology 52:456-467)若しくは電気穿孔法(Neu 40mann, B. 他 (1982) EMBO J. 1:841-845)を用いて形質転換を行う。初代細胞にDNAを導入するためには、これらの標準的な哺乳動物トランスフェクションプロトコルを変更する必要がある。

【0196】

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 50

。レトロウイルスベクター（例えば、PFB及びPFBNEO）はStratagene社から入手可能であり、公表データ（Riviere, I. 他. (1995) Proc. Natl. Acad. Sci. U. S. A. 92:6733-6737）に基づいている。上記データを引用することをもって本明細書の一部とする。このベクターは、VSVg（Armentano, D. 他 (1987) J. Virol. 61:1647-1650; Bender, M. A. 他 (1987) J. Virol. 61:1639-1646; Adam, M. A. and A. D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. 他 (1998) J. Virol. 72:8463-8471; Zufferey, R. 他 (1998) J. Virol. 72:9873-9880）等の乱交雑エンベロープタンパク質若しくは標的細胞上の受容体に対する親和性を有するエンベロープ遺伝子を発現する好適なベクター産生細胞系（VPCCL）において増殖される。RIGGに付与された米国特許第5, 910, 434号（「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」）において、レトロウイルスパッケージング細胞系を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、ある細胞集団（例えば、CD4⁺T細胞）の形質導入、並びに形質導入した細胞を患者に戻す方法は、遺伝子治療の分野では周知であり、多数の文献に記載されている（Ranga, U. 他. (1997) J. Virol. 71:7020-7029; Bauer, G. 他 (1997) Blood 89:2259-2267; Bonyhadi, M. L. (1997) J. Virol. 71:4707-4716; Ranga, U. 他 (1998) Proc. Natl. Acad. Sci. U. S. A. 95:1201-1206; Su, L. (1997) Blood 89:2283-2290）。

【0197】

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は当分野では周知である。複製欠損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を脾臓の無損傷の脾島の中に導入するために可変性であることが証明された（Csete, M. E. 他. (1995) Transplantation 27:263-268）。使用できる可能性のあるアデノウイルスベクターが、米国特許第5, 707, 618号（「Adenovirus vectors for gene therapy」）に記載されており、引用することをもって本明細書の一部とする。アデノウイルスベクターについてはまた、Antinuzzi, P. A. 他 (1999) Annu. Rev. Nutr. 19:511-544; and Verma, I. M. and N. Somia (1997) Nature 18:389:239-242を参照し、引用することをもって本明細書の一部とする。

【0198】

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する。単純疱疹ウイルス（HSV）系のベクターは、HSV親和性の中枢神経細胞にTRICHを導入する際に特に重要である。ヘルペス系ベクターの作製及びパッケージングは当分野では周知である。複製適格性の単純疱疹ウイルス（HSV）I型系のベクターは、霊長類の眼にレポーター遺伝子を送達するために用いられてきた（Liu, X. 他 (1999) Exp. Eye Res. 169:385-395）。HSV-1ウイルスベクターの作製は、DeLucaに付与された米国特許第5, 804, 413号（Herpes simplex virus swains for gene transfer）に記載されており、引用することをもって本明細書の一部とする。米国特許第5, 804

、413号には、ヒト遺伝子治療を含む目的のために、好適なプロモーターのコントロールの下で、細胞に導入される少なくとも1つの内在性遺伝子を含むゲノムからなる組換えHSV d92についての記載がある。また上記特許には、ICP4、ICP27及びICP22のために除去される組換えHSV株の作製及び使用方法が開示されている。HSVベクターについては、Goins, W. F. 他 (1999) J. Virol. 73:519-532 and Xu, H. 他 (1994) Dev. Biol. 163:152-161を参照し、引用することをもって本明細書の一部とする。クローニングされたヘルペスウイルス配列の操作や、巨大ヘルペスウイルスのゲノムの異なった部分を含む多数のプラスミドをトランスフェクトした後の組換えウイルスの継代、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は当分野で周知の技術である。 10

【0199】

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）の生物学的な研究が広範に行われ、遺伝子伝達ベクター（gene transfer vector）がSFVゲノムに基づいていることが分かった（Garoff, H. and K.-J. Li (1998) Cunn. Opin. Biotech. 9:464-469）。αウイルスRNAの複製中に、通常はウイルスカプシドタンパク質をコードするサブゲノムRNAが作り出される。このサブゲノムRNAが完全長のゲノムRNAより高いレベルで複製されるため、酵素活性（例えばプロテアーゼ及びポリメラーゼ）を有するウイルスタンパク質に対してカプシドタンパク質が過剰に産生される。同様に、TRICHをコードする配列をαウイルスゲノムのカプシドをコードする領域に導入することによって、ベクター導入細胞において多数の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ウイルス感染は2～3日以内の細胞溶解に関係するが、シンドビスウイルス（SIN）の変異体を有するハムスターの正常な腎細胞（BHK-21）の持続的な感染を確立する能力は、αウイルスの溶解性の複製が遺伝子治療に適用できるように好適に変更することが可能であることを示唆している（Dryga, S. A. 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。 30
ある集団における細胞のサブセットの特定の形質導入には、形質導入する前に細胞のソーティングを必要とする場合がある。αウイルスの感染性cDNAクローンの操作、αウイルスcDNA及びRNAのトランスフェクション、並びにαウイルスの感染方法は当分野で周知である。

【0200】

例えば開始部位から約-10から約+10までの転写開始部位に由来するオリゴヌクレオチドを用いて、遺伝子の発現を阻害することが可能である。同様に、三重らせん塩基対合法を用いて阻害することができる。三重らせん構造は、二重らせんがポリメラーゼ、転写因子、または調節分子の結合のために十分に広がるのを阻止するため有用である。三重鎖DNAを用いる最近の治療の進歩は文献に記載されている（例えば、Gee, J. E. 40
ら. (1994) In: Huber, B. E. 及び B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, NYを参照）。相補的な配列またはアンチセンス分子もまた、転写物がリボソームに結合するのを阻止することによってmRNAの翻訳を阻止するように設計できる。

【0201】

酵素性RNA分子であるリボザイムは、RNAの特異的切断を触媒するために用いることができる。リボザイム作用の機構には、相補的な標的RNAへのリボザイム分子の配列特異性ハイブリダイゼーションが含まれ、ヌクレオチド鎖切断が続く。例えば、TRICHをコードする配列のヌクレオチド鎖切断を、特異的且つ効果的に触媒する組換え型のハン 50

マーヘッド型リボザイム分子が含まれる。

【0202】

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

【0203】

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

【0204】

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

【0205】

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。特定のポリヌクレオチドの発現の変化に有効な化合物には、限定するものではないが、特定のポリヌクレオチド配列と相互作用可能な非高分子化学物質、オリゴヌクレオチド、アンチセンスオリゴヌクレオチド、三重らせん形成オリゴヌクレオチド、転写因子やその他のポリペプチド転写調節因子が含まれる。有効な化合物は、ポリヌクレオチド発現のインヒビター或いはエンハンサーとして作用し、ポリヌクレオチドの発現を変化させ得る。従って、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用であり、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用であり得る。

【0206】

特定のポリヌクレオチドの発現の変化の有効性を調べるために、少なくとも1個から複数個の試験化合物をスクリーニングすることができる。試験化合物は、有効な化合物の化学修飾を含む当分野で周知の任意の方法で得ることができる。このような方法は、ポリヌクレオチドの発現を変化させる場合、一般に市販されている或いは専売の天然または非天然の化合物ライブラリから選択する場合、標的ポリヌクレオチドの化学的及び/または構造的な特性に基づいて化合物を合理的にデザインする場合、更に組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合に有効である。TRICHをコードするポリヌクレオチドを含むサンプルは、少なくとも1つの試験化合物に曝露して得る。サンプルには、例えば無傷細胞、透過化処理した細胞、*in vitro*細胞遊離系または再構成

10

20

30

40

50

生化学系が含まれ得る。T R I C Hをコードするポリヌクレオチドの発現における変化は、当分野で周知の任意の方法でアッセイする。通常、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の収量を定量し、その値が1或いは複数の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較における基準となり得る。試験化合物に曝露されるポリヌクレオチドの発現の変化が検出される場合は、ポリヌクレオチドの発現の変化に試験化合物が有効であることを示している。特定のポリヌクレオチドの発現の変化に有効な化合物を調べるために、例えば Schizosaccharomyces pombe 遺伝子発現系 (Atkins, D. 他 (1999) 米国特許第5, 932, 435号、Arndt, G. M. 他 (2000) *Nucleic Acids Res.* 28: E15) または HeLa 細胞等のヒト細胞株 (Clarke, M. L. 他 (2000) *Biochem. Biophys. Res. Commun.* 268: 8-13) を用いてスクリーニングする。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性を調べるための、各オリゴヌクレオチド (デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、及び修飾オリゴヌクレオチド) の組み合わせライブラリのスクリーニングを含む (Bruice, T. W. 他 (1997) 米国特許第5, 686, 242号、Bruice, T. W. 他 (2000) 米国特許第6, 022, 691号)。

【0207】

ベクターを細胞または組織に導入する多数の方法が利用でき、in vivo、in vitro、及び ex vivoでの使用に等しく適している。ex vivoでの治療の場合、患者から採取された肝細胞の中にベクターを導入して、自家移植で同じ患者に戻すためにクローニング増殖される。トランスフェクション、リボソーム注入またはポリカチオンアミノポリマーによる運搬は、当分野で周知の方法を用いて実行することができる (例えば、Goldman, C. K. 他. (1997) *Nature Biotechnology* 15: 462-66: を参照)。

【0208】

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

【0209】

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

【0210】

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

【0211】

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

要でなくなる。

【0212】

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

【0213】

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輸送を促進し得る。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。このようにして作製された融合タンパク質は、マウスモデル系の脳を含む全ての組織の細胞に形質導入されることが確認されている (Schwarze, S. R. 他 (1999) Science 285: 1569-1572)。

10

【0214】

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

【0215】

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

20

30

【0216】

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

【0217】

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

40

【0218】

(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

50

ト、インヒビターで治療を受けている患者をモニターするためのアッセイに用いられる。診断に有用な抗体は、治療のところで記載した方法と同じ方法で製剤される。T R I C H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T R I C Hを検出する方法が含まれる。これらの抗体は、修飾をして或いはしないで使用され、レポーター分子の共有結合性或いは非共有結合性の接着によって標識化され得る。当分野で周知の種々のレポーター分子が用いられるが、その内の幾つかは上記した。

【0219】

T R I C Hを測定するためのE L I S A, R I A, 及びF A C S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T R I C H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T R I C H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T R I C Hに対する抗体とを結合させることによって決定する。標準的な複合体形成の量は、測光法 (p h o t o m e t r i c) などの種々の方法で定量され得る。被験者のT R I C Hの発現の量、制御及び疾患、生検組織からのサンプルが基準値と比較される。基準値と被験者との間の偏差が、診断の指標となる。

【0220】

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

【0221】

一実施形態では、T R I C H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P C Rプローブを用いたハイブリダイゼーションによって、T R I C H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T R I C Hをコードする自然界の配列のみを同定するかどうか、或いはアレルや関連配列コードする自然界の配列のみを同定するかどうかによって決まるであろう。

【0222】

プローブはまた、関連する配列の検出に利用され、T R I C H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、D N AあるいはR N Aが可能であり、S E Q I D N O: 16-30の配列、或いはT R I C H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。

【0223】

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

【0224】

T R I C Hをコードするポリヌクレオチド配列を用いて、T R I C Hの発現に関連する疾

患を診断することが可能である。限定するものではないが、このような疾患には輸送障害、神経の疾患、筋疾患、免疫疾患、及び細胞増殖異常が含まれ、輸送障害の中には運動不能症、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、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、シスチン尿症、イミノグリシン尿症、Hartup病、ファンコニ病が含まれ、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、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、および酸性マルターゼ欠損症(AMD、ポンペ病としても知られる)を含む筋障害が含まれ、免疫疾患の中には、炎症及び日光性角化症、後天性免疫不全症候群(AIDS)及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー(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

症、骨関節炎、骨粗しょう症、膵炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病（MCTD）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、子宮の癌などが含まれる。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、ドットプロット法、或いはその他の膜系の技術、PCR法、ディップスティック（dipstick）、ピン（pin）、ELISA式アッセイ、及び変異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。このような質的或いは量的方法は、当分野では周知である。

【0225】

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

【0226】

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

【0227】

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

【0228】

癌では、個体からの生体組織における異常な量の転写物が、疾患の発生の素因を示し、また実際に臨床的症状が出る前に疾患を検出する方法を提供することが可能である。この種のより明確な診断により、医療の専門家が予防方法或いは積極的な治療法を早くから利用して、癌の発生または進行を防ぐことが可能となる。

【0229】

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。このようなオリゴマーは、化学的な合成、酵素を用いた生成、或いは *in vitro* で生成され得る。オリゴマーは、好ましくは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オチド

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

【0230】

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

【0231】

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

【0232】

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

【0233】

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、上記のようにタンパク質間相互作用、薬剤-標的相互作用及び遺伝子発現プロファイルをモニタリング及び測定することが可能である。

【0234】

特定の実施例は、或る組織または細胞型の転写イメージを生成する本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞型により遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所定の条件下で所定の時間に発現した遺伝子の数及び相対存在量を定量することにより分析される（Seilliamer 他、米国特許第 5, 840, 484 号の "Comparative Gene Transcript Analysis" を参照。この特許に言及することを以って本明細書の一部とする）。従って、特定の組織または細胞型の転写物または逆転写物の全てに本発明のポリヌクレオチドまたはその相補配列をハイブリダイズすることにより、転写イメージが生成され得る。或る実施例では、本発明のポリヌクレオチドまたはその相補配列がマイクロアレイ上に複数のエレメントのサブセットを構成するハイスループット型でハイブリダイゼーションさせる。結果として得られる転写イメージは、遺伝子活性のプロファイルとなり得る。 10

【0235】

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

【0236】

本発明のポリヌクレオチドの発現プロファイルを示す転写イメージはまた、合成化合物または天然化合物の毒性試験のみならず、in vitro モデル系及び薬剤の前臨床評価 20 に関連して使用され得る。全ての化合物は、作用及び毒性の機構を示唆する、頻繁に分子フィンガープリント若しくは毒性シグネチャ（signature）と称されるような特徴的な遺伝子発現パターンを引き起こす（Nuwaysir, E. F. 他（1999）Mol. Carcinog. 24:153-159、Steiner, S. and N. L. Anderson（2000）Toxicol. Lett. 112-113:467-471、また言及することを以って本明細書の一部とする）。試験化合物が、毒性を有する既知の化合物のシグネチャと同一のシグネチャを有する場合には、毒性特性を共有している可能性が高い。フィンガープリンまたはシグネチャが、より多くの遺伝子及び遺伝子ファミリーからの発現情報を含んでいれば、より有用かつ正確になる。理想としては、発現のゲノム全域にわたって測定し、最高品質のシグネチャを提供すること 30 である。任意の試験化合物によっても発現が変化しない遺伝子も同様に重要である。それは、これらの遺伝子の発現レベルを用いて残りの発現データを標準化することができるためである。標準化処理は、異なる化合物で処理した後の発現データの比較に有用である。毒性シグネチャのエレメントへの遺伝子機能を割り当てることは毒性機構の解明に役立つが、毒性の予測につながるシグネチャの統計的な一致には遺伝子機能の知識は必要ではない（例えば2000年2月29日にNational Institute of Environmental Health Sciencesより発行されたPress Release 00-02を参照されたい。これについては <http://www.niehs.nih.gov/oc/news/toxchip.htm> で入手可能である）。従って、毒性シグネチャを用いる毒性スクリーニングにおいて、全ての発現した遺伝子配列 40 を含めることは重要でありまた望ましいことである。

【0237】

一実施例では、試験化合物の毒性は、核酸を含有する生体サンプルをその試験化合物で処理して評価する。処理した生体サンプル中で発現した核酸は、本発明のポリヌクレオチドに特異的な1若しくは複数のプローブでハイブリダイズさせ、それによって本発明のポリヌクレオチドに対応する転写レベルを定量することができる。処理した生体サンプル中の転写レベルを、非処理生体サンプル中のレベルと比較する。両サンプルの転写レベルの差が、処理されたサンプル中で試験化合物が引き起こす毒性反応を示唆する。

【0238】

別の実施例は、本発明のポリペプチド配列を用いて組織または細胞型のプロテオームを分 50

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

10

20

【0239】

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

30

【0240】

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

40

【0241】

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

【0242】

別の実施例では、試験化合物の毒性は、タンパク質を含む生体サンプルをその試験化合物

50

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

【0243】

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

【0244】

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

【0245】

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

【0246】

染色体標本のin situ ハイブリダイゼーション、及び確定した染色体マーカーを用いた結合分析などの物理的マッピング技術を用いて、遺伝子地図を拡張することもできる。マウスなどの別の哺乳動物の染色体上に遺伝子を配置させることにより、たとえ正確なヒト染色体の位置が分かっていなくても、関連するマーカーが明らかになる場合が多い。この情報は、位置クローニング或いは別の遺伝子発見技術を用いて遺伝的疾患の研究をして

いる研究者にとって価値がある。疾患や症候群に關与する1つ或いは複数の遺伝子の位置が、例えば血管拡張性失調症の11q22-23などの特定の遺伝子領域に遺伝子結合によって大まかに決定されると、その領域に対するどの配列マッピングも、さらなる調査のための關連する遺伝子或いは調節遺伝子を表す（例えば、G a t t i, R. A. 他による（1988）N a t u r e 336:577-580を参照）。また、目的の本発明のヌクレオチド配列を用いて、正常者、保有者、即ち感染者の間の、転位置、反転などによる染色体位置の違いを検出することもある。

【0247】

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。このようなスクリーニングに用いる断片は、溶液に遊離、固体支持物に固定、細胞の表面上に保持、或いは細胞内に存在する。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

【0248】

薬剤スクリーニングに用いる別の方法は、目的のタンパク質に対して、好適な結合親和性を有する化合物のスクリーニング処理能力を高めるために用いられる（例えば、G e y s e n, 他による（1984）P C T出願番号 W O 84/03564を参照）。この方法では、相当な数の異なる小さな試験用化合物が、プラスチックピン或いは他の基板の上に合成される。試験用化合物は、T R I C H、或いはその断片と反応してから洗浄される。次ぎに、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、ペプチドを捕らえ、固体支持物に固定することもできる。 20

【0249】

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

【0250】

別の実施例では、発展途上の分子生物学技術にT R I C Hをコードするヌクレオチド配列を用いて、限定はされないが、現在知られているトリプレット暗号及び特異的な塩基対相互作用などのヌクレオチド配列の特性に依存する新しい技術を提供することができる。 30

【0251】

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

【0252】

前述した及び以下に記載する全ての特許出願、特許、刊行物、特に米国特許出願第60/195,595号、同第60/196,872号、同第60/199,020号、同第60/200,552号、及び同第60/202,348号に言及することをもって本明細書の一部とする。 40

【0253】

（実施例）

1 cDNAライブラリの作製

インサイトcDNA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）に含まれているcDNAライブラリに由来し、表4の列5に示されている。ある組織をグアニジニウムイソチオシアネート溶液においてホモジナイズして溶解する一方、別の組織を、フェノールにおいて、或いはグアニジニウムイソチオシアネート及びフェノールの単相溶液であるT R I Z O L（L i f e T e c h n o l o g i e s）などの変性剤の好適な混合液においてホモジナイズして溶解した。こ 50

の溶解物を塩化セシウムにおいて遠心分離によって、或いはクロロホルムで抽出した。RNAは、イソプロパノール或いは酢酸ナトリウムとエタノールのどちらか、或いは別の一般的な方法でこの溶解物から沈殿させた。

【0254】

RNAの純度を高めるためにRNAのフェノールによる抽出及び沈殿を必要な回数繰り返した。場合によっては、DNA分解酵素でRNAを処理する。殆どのライブラリでは、オリゴd(T)連結常磁性粒子(Promega)またはOLIGOTEXラテックス粒子(QIAGEN, Valencia CA)、OLIGOTEX mRNA精製キット(QIAGEN)を用いてポリ(A+) RNAを単離した。別法では、POLY(A) PURE mRNA精製キット(Ambion, Austin TX)などの別のRNA単離キットを用いて組織溶解物から直接単離した。 10

【0255】

ある場合には、Stratagene社にRNAを提供し、Stratagene社が対応するcDNAライブラリを作製した。そうでない場合は、UNIZAPベクターシステム(Stratagene)またはSUPERSCRIPT プラスミドシステム(Life Technologies)を用いて当分野で周知の推奨方法または類似の方法でcDNAを合成してcDNAライブラリを作製した。(例えば、Ausubel, 1997, 前出, ユニット5.1-6.6を参照)。逆転写は、オリゴd(T)またはランダムプライマーを用いて開始させた。合成オリゴヌクレオチドアダプターを二本鎖cDNAに結合させてから、好適な1或いは複数の制限酵素でcDNAを消化した。殆どのライブラリでは、SEPHACRYL S1000、SEPHAROSE CL2B、SEPHAROSE CL4Bカラムクロマトグラフィー(Amersham Pharmacia Biotech)、またはアガロースゲル電気泳動法によってcDNAの大きさ(300~1000bp)を選択した。PBLUESCRIPTプラスミド(Stratagene)、pSPORT1プラスミド(Life Technologies)、pcDNA2.1プラスミド(Invitrogen Carlsbad CA)、PBK-CMVプラスミド(Stratagene)、pINCYプラスミド(Incyte Pharmaceuticals, Palo Alto CA)、またはそれらの誘導体などの好適なプラスミドのポリリンカーの適合性制限酵素部位にcDNAを結合させた。この組換えプラスミドを、Stratagene社のXL1-Blue, XL1-BlueMRF 30、SOLR、またはLife Technologies社のDH5αまたはDH10B、ELECTROMAX DH10Bを含むコンピテント大腸菌細胞に導入し形質転換した。

【0256】

2 cDNAクローンの単離

上記実施例1に記載したように得たプラスミドは、UNIZAPベクターシステム(Stratagene)或いは細胞溶解を利用した*in vivo*切除によって宿主細胞から回収した。プラスミドの精製には、MagicまたはWIZARD Minipreps DNA精製システム(Promega)、AGTC Miniprep精製キット(Edge Biosystems, Gaithersburg MD)、QIAGEN社のQIAWELL 8 Plus、QIAWELL 8 Plus Plasmid、またはQIAWELL 8 Ultra Plasmid 精製システム、またはREAL Prep 96 プラスミドキットの内の少なくとも1つを用いた。沈殿させた後、0.1 mlの蒸留水に再懸濁して、凍結乾燥して或いは凍結乾燥しないで4℃で保管した。 40

【0257】

別法では、ハイスルーブットの直接結合PCR法によって宿主細胞溶解物からプラスミドDNAを増幅した。(Rao, V. B. (1994) Anal. Biochem. 216:1-14)。宿主細胞の溶解及び熱サイクリングステップは、1つの反応混合液で行った。サンプルを処理してから384-ウェルプレートに移して保管し、増幅したプラスミドDNAの濃度を、PICOGREEN色素(Molecular Probes 50

, Eugene OR) 及び Fluoroskan II 蛍光スキャナ (Labsystems Oy, Helsinki, Finland) を用いて蛍光定量的に測定した。

【0258】

3 シークエンシング及び分析

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

【0259】

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

ftware Engineering, South San Francisco CA) および LASERGENE ソフトウェア (DNASTAR) を用いて分析した。ポリヌクレオチド配列およびポリペプチド配列のアラインメントを、アラインメントした配列間のパーセント同一性も計算する MEGALIGN マルチシーケンシングアラインメントプログラム (DNASTAR) に組み込まれた CLUSTAL アルゴリズムによって指定されたデフォルトパラメータを用いて作成した。

【0260】

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

10

【0261】

完全長のポリヌクレオチド配列およびポリペプチド配列の組み立て及び分析に用いる上記のプログラムは、SEQ ID NO: 16-30 のポリヌクレオチド配列断片の同定にも利用できる。ハイブリダイゼーション及び増幅技術に有用である約20～約4000ヌクレオチドの断片を表4の列4に示した。

【0262】

20

4 ゲノムDNAからのコード配列の同定および編集

推定上の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、gbprri や gbhtg) において Genscan 遺伝子同定プログラムを実行して初めに同定された。Genscan は、様々な生物に由来するゲノムDNA配列を分析するための汎用遺伝子同定プログラムである (Burge, C. および S. Karlin (1997) J. Mol. Biol. 268 : 78-94、Burge, C. および S. Karlin (1998) Curr. Opin. Struct. Biol. 8 : 346-354 を参照)。このプログラムは推定エキソンを連結して、メチオニンから停止コドンまで伸長した組み立て cDNA 配列を構築する。Genscan により得られる配列は、FASTA データベースのポリヌクレオチド配列およびポリペプチド配列になる。Genscan によって一回で解析できる配列の最大長さは30kbに設定されている。これらのGenscan推定cDNA配列の内、どの配列が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アノテーションが付けられたインサイト cDNA 配列に対する相同性を基に同定された。次に、これらの選択されたGenscan推定配列を、BLAST解析を用いて genept および gbprri 公共データベースの配列と比較した。必要に応じて、Genscan推定cDNA配列を、geneptにおいてBLASTで最もヒットした配列と比較して、Genscan推定配列における余分なエキソンや省いてしまったエキソンなどのエラーを修正し、編集した。BLAST解析を用いてGenscan推定cDNA配列を含むインサイト cDNA または公共の cDNA を見つけ出すことにより、転写の証拠が得られる。インサイト cDNA が Genscan 推定 cDNA 配列を含む場合、この情報を用いて Genscan 推定配列を修正或いは確認できる。完全長ポリヌクレオチド配列は、実施例3に説明した組み立て方法で Genscan 推定コード配列とインサイト cDNA および／または公共の cDNA 配列を組み立てて作製した。別法では、完全長ポリヌクレオチド配列は、その全てが編集した或いは未編集の Genscan 推定コード配列から作製した。

30

40

【0263】

5 cDNA 配列データを用いるゲノム配列データの組み立て ステッチ配列 (Stitched Sequence)

50

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

【0264】

ストレッチ配列 (Stretched Sequence)

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

【0265】

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: 16-30 を組み立てるために用いた配列を、BLAST 及び Smith-Waterman アルゴリズムを用いて、インサイト LIFESEQ データベース及び公共のドメインデータベースの配列と比較した。SEQ ID NO: 16-30 と一致するこれらのデータベースの配列を、Phrap (表 7) などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。Stanford Human Genome Center (SHGC)、Whitehead Institute for Genome Research (WIGR) 及び Genethon などの公共の情報源から入手できる放射線ハイブリッド (radiation hybrid) 及び遺伝子マッピングのデータを用いて、クラスター化した配列がすでにマッピングされているかを調べる。クラスターにマッピングされた配列が含まれている場合は、そのクラスターの全ての配列 (特定の SEQ ID NO を含む) をそのマッピング位置に割り当てた。

【0266】

遺伝子地図の位置は、範囲、区間、またはヒト染色体によって表される。センチモルガン

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

【0267】

10

7 ポリヌクレオチド発現の分析

ノーザン分析は、遺伝子の転写物の存在を検出するために用いられる実験用技術であり、特定の細胞種或いは組織からのRNAが結合されている膜への標識されたヌクレオチド配列のハイブリダイゼーションを伴う（例えば、Sambrook, 前出, 7章; 及び Ausubel, F. M. 他、前出, 4章及び16章を参照）。

【0268】

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

20

【0269】

【数1】

(BLAST スコア×配列一致率)

5× (長さ(配列1), 長さ(配列2))の最小値

【0270】

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

30

40

【0271】

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

50

各ヒト組織は、癌、細胞系、発生、炎症、神経、外傷、心血管、プール (p o o l e d) などの1つの疾患／症状のカテゴリに分類され、各カテゴリにおけるライブラリ数をカウントし、その合計数を全カテゴリのライブラリ数で除す。得られるパーセンテージは、T R I C H をコードする c D N A の疾患特異的な発現を反映する。c D N A 配列および c D N A ライブラリ／組織の情報は、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) から得ることができる。

【0272】

8 T R I C H をコードするポリヌクレオチドの伸長

完全長のポリヌクレオチド配列は、完全長分子の好適な断片から設計したオリゴヌクレオチドプライマーを用いてその完全長分子の好適な断片を伸長して作製した。一方のプライマーは既知の断片の5'の伸長を開始するために合成し、他方のプライマーは既知の断片の3'の伸長を開始するために合成した。開始プライマーは、O L I G O 4 . 0 6 ソフトウェア (N a t i o n a l B i o s c i e n c e s) 或いは他の適切なプログラムを用いて、約22個から約30個のヌクレオチドの長さで約50%以上のGC含量を有し、かつ約68～72℃の温度で標的配列にアニールするように設計した。ヘアピン構造及びプライマー-プライマー二量体が生じないようにヌクレオチドを伸長した。

【0273】

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

【0274】

当分野で既知の方法を利用した P C R 法で高い忠実度で増幅した。P C R は P T C - 2 0 0 t h e r m a l c y c l e r (M J R e s e a r c h , I n c .) 用いて96ウェルブロックプレートで行った。反応混合液は、鋳型DNA及び200 nmol の各プライマー、Mg²⁺と(NH₄)₂SO₄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 (A m e r s h a m P h a r m a c i a B i o t e c h) 、E L O N G A S E 酵素 (L i f e T e c h n o l o g i e s) 、P f u DNAポリメラーゼ (S t r a t a g e n e) を含む。プライマーの組、P C I A と P C I B に対して以下のパラメータで増幅を行った。

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

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

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

【0275】

各ウェルのDNA濃度は、1X TE及び0.5 μlの希釈していないPCR産物に溶解した100 μlのP I C O G R E E N 定量試薬 (0.25% (v/v) P I C O G R E E N ; M o l e c u l a r P r o b e s , E u g e n e O R) を不透明な蛍光光度計プレート (C o m i n g C o s t a r , A c t o n M A) の各ウェルに分配してDNAが試薬と結合できるようにして測定する。このプレートをF l u o r o s k a n I I (L a b s y s t e m s O y , H e l s i n k i , F i n l a n d) でスキャンして

、サンプルの蛍光を計測してDNAの濃度を定量化する。反応混合物の5～10 μ lのアリコットを1%のアガロースミニゲル上での電気泳動によって解析し、何れの反応物が配列を伸長することに成功したかを決定する。

【0276】

伸長したヌクレオチドを脱塩及び濃縮してから384ウェルプレートに移し、CviJIコレラウィルスエンドヌクレアーゼ (Molecular Biology Research, Madison WI) で消化し、pUC18ベクター (Amersham Pharmacia Biotech) に再連結する前に音波処理またはせん断を行った。ショットガンシーケンシングのために、消化したヌクレオチドを低濃度 (0.6～0.8%) のアガロースゲル上に分離して断片を切断し、寒天をAgar ACE (Promega) で消化した。T4リガーゼ (New England Biolabs, Beverly MA) を用いて伸長したクローンをpUC18ベクター (Amersham Pharmacia Biotech) に再連結し、Pfu DNAポリメラーゼ (Stratagene) で制限部位の伸び出しを処理してコンピテント大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りとって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。

【0277】

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

【0278】

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

【0279】

9 個々のハイブリダイゼーションプローブの標識化及び使用法

SEQ ID NO: 16-30から導き出されたハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記すが、より大きなcDNAフラグメントの場合でも基本的に同じ手順を用いる。オリゴヌクレオチドを、OLIGO4.06ソフトウェア (National Bioscience) のような最新式のソフトウェアを用いてデザインし、50 pmolの各オリゴマーと、250 μ Ciの [γ - 32 P] アデノシン三リン酸 (Amersham, Chicago, IL) 及びT4ポリヌクレオチドキナーゼ (DuPont NEN, Boston MA) とを組み合わせて用いることにより標識する。標識されたオリゴヌクレオチドを、SEPHADEX G-25超精細排除デキストランビードカラム (Amersham Pharmacia Biotech) を

用いて実質的に精製する。毎分 10^7 カウントの標識されたプローブを含むアリコットを、次のエンドヌクレアーゼ、Ase I、Bgl II、Eco RI、Pst I、Xba I 或いは Pvu II (DuPont NEN) の 1 つを用いて切断したヒトゲノム DNA の典型的な膜ベースのハイブリダイゼーション解析において用いる。

【0280】

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

10

【0281】

10 マイクロアレイ

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

20

【0282】

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

40

【0283】

組織または細胞サンプルの調製

グアニジウムチオシアネート法を用いて組織サンプルから全 RNA を単離し、オリゴ (dT) セルロース法を用いてポリ (A)⁺ RNA を精製する。各ポリ (A)⁺ RNA サンプルは、MMLV 逆転写酵素、0.05 pg/ μl のオリゴ (dT) プライマー (21 mer)、 $1 \times$ 第 1 鎖緩衝液、0.03 単位/ μl の RNアーゼインヒビター、500 μM dATP、500 μM dGTP、500 μM dTTP、40 μM dCTP、40 μM dCTP-Cy3 (BDS) または dCTP-Cy5 (Amersham Pharmacia Biotech) を用いて逆転写する。この逆転写反応は、GEMBRIGHT キット (Incyte) を用いて、200 ng のポリ (A)⁺ RNA を含む 25 ml 容

50

量で行う。特異的なコントロールポリ(A)⁺RNAは、in vitro転写により非コーディング酵母ゲノムDNAから合成する。370℃で2時間インキュベートした後、各反応サンプル（一方はCy3標識、他方はCy5標識）は、2.5mlの0.5M水酸化ナトリウムで処理し、850℃で20分間インキュベートし、反応を停止させてRNAを変性する。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピンカラム（CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA)を用いて精製する。結合後、2つの反応サンプルを、1mlのグリコーゲン(1mg/ml)、60mlの酢酸ナトリウム及び300mlの100%エタノールを用いてエタノール沈殿させる。サンプルは次に、Speed VAC (Savant Instruments Inc., Holbrook NY)を用いて乾燥して仕上げ、14μl 5×SSC/0.2% SDS中で再懸濁する。 10

【0284】

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

【0285】

精製したアレイエレメントを、ポリマーコートされたスライドガラス上に固定する。顕微鏡スライドガラス(Corning)は、処理中及び処理後に大量の蒸留水での洗浄と、0.1%のSDS及びアセトン中で超音波による洗浄を行う。スライドガラスは、4%フッ化水素酸(VWR Scientific Products Corporation (VWR), West Chester PA)中でエッチングし、蒸留水中で広範囲にわたって洗浄し、95%エタノール中の0.05%アミノプロピルシラン(Sigma)でコーティングする。コーティングしたスライドガラスは、110℃の天火で硬化させる。 20

【0286】

米国特許第5,807,522号に記載されている方法を用いて、コーティングしたガラス基板にアレイエレメントを付加する。この特許に引用することを以って本明細書の一部とする。平均濃度が100ng/μlのアレイエレメントDNA1μlを高速機械装置により開放型キャピラリープリンティングエレメント(open capillary printing element)に充填する。次にこの装置が、スライド毎に約5nlのアレイエレメントサンプルを分注する。 30

【0287】

マイクロアレイには、STRATALINKER UVクロスリンカー(Stratagene)を用いてUV架橋する。マイクロアレイは、室温において0.2%SDSで1回洗浄し、蒸留水で3回洗浄する。非特異的な結合部位は、リン酸緩衝生理食塩水(PBS)(Tropix, Inc., Bedford MA)における0.2%カゼイン中で60℃で30分間マイクロアレイをインキュベートし、その後上述したように0.2%SDS及び蒸留水で洗浄することによってブロックする。 40

【0288】

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0.2μg含む9μlのサンプル混合体を含めたものである。サンプル混合液を、65℃で5分間加熱し、マイクロアレイ表面上に一定量分注してから1.8cm²のカバーガラスで覆う。このアレイを、顕微鏡スライドより僅かに大きいキャビティを有する防水チェンバーに移す。チェンバーの角に140μlの5×SSCを加えて、チェンバー内を湿度100%に保持する。このア 50

レイを含むチャンバーを、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、第2洗浄緩衝液中(0.1×SSC)において45℃で10分間それぞれ3回洗浄し、その後乾燥させる。

【0289】

検出

レポーター標識されたハイブリダイゼーション複合体は、Cy3を励起するための488nm、及びCy3を励起するための632nmのスペクトル線を生成し得るInnova 70混合ガス10Wレーザー(Coherent, Inc., Santa Clara CA)を備えた顕微鏡で検出する。20倍の顕微鏡対物レンズ(Nikon, Inc., Melville NY)を用いて、アレイ上に励起レーザー光を集中させる。このアレイを含むスライドを顕微鏡のコンピュータ制御X-Yステージに置き、対物レンズを通してラスタスキャンする。本実施例で用いた1.8cm×1.8cmのアレイは、20μmの解像度でスキャンする。

【0290】

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

【0291】

スキャンの感度は通常、既知濃度のサンプル混合体に添加されるcDNAコントロール種により生成されるシグナル強度を用いて校正する。アレイ上の特定の位置には相補的DNA配列を含め、その位置におけるシグナルの強度がハイブリダイズする種の重量比1:100, 000に相関するようにする。異なる試料(例えば検査細胞及びコントロール細胞を代表する)からの2つのサンプルを、各々異なる蛍光体で標識し、他と異なって発現する遺伝子を同定するために単一のアレイにハイブリダイズさせる場合には、校正は2つの蛍光体を有する校正するcDNAのサンプルを標識して、ハイブリダイゼーション混合液に各々等量を加えて行う。

【0292】

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

【0293】

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

【0294】

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～約30個の塩基

対を含むオリゴヌクレオチドの使用について記すが、より小さな或いはより大きな配列の断片の場合でも本質的に同じ方法を用いることができる。O l i g o 4 . 0 6 ソフトウェア (N a t i o n a l B i o s c i e n c e s) 及び T R I C H のコーディング配列を用いて、適切なオリゴヌクレオチドを設計する。転写を阻害するためには、最も独特な 5 ' 配列から相補的なオリゴヌクレオチドを設計し、これを用いてプロモーターがコーディング配列に結合するのを阻害する。翻訳を阻害するためには、相補的なオリゴヌクレオチドを設計して、リボソームが T R I C H をコードする転写物に結合するのを阻害する。

【0295】

12 T R I C H の発現

T R I C H 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 T R I C H が発現するために、抗生物質耐性及び c D N A の転写レベルを高める誘導性のプロモーターを含む好適なベクターに c D N A をサブクローニングする。このようなプロモーターには、l a c オペレーター調節エレメントに関連する T 5 または T 7 バクテリオファージプロモーター及び t r p - l a c (t a c) ハイブリッドプロモーターが含まれるが、これらに限定されるものではない。組換えベクターを、B L 2 1 (D E 3) などの好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピル β - D チオガラクトピラノシド (I P T G) で誘発されると T R I C H を発現する。真核細胞での T R I C H の発現は、昆虫細胞株または哺乳動物細胞株に一般にバキュロウイルスとして知られている A u t o g r a p h i c a c a l i f o r n i c a 核多面性ウイルス (A c M N P V) を感染させて行う。バキュロウイルスの非必須ポリヘドリン遺伝子を、相同組換え或いは転移プラスミドの媒介を伴う細菌の媒介による遺伝子転移のどちらかによって、T R I C H をコードする c D N A と置換する。ウイルスの感染力は維持され、強いポリヘドリンプロモーターによって高いレベルの c D N A の転写が行われる。組換えバキュロウイルスは、多くの場合は S p o d o p t e r a f r u g i p e r d a (S f 9) 昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる。(例えば、E n g e l h a r d , E . K . 他 (1 9 9 4) P r o c . N a t l . A c a d . S c i . U S A 9 1 : 3 2 2 4 - 3 2 2 7 ; S a n d i g , V . 他 (1 9 9 6) H u m . G e n e T h e r . 7 : 1 9 3 7 - 1 9 4 5 . を参照)。

【0296】

殆どの発現系では、T R I C H が、例えばグルタチオン S トランスフェラーゼ (G S T) 、または F L A G や 6 - H i s などのペプチドエピトープ標識で合成された融合タンパク質となるため、未精製の細胞溶解物からの組換え融合タンパク質の親和性ベースの精製が素早く 1 回で行うことができる。 S c h i s t o s o m a j a p o n i c u m からの 2 6 キロダルトンの酵素 G S T によって、タンパク質の活性及び抗原性を維持した状態で固定されたグルタチオンで融合タンパク質の精製が可能となる (A m e r s h a m P h a r m a c i a B i o t e c h) 。精製の後、G S T 部分を特定の操作部位で T R I C H からタンパク質分解的に切断できる。アミノ酸 8 個のペプチドである F L A G で、市販のモノクローナル及びポリクローナル抗 F L A G 抗体 (E a s t m a n K o d a k) を用いた免疫親和性の精製が可能となる。6 個の連続するヒスチジン残基のストレッチである 6 - H i s によって、金属キレート樹脂 (Q I A G E N) で精製が可能となる。タンパク質の発現及び精製の方法は、A u s u b e l (1 9 9 5 , 前出 , c h 1 0 , 1 6) に記載されている。これらの方法で精製した T R I C H を直接用いて該当する以下の 実施例 1 6、1 7、及び 1 8 に示されているアッセイを行うことができる。

【0297】

13 機能のアッセイ

T R I C H 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 T R I C H をコードする配列の発現によって評価する。c D N A を、c D N A を高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにサブクローニングする。このようなベクターには、p C M V S P O R T ^{T M} (L i f e T e c h n o l o g i e s .

10

20

30

40

50

）及び pCR 3.1 (Invitrogen, Carlsbad, CA) が含まれ、どちらもサイトメガロウイルスプロモーターを含んでいる。5～10 µg の組換えベクターを、例えば内皮由来か造血由来のヒト細胞株にリポソーム製剤或いは電気穿孔法によって一時的に形質移入する。更に、標識タンパク質をコードする配列を含む 1～2 µg のプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入された細胞と形質移入されていない細胞とを区別できる。また、標識タンパク質の発現によって、cDNA の組換えベクターからの発現を正確に予想できる。このような標識タンパク質には、緑色蛍光タンパク質 (GFP; Clontech)、及び CD64 または CD64-GFP 融合タンパク質が含まれる。レーザー光学に基づいた技術を利用した自動流動細胞計測法 (FCM) を用いて、GFP または CD64-GFP を発現する形質移入された細胞を同定し、その細胞のアポトーシス状態や他の細胞特性を評価する。また、FCM で、先行した或いは同時の細胞死の現象を診断する蛍光分子の取り込みを検出して計量する。これらの現象には、プロピジウムヨウ化物での DNA の染色によって計測される核 DNA 内容物の変化と、プロモデオキシウリジンの取り込み量の低下によって計測される DNA 合成の下方調節と、特異的な抗体との反応性によって計測される細胞表面及び細胞内のタンパク質の発現の変化と、蛍光複合アネキシン V タンパク質の細胞表面への結合によって計測される原形質膜組成の変化とが含まれる。流動細胞計測法は、Ormerod, M. G. による (1994) Flow Cytometry Oxford, New York, NY. に記載されている。

10

【0298】

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) の保存された領域と結合する。形質転換された細胞と形質転換されない細胞とは、ヒト IgG か CD64 に対する抗体のどちらかで被覆された磁気ビードを用いて分離することができる (DYNAL, Lake Success, NY)。mRNA は、当分野で周知の方法で細胞から精製することができる。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。

20

【0299】

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。

30

【0300】

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)。

40

【0301】

通常、約 15 残基の長さのオリゴペプチドを、Applied Biosystems の ABI 431A ペプチドシンセサイザー (PE Biosystems) を用いて fmoc 法のケミストリにより合成し、N-マレイミドベンゾイル-N-ヒドロキシスクシンイミドエステル (MBS) を用いた反応により KLH (Sigma-Aldrich, St. Louis MO) に結合させて、免疫原性を高める (例えば、前出の Ausubel, 1995 を参照)。フロイントの完全アジュバントにおいてオリゴペプチド-KLH 複合体を用いてウサギを免疫化する。得られた抗血清の抗ペプチド活性及び抗 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

50

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。

【0302】

1.5 特異的抗体を用いる天然 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。結合の後、そのレジンを製造者の使用説明書に従ってブロッキング処理し洗浄する。

10

【0303】

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、pH 2～3 のバッファー、或いは高濃度の尿素またはチオシアン酸塩イオンのようなカオトロピックイオンで）溶出させ、TRICH を回収する。

【0304】

1.6 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関与するタンパク質が含まれる。TRICH または生物学的に活性なその断片を、¹²⁵I ボルトンハンター試薬（例えば、Bolton A. E. 及び W. M. Hunter (1973) Biochem. J. 133: 529 を参照）で標識する。マルチウェルプレートに予め配列しておいた候補の分子を、標識した 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。

20

【0305】

別法では、TRICH と相互作用するタンパク質が 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。

30

40

【0306】

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)。

【0307】

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。

50

【0308】

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13-ガラクトシダーゼを発現し蓄積するのに適した条件下で、48-72時間10に渡ってインキュベートされる。

【0309】

β-ガラクトシダーゼを発現させる形質転換細胞は、適切な比色基質が本技術分野で既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、膜電気伝導度差を試験される。形質転換されていない細胞、及び／又はベクター配列単独、若しくはβ-ガラクトシダーゼ単独のいずれかで形質転換された細胞が対照として用いられ、並行に試験される。TRICHを発現する細胞は、対象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しうる。コンダクタンスに対するTRICHの寄与は、TRICHの抗体特性を用いて細胞をインキュベートすることで確認されてもよい。抗体は、TRICHの細胞外側面に結合されてもよく、それによってイオンチャネル内の孔及び関連のコンダクタンスをブロッキングする。20

【0310】

別法では、TRICHのイオンチャネル活性は、二電極電位固定技術(Ishiiら(既出)、Jegla, T. 及びL. Salter (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組織にかか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ニングされ、また0, 5-5ng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、116mMのK-グルコン酸、4mMのKCl、および10mMのHepes (pH 7.2)を含む細胞内液へと抽出される。細胞内液は、適切な、cAMP、cGMP、またはCa²⁺ (CaCl₂の形態)のような様々な濃度のTRICHメディエータで補われる。電極の抵抗は2-5MΩでセットされ、細胞内液欠乏メディエータで満たされる。実験は、外界温度で0mVの保持電位で実行される。-100mVから100mVの電圧ランプが、サンプリング周波数500Hzで得られる。電流測定はアッセイ中のTRICHの活性に比例する。30

【0311】

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10ng)で注入され、OR2培地(82, 5mMのNaCl、2, 5mMのKCl、1mMのCaCl₂、1mMのMgCl₂、1mM Na₂HPO₄、5mMのHepes、3, 8mMのNaOH、50μg/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(100mMのNaCl、2mMのKCl、1mMのCaCl₂、1mMのMgCl₂、10mM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と共に放射標識されていたり、ローダミンと共に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、組込ラベルを測定し対照と比較するNa⁺遊離(free)培地内で、50

卵母細胞を3回洗浄することで終了する。T R I C H活性は内部移行されたラベル基質のレベルに比例する。

【0312】

T R I C Hに関連するA T P a s e活性は、放射性ラベルされたA T P - [γ - 3 2 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 3 2 Pの定量化によって測定される。反応混合物は、A T P - [γ - 3 2 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T R I C H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、膜若しくはフィルターの紙ベースクロマトグラフィーにかけられる。開放された 3 2 Pの量は、シンチレーション計数器で測定される。回復した放射活性の量はアッセイ中のT R I C HのA T P a s e活性に比例する。

【0313】

18 T R I C Hアゴニスト及びアンタゴニストの同定

T R I C Hは、C H O (チャイニーズハムスターの卵巣) またはH E K (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、細胞膜にかかるイオンの流れを測定する蛍光技術を用い検定される (V e l i c e l e b i, G. ら (1999) M e t h. E n z y m o l. 294 : 20-47 ; W e s t, M. R. 及びC. R. M o l l o y (1996) A n a l. B i o c h e m. 241 : 51-58)。これらアッセイは、マイクロプレートを用いる高スループット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、C a $^{2+}$ 指示薬F l u o - 4 A M、S B F I 及びナトリウムグリーンのようなナトリウム感知染料、C l 指示薬M Q A E (すべてM o l e c u l a r P r o b e s より入手可能) のような蛍光染料を用い、F L I P R 蛍光定量プレートリーディングシステム (分子デバイス) と併せて測定される。アッセイのさらに一般的な変法では、原形質膜にかかるイオンの流れによって生じる膜電位の変化が、D i B A C $_4$ (分子プローブ) のようなo x o n y l 染料を用いて測定される。D i B A C $_4$ は、細胞の外側の溶液と細胞膜電位による細胞部位との間を釣り合わせる。染料の蛍光強度は、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D i B A C $_4$ の取り込みが検出可能となる (G o n z a l e z, J. E. 及び P. A. N e g u l e s c u (1998) C u r r. O p i n. B i o t e c h n o l. 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、若しくはコンビケムライブラリより選択されてもよい。

【0314】

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

【0315】

(表の簡単な説明)

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

【0316】

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

【0317】

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

【0318】

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

【0319】

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。

【0320】

表 6 は、表 5 に示す cDNA ライブラリの作製に用いた組織およびベクターを示す付録である。 10

【0321】

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

【表 1】

表1

インサイト プロジェクトID	ポリペプチド SEQ ID NO:	インサイト ポリペプチドID	ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチドID
1784775	1	1784775CD1	16	1784775CB1
7473034	2	7473034CD1	17	7473034CB1
1878581	3	1878581CD1	18	1878581CB1
2246292	4	2246292CD1	19	2246292CB1
5151730	5	5151730CD1	20	5151730CB1
7472584	6	7472584CD1	21	7472584CB1
7472536	7	7472536CD1	22	7472536CB1
7473422	8	7473422CD1	23	7473422CB1
2864715	9	2864715CD1	24	2864715CB1
1734724	10	1734724CD1	25	1734724CB1
1563237	11	1563237CD1	26	1563237CB1
7473443	12	7473443CD1	27	7473443CB1
7473438	13	7473438CD1	28	7473438CB1
7474286	14	7474286CD1	29	7474286CB1
7472589	15	7472589CD1	30	7472589CB1

10

20

30

40

表2-1

ポリペプチド SEQ ID NO:	インサイト ポリペプチドID	GenBank ID NO:	確率スコア	GenBank 相合体
1	1784775CD1	g4902730	4.00E-94	[ヒト] 多剤耐性関連タンパク質7
2	7473034CD1	g2317274	9.00E-168	[ヒト] アクアポリン脂質 Kuriyama, H.ら (1997) Biochem. Biophys. Res. Commun. 241, 53-58
3	1878581CD1	g545998	1.10E-164	[Rattus sp.] tricarboxylate 担体ラット、 肝臓、ミトコンドリアのペプチドの一部、 357個のアミノ酸 Azzi, A.ら (1993) J. Bioenerg. Biomembr. 25, 515-524
4	2246292CD1	g6045150	0	[ドブネズミ] TAP様ABC輸送体 Yamaguchi, Y.ら (1999) FEBS Lett. 457, 231-236
5	5151730CD1	g475542	0	[ドブネズミ] グルタミン酸受容体 δ -1サブユニット
		g220418	0	[マウス] グルタミン酸受容体 チャネルサブユニット δ -1 Yamazaki, M.ら (1992) Biochem. Biophys. Res. Commun. 183:886-892
6	7472584CD1	g7546839	0	[モルモット] カリウムチャネルTASK3 Rajan, S.ら (2000) J. Biol. Chem. 275:16650-16657
7	7472536CD1	g2687858	4.30E-129	[Pseudopleuronectes americanus] 腎有機アニオン輸送体 Wolff, N.A.ら (1997) Expression cloning and characterization of a renal organic anion transporter from winter flounder, FEBS Lett 417:287-291
8	7473422CD1	g9950945	0	[総論] 鉄(III)-輸送体系パーミアーゼ HitB Stover, C.K.ら (2000) Nature 406:959-964

10

20

30

【表3】

表2-2

ポリペプチド SEQ ID NO:	インサイト ポリペプチドID	GenBank ID NO:	確率スコア	GenBank 相合体
9	2864715CD1	g531459	0	[ドブネズミ] 腎臓ストレス誘発性 Na-Gl 有機溶質共輸送体
		g3347922	5.00E-190	[マウス] オーファン輸送体アイソフォーム A12 Nash, S.R. & (1998) Cloning, gene structure and genomic localization of an orphan transporter from mouse kidney with six alternatively-spliced isoforms, Receptors Channels 6:113- 128
10	1734724CD1	g3980315	5.00E-46	[Oryctolagus cuniculus] 肝臓ナトリウム依存性胆汁酸
11	1563237CD1	g2208839	2.00E-125	[ドブネズミ] ペプチド/ヒスチジン輸送体 Yamashita, T. & (1997) Cloning and functional expression of a brain peptide/histidine transporter, J. Biol. Chem. 272:10205-10211
		g7415511	0	[ヒト] ペプチド輸送体 3
12	7473443CD1	g4186073	3.20E-280	[マウス] カルシウムチャネル α -2- δ -0サブユニット
13	7473438CD1	g2465542	7.60E-84	[ヒト] TWIK 関連酸感受性 K ⁺ チャネル Duprat, F. & (1997) TASK, a human background K ⁺ channel to sense external pH variations near physiological pH, EMBO J. 16:5464- 5471
		g11228686	0	[ヒト] 2 孔カルシウムチャネル KT3.3

【表 4】

表2-3

ポリペプチド SEQ ID NO:	インサイト ポリペプチドID	GenBank ID NO:	確率スコア	GenBank 相合体
14	7474286CD1	g306473	1.90E-39	[ヒト] カルシウムチャネル α サブユニット Powers, P.A.ら (1993) Molecular characterization of the gene encoding the gamma subunit of the human skeletal muscle 1,4- dihydropyridine-sensitive Ca^{2+} channel (CACNLG), cDNA sequence, gene structure, and chromosomal location, J. Biol. Chem. 268:9275- 9279
15	7472589CD1	g11177514	0	[ヒト] タンデムリピドメインカリウムチャネル THIK-2 Rajan, S.ら (2001) J. Biol. Chem. 276:7302-7311

【表 5】

10

20

30

40

表3-1

SEQ ID NO.	インサイトポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ(signature)配列, ドメイン及びモチーフ	分析方法及びデータベース
1	1784775CD1	520	T343 S354 S404 T496 S345		膜貫通ドメイン: L25-L42; P120-Y142; L277-L297 L365-P392; Y465-M485 ABC 輸送体膜貫通領域: Y242-E499 RESISTANCE; MULTIDRUG; SAUROLEISHMANIA; C3F10.11C; DM01742 P33527 195-653: P160-P201; Q221-V483 膜貫通ドメイン: M46-L64	HMW HMW-PFAM BLAST-DOMO
2	7473034CD1	346	Y297 T11 S146 T192 T194 S252 S10	N65 N327	主内在性タンパク質: E31-Y276 MIP family proteins BL00221A:A43-V53 BL00221B:I92-T102 BL00221C:E179-D195 BL00221D:T225-I239 BL00221E:W258-G268 MAJOR INTRINSIC PROTEIN PR00783A:R39-S58 PR00783B:F78-T102 PR00783C:K115-I134 PR00783D:N178-Q196 PR00783E:G79-V101 PR00783F:W259-F279 AQUAPORIN ADIPOSE AQPAP TRANSPORT TRANSMEMBRANE PD078705:L277-F346 MIP FAMILY DM00228 P47862 15-263:K36-Y276 DM00228 I59266 15-263:K36-Y276 DM00228 P43549 340-587:R35-F279 DM00228 F11244 1-253:L42-Y276	HMW HMW-PFAM BLIMPS_BLOCKS BLIMPS_PRINTS BLAST_PRODOM BLAST-DOMO

【表 6】

表3-2

SEQ ID NO.	インサイト ポリペプチドID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列, ドメイン及びモチーフ	分析方法及び データベース
3	1878581CD1	322	T30 T64 T143 T296 S309 S46 T134	N100 N124 N132 N135	PROTEIN TRANSMEMBRANE CHROMOSOME PUTATIVE TRANSPORTER PD006986:L5-L251	BLAST_PRODOM
4	2246292CD1	723	T139 S161 T209 T311 T377 T469 S500 S720 S354 Y559 S28 S33 S46 T153 T181 S275 T367 S528 S552 S628 S659	N280 N465 N481 N556 N718	ABC輸送体: L600-L614 ATP/GTP結合部位モチーフA: G496-S503 膜貫通ドメイン V85-F104, V185-F204, L328- G347 ABC輸送体: G489-G673 ABC輸送体膜貫通領域: L188-G450 ABC輸送体ファミリー BL00211A:L494-C505 BL00211B:L600-D631 ABC輸送体ファミリーシグネチャ: I582-D631 ATP-BINDING TRANSPORT PROTEIN TRANSMEMBRANE GLYCOPROTEIN TRANSPORTER: PD000130:V229- A406 MT2 PROTEIN DM07894 Q03519 1-685:L395-V675 DM07894 S25577 1-703:V422-M699 ABC TRANSPORTERS FAMILY DM000008 Q03518 502-717:V461- G673 DM000008 S13426 479-695:V461- G673	MOTIFS HMMER HMMER_PFMAM BLIMPS_BLOCKS PROFILESSCAN BLAST_PRODOM BLAST_DOMO

【表7】

表3-3

SEQ ID NO:	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列, ドメイン及びモチーフ	分析方法及び データベース
5	5151730CD1	1009	S17, S49, S57, S63, S104, S123, S171, S180, T211, T223, S263, S287, S355, T356, S364, T368, T400, S417, S424, T443, T438, S510, S528, S533, S659, T667, T681, T704, S724, S750, T752, S766, S785, S825, S857, T864, S882, S901, T939, Y220, Y679	N200, N422, N498	代謝調節型グルタミン酸受容体: PR00593C: S518-S533 N-メチル-D-アスパラギン酸 (NMDA): PR00177D: I632-S659 HYDROXYTRYPTAMINE 2A RECEPTOR (5-HT-2A) / (SEROTONIN RECEPTOR) (5-HT-2): PR00177D: I632-S659 GLUTAMATE RECEPTOR SUBUNIT DELTA2 DELTA1 CHANNEL SUBTYPE DELTA: PD009864: M83-L437, PD009890: W854-I1009, PD000500: V565-V848, PD000273: L437-E550 GLUTAMATE RECEPTOR: DM08722 S28857 1-421:M1-N422, DM00247 S28857 627-897:S627-I898, DM08722 S28858 2-426:A20-N422, DM00247 S28858 628-900:S627-I898 シグナルペプチド: M1-A20 膜貫通ドメイン (transmem domain): L557-L583, S830-I847 受容体ファミリーリガンド結合領域 (ANF_receptor): I12-S424	BLIMPS-PRINTS

【表 8】

表3-4

SEQ ID NO.	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列, ドメイン及びモチーフ	分析方法及び データベース
6	7472584CD1	374	S31, S55, S127, S179, S251, S319, S331, T341, S360, S373,	N53	TASK K+チャネルシグネチャ: PR01095A: V6-V25, PR01095B: F26-K51, PR01095C: E63-Q77, PR01095D: Q126-K145, PR01095E: C146-T161, PR01095F: Q209- V230, PR01095G: G236-E254, PR01095H: R283-D306, PR01095K: P348-V374, CHANNEL IONIC TWIKRELATED ACIDSENSITIVE K+ CTBAK F34D6.3 PROTEIN PUTATIVE POTASSIUM: PD013020: M1-A74 シグナルペプチド (signal_peptide): M1-V25 膜貫通ドメイン (transmem_domain): F225-V243 TASK K+チャネル (TWIK_channel): M1-V374 イオン輸送タンパク質 (ion_trans): I54-L244 糖輸送タンパク質: DM00135 P38142 145-478: R230- E488 (p = 6.1e-06) 膜貫通ドメイン: G400-M425; T424-A447 糖(及び他の)輸送体: S133-A553	ELIMPS-PRINTS
7	7472536CD1	589	S94, T121, T170, S192, S246, S308, S362, T376, T433, T513, S563, S564	N54, N74, N89, N96, N131, N149	Transporter: organic cation transport protein: putative ion renal cationic transmembrane protein: PD003661: M34-W125	BLAST-DOMO HMMER- TRANSMEMBRANE HMMER-PTAM BLAST-PRODOM

【表 9】

10

20

30

表3-5

SEQ ID NO.	インサート ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列, ドメイン及びモチーフ	分析方法及び データベース
8	7473422CD1	549	S3, S19, S49, S155, T182, T188, S189, S193 T449, T541		結合タンパク質依存性輸送体シグネチャ: BPD_transp: T182-E252	HMMER-PFAM
					結合タンパク質依存性 輸送体の内制成分: bpd_transp_inn_membr.prf: L432-E231	PROFILES CAN
					タンパク質: 鉄 III 輸送体系, パーミアーゼ輸送体、膜貫通内側膜 HITB PD031934: K46-L183	BLAST-PRODOM
					DOMAIN: SFUB; IRON; TRANSPORT; PERIPLASMIC; DM05963 D64048 210-473: A316- L540	BLAST-DOMO
					BPD 輸送体内在性 膜タンパク質シグネチャ: L183-P211, L443-P471	MOTIFS
					膜貫通ドメイン: P48-V70; N97-A117; M262-L279; V315- S334; V356-V375	HMMER- TRANSMEMBRANE

10

20

30

【表 10】

表3-6

SEQ ID NO:	インサイト ポリペプチドID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列, ドメイン及びモチーフ	分析方法及び データベース
9	2864715CD1	634	S17, S30, S100, S143, S170, S186, T212, T219, T243, T267, T330, S483, T533, S561, S573	N158 N182 N258 N354 N368	SODIUM/NEUROTRANSMITTER SYMPORTER FAMILY: DM00572 S50998 19-616:S31-G614 ナトリウム：神経伝達物質： BL00610A: Q40-E89; BL00610B: W104-W153; BL00610C: W195- G246; BL00610D: E261-T313; BL00610E: A406-V448; BL00610F: M502-Q556; BL00610G: I576-P598 ナトリウム：神経伝達物質： 共輸送体ファミリーシグネチャ： neurotransp_1: D36-F90 SODIUM/NEUROTRANSMITTER: PR00176A: Q40-L61; PR00176B: A69-L88; PR00176C: G113-Y139; PR00176D: A222-I239; PR00176E: V304-V324; PR00176F: M410- L429; PR00176G: S491-V511; PR00176H: Q531-V551 TRANSPORTER NEUROTRANSMITTER TRANSPORT TRANSMEMBRANE SYMPORT GLYCOPROTEIN SODIUM CHLORIDEDEPENDENT SODIUMDEPENDENT GABA: PD000448:C383-I606 膜貫通ドメイン: F70-L88, F123-F141, I305-V324, S417- V440, I462-Y480, T533-F550 ナトリウム神経伝達物質 共輸送体ファミリー (SNF): R32-R332, L378-N608	BLAST-DOMO BLIMPS-BLOCKS PROFILES CAN BLIMPS-PRINTS BLAST-PRODOM HMMER- TRANSMEMBRANE HMMER-PFAM

10

20

30

【表 1 1】

表3-7

SEQ ID NO.	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列, ドメイン及びモチーフ	分析方法及び データベース
10	1734724CD1	491	T227, S233, S459, T475	N6, N18, N24, N180	SODIUM; ACID; BILE; TRANSPORTER DOMAIN; DM03972 I38655 8-318:S251-Y439 膜貫通ドメイン: V241- L261, L288-F307, W325-G343, M416-I435 ナトリウム胆汁酸共輸送体ファミリー (SBF): L113-E345 (score=6.0, e-value=0.00017) ACID COTRANSPORTING POLYPEPTIDE; SODIUM/BILE COTRANSPORTER; NA+/BILE SODIUM/TAUROCHOLATE TRANSMEMBRANE TRANSPORT SYMPORTER: PD007533:L342-Y439 PTR2ファミリー (putative protein transp.), proton/oligo: BL01022A:E42-L60; BL01022B:A72-L117; BL01022C:C164-V187; BL01022D:F199-V211; BL01022E:E416-S451 グルコース輸送体シグネチャ: PR00172C: V241-L261.	BLAST-DBO HMMER HMMER-PFAM BLAST-PRODOM BLIMPS-BLOCKS BLIMPS-PRINTS

【表 1 2】

表3-8

SEQ ID NO.	インサート ポリペプチドID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列, ドメイン及びモチーフ	分析方法及び データベース
11	1563237CD1	525	S281, S302, S367, S174, T195, S281, S371	N61, N66, N178, N223, N356, N383	PEPTIDE OLIGOPEPTIDE TRANSPORTER; PROTEIN SYMPORT ISOFORM H+/PEPTIDE COTRANSPORTER: PD001550: Y101- S262 PTR2 FAMILY PROTON/OLIGOPEPTIDE SYMPORTERS: M01990 P46032 46- 551:E42-L269 POTファミリー(推定上のタンパク質輸送体) PTR2:Y101-S259; PTR2:N396-S447	BLAST-PRODOM
12	7473443CD1	1310	S52, T81, T96, S108, S109, S119, S176, S213, S330, T375, S410, S426, S459, S511, S550, S601, T725, S748, S761, T776, T780, S897, T918, S976, S1026, S1052, S1062, T1125, S1126, T1150, S1160, T1170, T1235, T1239, T1245, T1287	N16, N457, N661, N979, I124, I206	カルシウムチャネルとの類似性: DM06895 P54289 261- 693:S759-L1187 CHANNEL CALCIUM PRECURSOR; IONIC TRANSMEMBRANE ION TRANSPORT; VOLTAGE-GATED DIHYDROPYRIDINESENSITIVE LTYPE: PD013837:I631-K750	BLAST-DOMO BLAST-PRODOM

【表 1 3】

10

20

30

表3-9

SEQ ID NO.	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列, ドメイン及びモチーフ	分析方法及び データベース
13	7473438CD1	400	S75 S101 S125 T162 S197 S249 S7 S105 S175 S332		膜貫通ドメイン: S75-F96; F295-V313 TASK K+ チャネルドメイン: M71-I400 TASK K+ チャネルシグネチャ PR01095: V76-V95; F96-K121; E133-Q147; Q196-C215; C216- V231; Q279-L300; G306-P324 CHANNEL PROTEIN IONIC POTASSIUM SUBUNIT K+ PUTATIVE SUBFAMILY K MEMBER PD021430: L229-V312	HMER HMER-PFAM BLIMPS-PRINTS BLAST-PRODOM
14	7474286CD1	260	T121 T133 T67 T88 T132 S195	N71 N113	シグナルペプチド: M1-A53 膜貫通ドメイン: V146-S164; F169-V187; V225- T242 CALCIUM CHANNEL GAMMA SUBUNIT DIHYDROPYRIDINESENSITIVE LTYPE SKELETAL MUSCLE IONIC TRANSMEMBRANE PD016829: L31-C223	SPSCAN HMER BLAST-PRODOM
15	7472589CD1	489	S257 S373 S424 S58 S62 S260 T295 S416 S429 S2 S5 S13 T36 T75 T149 T156 T252 S436	N78	膜貫通ドメイン: F204-F223; P270-M293; F340- I359 TASK K+ チャネルドメイン: L32-E484 (Score = -115.2; E-value = 0.0078)	HMER HMER-PFAM

【表 1 4】

表4-1

ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチドID	配列長さ	選択された断片	配列断片	5'位置	3'位置
16	1784775CB1	1735	233-901, 1-59, 1252-1513	60118407D2 60209102U1 6756053J1 (SINTFER02) 6431853H1 (LUNGNON07) GBI:96850429_000023. fasta.edit.comp GBI:96742097_000003. fasta.edit	948 1053 1057 324 1 1	1303 1416 1735 1021 771
17	7473034CB1	1041	158-280	71456748V1 71461790V1 71465463V1 71464816V1 71465977V1 71456476V1 71466268V1	228 1602 300 1499 807 1723 908	852 2235 863 2193 1533 2367 1564
18	1878581CB1	2367	1-42, 1275-2367	4202150H1 (BRAITUT29) 5531403F6 (HEARFET05) 60111369D2 95396013 3811252F6 (CONTTUT01) 2246292R6 (HIPONON02) 6890917J1 (BRAITDR03) 7168417H1 (MCLRNOC01) 4691312T6 (BRAENOT02) 6120730H1 (BRAHNON05) 3334167H1 (BRAIFET01) 4691312P6 (BRAENOT02) 5014991H1 (BRAXNOT03) 2395385F6 (THPIAZT01)	1 1747 1 2884 1387 2978 650 100 2689 2365 549 1959 1713 1026	280 2162 473 3343 1948 3333 1293 648 3320 3044 818 2552 1968 1490
19	2246292CB1	3343	1-867, 2606-2809, 1501-1983, 960-1013			

【表 15】

10

20

30

表4-2

ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチドID	配列長さ	選択された断片	配列断片	5'位置	3'位置
20	5151730CB1	3517	863-1689, 1-113, 3288-3339	6885903J1 (BRAHTR03) 5151730F6 (HEARFET03) 1675978F6 (BLADNOT05) 3907290H1 (LUNGN0T23) 4406394H1 (PROSDIT01) 6974079F8 (BRAHTR04) 7288330H1 (BRAIFER06) 6250428F8 (LUNPTUT02) 7326591H2 (SPLANTUE01) 7174088H1 (BRSTTWC01) 3486391T6 (EPIGNOT01) 6891842H1 (BRAHTR03) 449149H1 (TLYMNOT02) GBF.g6164987.raw g2001843 7693385H2 (LNODTUE01) 7693385J2 (LNODTUE01) GNN.g5776606_000002_002 .edit GNN.g5525050_000002_004 70870289V1 70529541V1 6783918H1 (SINITWC01) 6829267J1 (SINTNOR01) 70792490V1 6783311H1 (SINITWC01) 6783435H2 (SINITWC01) 1734724F6 (COLNNOT22) 60123116D1 4222214H1 (PANCNOT07) 3495639T7 (ADRETUT07) 3495639F6 (ADRETUT07) GNN.g7344269_000021_002 6147668H1 (BRANDIT03)	2374 604 2973 3381 2229 1 1855 183 804 1082 2821 1458 959 1 946 1 253 1 1 1771 550 1254 1088 2213 1710 1 1063 902 1942 1525 963 69 1	2952 927 3502 3517 2485 582 2316 856 1409 1683 3440 1993 1128 1248 743 879 1770 2544 2335 1154 1738 1616 2871 2261 466 1704 987 2141 2120 1551 1544 545
21	7472584CB1	1248	751-945, 96-219			
22	7472536CB1	1770	1-481, 1136-1770, 573-1030			
23	7473422CB1	2544	1-2544			
24	2864715CB1	2871	1-32, 750- 1421, 1950- 2871			
25	1734724CB1	2141	658-817			

10

20

30

【表 16】

表4-3

ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチドID	配列長さ	選択された断片	配列断片	5'位置	3'位置
26	1563237CB1	1902	1-1283	7127258H1 (COLNDIY01) 2207658F6 (SINTFET03) 1563237T6 (SPLNNOT04) 2118071T6 (BRSTTUT02) 7679862H1 (BRAFTUE01)	689 818 1332 1410 1	1258 1379 1899 1902 722
27	7473443CB1	4125	1-41, 3763- 4125, 407- 435, 491- 2446	71153376V1 FL722155_00001.raw 71395953V1 GBL.g3810573.raw 71151716V1 71158575V1 GBL.g3810573.edit	3186 1864 3713 1 3338 2741 634	3857 2161 4125 804 3892 3346 3099
28	7473438CB1	2460	1628-1717, 1-374, 998- 1020, 515- 540, 2181- 2460, 1182- 1278	6758949H1 (HEAONOR01) 2230573F6 (PROSNOT16) FL767399_00001 CpG_991027_B15_masked_ fa.Contig52563	1907 1541 722 1	2460 2120 1974 1162
29	7474286CB1	896	86-487	GNN.g7523629_000001_ 002	1	896
30	7472589CB1	2080	1-492, 1002-1177, 618-691	7289171H1 (BRAIFER06) GNN.g6560849_000001_ 002	1 285	599 2080

【表 17】

10

20

30

40

表5

ポリヌクレオチド SEQ ID NO:	インサート プロジェクトID	代表的ライブラリ
16	1784775CB1	BRAINOT10
18	1878581CB1	SPLNFET02
19	2246292CB1	HIPONON02
20	5151730CB1	PROSTMT07
21	7472584CB1	TLYMNOT02
22	7472536CB1	LNODTUE01
24	2864715CB1	SINITUT03
25	1734724CB1	ADRETUT07
26	1563237CB1	DRGLNOT01
27	7473443CB1	THP1AZS08
28	7473438CB1	PROSNOT16
29	7474286CB1	PTHYNOT03
30	7472589CB1	BRAIFER06

【表 18】

10

20

30

表6-1

ライブラリ	ベクター	ライブラリの説明
ADRETUT07	pINCY	このライブラリは、43歳の白人女性から片側副腎摘出の際に採取した副腎腫瘍組織から単離したRNAを用いて作製した。病理学的には、褐色細胞腫を示していた。
BRAIFER06	PCDNA2.1	このランダムプライミングによるライブラリは、妊娠23週目で発育不全で死産した白人男子胎児から採取した脳組織から単離したRNAを用いて作製した。血清検査は陰性であった。
BRAINOT10	pINCY	このライブラリは、アルツハイマー病で死亡した74歳白人男性の脳から採取した病変小脳組織から単離したRNAを用いて作製した。
DRGLNOT01	pINCY	このライブラリは、急性肺水腫及び気管支肺炎、両側肺の胸水及び心外膜液、及び悪性リンパ腫（ナチュラルキラー細胞型）で死亡した年齢32歳の白人男性の頸椎から採取した後根神経節組織から単離されたRNAを用いて作製した。患者の既往症には、サイトメガロウイルス感染の可能性、肝臓のうっ血及び脂肪症、脾腫、出血性膀胱炎、甲状腺出血、及びベル麻痺があった。過去の外科処置には、結腸鏡検査、大腸の生検、アデノイド口蓋扁桃摘出術、及び鼻咽頭の内視鏡検査及び生検があり、放射線治療も受けていた。
HIPONON02	PSPORT1	この標準化された海馬ライブラリは、海馬組織ライブラリからの113万個の独立クローンから作製された。RNAは、頭蓋出血で死亡した72歳白人女性の海馬組織から単離した。既往症には、鼻の癌、高血圧、及び関節炎があった。標準化及びハイブリダイゼーションの条件は、Soares 他の方法を利用した(PNAS (1994)91:9928)。
LNODTUE01	PCDNA2.1	この5'に重点をおいたランダムプライミングによるライブラリは、67歳白人女性の所属リンパ節の切除、舌の直視下生検、及び部分舌切除の際に採取したリンパ節腫瘍組織から単離したRNAを用いて作製した。病理学的には、転移性基底細胞扁平上皮癌であった。関連腫瘍組織は、病理学的には転移性基底細胞扁平上皮癌であって、舌の基底に潰瘍化腫瘍を形成していた。腫瘍が、骨格筋の下側に浸潤していた。患者は、舌癌であった。既往症には、良性高血圧及びアルコールの乱用があった。患者は薬物の投与を受けていなかった。家族歴は、息子の糖尿病があった。

10

20

30

40

表6-3

ライブラリ	ベクター	ライブラリの説明
THP1A2S08	PSP0RT1	このサブトラクトされたTHP-1前単核細胞株ライブラリは、5'-aza-2'-デオキシシチジン(AZ)処理したTHP-1ライブラリからの 5.76×10^6 個のクローンをを用いて作製した。開始RNAは、0.8 μ MのAZで3日間処理したTHP-1前単核細胞から作製した。ドナーは、急性単球白血病であった。サブトラクションのためのハイブリダイゼーションプロトコルは、未処理のTHP-1細胞ライブラリから単離した1 μ gのポリA RNAから作製した、同様に作製したライブラリに由来する。次に、AZ処理したTHP-1細胞ライブラリから得た 5.76×10^6 個のクローンを、未処理のTHP-1細胞ライブラリから得た 5.0×10^6 個のクローンを用いた2回のサブトラクション。サブトラクションにかけた。サブトラクションを用いた2回のサブトラクション条件は、Swaroopら (NAR, (1991) 19:1954) 及びBonaldoら (Genome Research (1996) 6:791) の方法論に基づく。
TLYMN0T02	Lambda UniZAP	このライブラリは、非粘着性末梢血の単核細胞から単離したRNAを用いて作製した。この末梢血は、血縁関係のない男女ドナーに由来し、LPSで0時間処理した。各ドナーからの細胞をフィコール-ハイバーク法により精製し、遠心分離器にかけ、GusONを含むバッファーで溶解し、塩化セシウムで吸収してRNAを得、ライブラリの作製に用いた。ポリA RNAは、Qiagen Oligotex キットを用いて単離した。cDNAの合成は、XhoI-オリゴ(dT)プライマーを用いて開始した。二本鎖cDNAを平滑化し、Lambda UniZAPベクターのXhoI部位及びEcoRI部位にクローニングした。

表7-1

プログラム名	説明	引用文献	パラメーター関連値
ABI FACTURA	核酸配列においてベクター配列を除去して特定の塩基をマスクするプログラム。	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finderは、アミノ酸または核酸配列の比較及び注釈付け (annotation) に有用である。	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler BLAST	核酸配列を構築するプログラム。 Basic Local Alignment Search Tool は、アミノ酸及び核酸配列の配列類似性検索に有用であり、blastp及びblastn、blastx、tblastn、tblastxの5つのファンクションがある。	Applied Biosystems, Foster City, CA. Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: 確率値=1.0E-8 以下 完全長配列: 確率値=1.0E-10 以下
FASTA	Pearson 及び Lipman アルゴリズムは、問い合わせの配列と同種の配列群との類似性を検索する。FASTA は、fasta 及び tfasta、fastx、tfastx、ssearch の少なくとも5つのファンクションを含む。	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; Smith, T.F. and W. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E 値=1.0E-6 構築された ESTs: fasta 同一性=95%以上、一致長さ=200 塩基以上、fastx E 値=1.0E-8 以下 完全長配列: fastx スコア=100 以上
BLIMPS	Blocks IMPROVED Searcher は、BLOCKS 及び PRINTS、DOMO、PRODOM、PFAM データベースにおける配列に対して、ある配列の一致性を調べ、遺伝子ファミリー及び配列相同性、構造的フィンガープリント領域を探索する。	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.C.; Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; Artwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	確率値=1.0E-3 以下
HMMER	PFAM などのタンパク質ファミリーコンセンサス配列の隠れマルコフモデル(HMM)に基づいたデータベースに対して問合せ配列を検索するアルゴリズム。	Krogh, A. 他 (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. 他 (1998) Nucleic Acids Res. 26:320-322. Durbin, R. 他 (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM ヒット: 確率値=1.0E-3 以下 シングルペルバチドヒット: スコア=0 以上

10

20

30

40

表7-2

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

10

20

30

40

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
18 October 2001 (18.10.2001)

PCT

(10) International Publication Number
WO 01/77174 A2

(51) International Patent Classification: C07K 14/705

(21) International Application Number: PCT/US01/11206

(22) International Filing Date: 6 April 2001 (06.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/195,595	6 April 2000 (06.04.2000)	US
60/196,872	12 April 2000 (12.04.2000)	US
60/199,020	20 April 2000 (20.04.2000)	US
60/200,552	28 April 2000 (28.04.2000)	US
60/202,348	5 May 2000 (05.05.2000)	US
60/203,495	11 May 2000 (11.05.2000)	US

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

US	60/195,595 (CIP)
Filed on	6 April 2000 (06.04.2000)
US	60/196,872 (CIP)
Filed on	12 April 2000 (12.04.2000)
US	60/199,020 (CIP)
Filed on	20 April 2000 (20.04.2000)
US	60/200,552 (CIP)
Filed on	28 April 2000 (28.04.2000)
US	60/202,348 (CIP)
Filed on	5 May 2000 (05.05.2000)
US	60/203,495 (CIP)
Filed on	11 May 2000 (11.05.2000)

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

(72) Inventors and

(75) Inventors/Applicants (for US only): REDDY, Roopa [IN/US], 1233 W. McKinley Avenue, #3, Sunnyvale, CA 94086 (US); THORNTON, Michael [US/US], 9 Midway Road, Woodside, CA 94062 2612 (US); BOROWSKY, Mark, L. [US/US], 122 Orchard Avenue, Redwood City, CA 94061 (US); TANG, V., Tom [US/US], 4330 Ramewick Court, San Jose, CA 95118 (US); KHAN, Farrah, A. [IN/US], 333 Escuela Avenue, #221, Mountain View, CA 94040 (US); TRIBOULEY, Catherine, M. [FR/US], 1121 Tennessee Street, #5, San Francisco, CA 94107 (US); GANDHI, Auneen, R. [US/US], 837 Roble Avenue, #1, Menlo Park, CA 94025 (US); YAO, Monique, G.

[US/US], 111 Frederick Court, Mountain View, CA 94043 (US); SAMANWALA, Madhus, S. [US/US], 210 Sylvia Court, Los Altos, CA 94024 (US); BAUGHN, Mariah, R. [US/US], 14244 Santiago Road, San Leandro, CA 94577 (US); NGUYEN, Daniel, B. [US/US], 1403 Ridgewood Drive, San Jose, CA 95118 (US); POLICKY, Jennifer, L. [US/US], 1511 Jarvis Court, San Jose, CA 95118 (US); YCE, Henry [US/US], 826 Lois Avenue, Sunnyvale, CA 94087 (US); SEILHAMER, Jeffrey, J. [US/US], 12555 La Cresta, Los Altos Hills, CA 94022 (US); WALIA, Narinder, K. [US/US], 890 Davis Street, #205, San Leandro, CA 94577 (US); LAL, Preeti [IN/US], P.O. Box 5142, Santa Clara, CA 95056 (US); KEARNEY, Liam [IE/US], 30 Woodside Avenue, San Francisco, CA 94127 (US); WALSH, Rudrick, T. [IE/GB], 3 Beaumont Court, St. Lawrence Road, Carletonbury, Kent CT1 3EZ (GB); LIU, Deyang, Aina, M. [US/US], 233 Coy Drive, San Jose, CA 95125 (US); LIU, Yan [CN/US], 3885 Corning Way, Palo Alto, CA 94303 (US); GREEFE, Barrie, D. [US/US], 1332 10th Avenue, #104, San Francisco, CA 94122 (US); RAUMANN, Brigitte, E. [US/US], 5801 South Dorchester Avenue, #3H, Chicago, IL 60637 (US); PATTERSON, Chandra [US/US], 490 Sherwood Way, #1, Menlo Park, CA 94025 (US).

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

(81) Designated States (national): AF, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CL, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GR, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, NZ, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

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

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

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

(54) Title: 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 01/77174 A2

WO 01/77174

PCT/US01/11206

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^+ , Na^+ , 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 01/77174

PCT/US01/11206

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 01/77174

PCT/US01/11206

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 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-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 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, 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 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 acid transport appears to occur via a high affinity, low capacity protein-mediated transport process.

WO 01/77174

PCT/US01/11206

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).

Mitochondrial carrier proteins are transmembrane-spanning proteins which transport ions and 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 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
15 Mitochondrial energy transfer proteins signature; Online Mendelian Inheritance in Man (OMIM)
*275000 Graves Disease).

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
20 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. (1999) J. Int. Med. 245:637-642).

Ion Channels

The electrical potential of a cell is generated and maintained by controlling the movement of
25 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
30 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
35 energy derived from ATP hydrolysis, they transport ions against the ion's concentration gradient.

WO 01/77174

PCT/US01/11206

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. P-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 (Forgacs, M. (1999) *J. Biol. Chem.* 274:12951-12954).

The resting potential of the cell is utilized in many processes involving carrier proteins and 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 Ca²⁺ 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⁺, Ca²⁺, and Cl⁻ channels) open their pores in response to changes in membrane potential; and ligand-gated channels (e.g.,

WO 01/77174

PCT/US01/11206

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

5 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).

20 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^{+} 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, $\beta 1$ and $\beta 2$. The $\beta 2$ subunit is a 35 integral membrane glycoprotein that contains an extracellular Ig domain, and its association with α

WO 01/77174

PCT/US01/11206

and $\beta 1$ 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; Egle, 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 Ca^{2+} and cAMP. In non-excitable 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 resting membrane 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. Jegla (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

WO 01/77174

PCT/US01/11206

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).

- 5 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
- 10 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
- 15 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
- 20 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 three subunits. The α_1 subunit forms the membrane pore and voltage sensor, while the $\alpha_2\delta$ and β subunits modulate the voltage-dependence, gating properties, and the current amplitude of the
- 25 channel. These subunits are encoded by at least six α_1 , one $\alpha_2\delta$, 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; McCleskey, E.W. (1994) *Curr. Opin. Neurobiol.* 4:304-312).

- The transient receptor family (Trp) of calcium ion channels are thought to mediate capacitative calcium entry (CCE). CCE is the Ca²⁺ influx into cells to resupply Ca²⁺ stores depleted
- 30 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 Ca²⁺ channels in the S3 through S6 regions. This suggests that Trp and/or related proteins may form mammalian CCE entry 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, and
- 35 whose expression in melanoma cells is inversely correlated with melanoma aggressiveness in vivo.

WO 01/77174

PCT/US01/11206

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).

5 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
10 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
15 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,
20 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
25 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
30 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
35 in determining channel permeation and selectivity (Sather, W.A. et al. (1994) Curr. Opin. Neurobiol.

WO 01/77174

PCT/US01/11206

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 (Ishii et al.,

5 *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
10 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
15 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 Ca²⁺ 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
20 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
25 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, P. and F.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
30 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. Bredt (1998) *Cell* 93:495-498).

Disease Correlation

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

WO 01/77174

PCT/US01/11206

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

5 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

10 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

15 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.H. 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.*

20 *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,

25 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

35 and allogenic responses in pigs, validating the idea of channel blockers as safe and efficacious

WO 01/77174

PCT/US01/11206

immunosuppressants (Cahalan, M.D. and K.G. Chandy (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," and "TRICH-15." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-15.

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

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a

WO 01/77174

PCT/US01/11206

biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

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

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). 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:16-30, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of

WO 01/77174

PCT/US01/11206

SEQ ID NO:16-30, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and
 5 which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

10 The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID
 15 NO:16-30, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

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

WO 01/77174

PCT/US01/11206

selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of 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:1-15, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting 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:1-15, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

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

WO 01/77174

PCT/US01/11206

acid sequence selected from the group consisting of SEQ ID NO:1-15, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15. The method comprises a) 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 sequence selected from the group consisting of SEQ ID NO:16-30, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, 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:16-30, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of

WO 01/77174

PCT/US01/11206

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.

5 **BRIEF DESCRIPTION OF THE TABLES**

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

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability score for the match between each
10 polypeptide and its GenBank homolog is also shown.

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

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

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

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

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

DESCRIPTION OF THE INVENTION

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

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

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

WO 01/77174

PCT/US01/11206

meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to 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

5 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

"TRICH" refers to the amino acid sequences of substantially purified TRICH obtained from

10 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

15 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

20 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

25 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

30 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

35 amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having

WO 01/77174

PCT/US01/11206

similar hydrophobicity values may include: asparagine and glutamine; and serine and threonine. 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 "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

WO 01/77174

PCT/US01/11206

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

5 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"

10 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,

15 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding TRICH or fragments of TRICH may be

20 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

25 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

30 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

35 as conservative amino acid substitutions.

WO 01/77174

PCT/US01/11206

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

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

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

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

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

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

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, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid

WO 01/77174

PCT/US01/11206

residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

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

A fragment of SEQ ID NO:1-15 is encoded by a fragment of SEQ ID NO:16-30. A fragment of SEQ ID NO:1-15 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-15. For example, a fragment of SEQ ID NO:1-15 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-15. The precise length of a fragment of SEQ ID NO:1-15 and the region of SEQ ID NO:1-15 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

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

WO 01/77174

PCT/US01/11206

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.

5 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
10 analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST
15 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

20 *Penalty for mismatch: -2*

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

25 *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
30 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
35 in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid

WO 01/77174

PCT/US01/11206

sequences that all encode substantially the same protein.

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

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

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

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

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

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

The term "humanized antibody" refers to an antibody molecule in which the amino acid

WO 01/77174

PCT/US01/11206

sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

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

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

WO 01/77174

PCT/US01/11206

hybridization complex may be formed in solution (e.g., C_{96} or R_{96} 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).

5 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
10 cellular and systemic defense systems.

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
15 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.

20 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
25 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
30 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.

35 PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript

WO 01/77174

PCT/US01/11206

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 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

WO 01/77174

PCT/US01/11206

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
 5 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
 10 unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

15 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
 20 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
 25 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,
 30 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.

35 An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear

WO 01/77174

PCT/US01/11206

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

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

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

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

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

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

WO 01/77174

PCT/US01/11206

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

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given 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

WO 01/77174

PCT/US01/11206

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.

5 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.

10 Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is
15 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
20 polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank homolog. Column 5 shows the annotation of the GenBank homolog along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

25 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
30 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 those
35 properties establish that the claimed polypeptides are transporters and ion channels. For example,

WO 01/77174

PCT/US01/11206

SEQ ID NO:5 is 98% identical, from residue M1 to residue I1009, to *Rattus norvegicus* glutamate receptor delta-1 subunit (GenBank ID g475542) 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:5 also contains a
 5 receptor family ligand binding region as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS and additional BLAST analyses provide further corroborative evidence that SEQ ID NO:5 is a glutamate receptor. SEQ ID NO:1-4 and SEQ ID NO:6-15 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ
 10 ID NO:1-15 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide
 15 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:16-30 or that distinguish between SEQ ID NO:16-30 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA
 20 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 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, 6431853H1 is the identification number of an Incyte cDNA sequence, and LUNGNON07 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., 71456748VI). Alternatively, the identification numbers in column 5
 30 may refer to GenBank cDNAs or ESTs (e.g., g5396013) which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. For example, GBI.g6742097_000003.fasta.edit is the identification number of a Genscan-predicted coding sequence, with g6742097 being the GenBank identification number of the sequence to which Genscan
 35 was applied. The Genscan-predicted coding sequences may have been edited prior to assembly. (See

WO 01/77174

PCT/US01/11206

Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and GenScan-predicted exons brought together by an "exon stitching" algorithm. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and GenScan-predicted exons brought together by an "exon-stretching" algorithm. (See Example V.) In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

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

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

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 NO:16-30 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:16-30. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of TRICH.

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

WO 01/77174

PCT/US01/11206

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.

5 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
10 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.

15 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.

20 Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:16-30 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kinmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in
25 "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
30 combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 (thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA
35 sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system

WO 01/77174

PCT/US01/11206

(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.)

- 5 The nucleic acid sequences encoding TRICH may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sackar, G. (1993) *PCR Methods Applic.* 2:318-322.)
- 10 Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-3060.) Additionally, one may use PCR, nested primers, and PROMOTER-FINDER libraries
- 20 (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of
- 25 about 68°C to 72°C.

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

30 into 5' non-transcribed regulatory regions.

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

35 emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate

WO 01/77174

PCT/US01/11206

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.

5 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.

10 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 Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. 20 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 25 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 30 homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding 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 35 Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.)

WO 01/77174

PCT/US01/11206

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 Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of 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. Rognier (1990) *Methods Enzymol.* 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, *supra*, pp. 28-53.)

In order to express a biologically active 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 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 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 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 control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. (See, e.g., Sambrook, J. et al. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

WO 01/77174

PCT/US01/11206

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);

5 plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, *supra*; Ausubel, *supra*; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO

10 J. 6:307-311; The McGraw-Hill Yearbook of Science and Technology (1992) McGraw-Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di

15 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. Sornia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending

20 upon the use intended for polynucleotide sequences encoding TRICH. For example, routine cloning, 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

25 transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of TRICH are needed, e.g. for the production of antibodies, vectors which direct high level expression of TRICH may be used. For example, vectors

30 containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of TRICH. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable

35 integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel,

WO 01/77174

PCT/US01/11206

1995, *supra*; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *BioTechnology* 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.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., *The McGraw Hill Yearbook of Science and Technology* (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding 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 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 DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of 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 confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine

WO 01/77174

PCT/US01/11206

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 chlorosulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci. USA* 77:3567-3570; Colbrete-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, e.g., Hampton, R. et al. (1990) *Serological Methods, a Laboratory Manual*, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) *Current Protocols in Immunology*, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) *Immunochemical Protocols*, Humana Press, Totowa NJ.)

WO 01/77174

PCT/US01/11206

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.

- 5 Alternatively, the sequences encoding TRICH, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega
10 (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

- Host cells transformed with nucleotide sequences encoding TRICH may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein
15 produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode 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
20 inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for
25 post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

- In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding TRICH may be ligated to a heterologous sequence resulting in translation of a
30 fusion protein in any of the aforementioned host systems. For example, a chimeric TRICH protein containing a heterologous moiety that can be recognized by a commercially available antibody may 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),
35 maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG,

WO 01/77174

PCT/US01/11206

c-myc, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the 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 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

WO 01/77174

PCT/US01/11206

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 Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) *Science* 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. et al. (1997) *Nucleic Acids Res.* 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

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

WO 01/77174

PCT/US01/11206

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 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 between regions of TRICH and transporters and ion channels. In addition, the expression of TRICH is closely associated with brain, prostate and thyroid tissues, neoplasms, and cancers of the small intestine. 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.

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, 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, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter.

WO 01/77174

PCT/US01/1206

- Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms,
- 5 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
- 10 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, cerebellar hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic
- 15 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, cataplexy, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses,
- 20 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
- 25 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, and acid maltase deficiency (AMD, also known as Pompe's disease); an immunological disorder such as acquired immunodeficiency syndrome
- 30 (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis,

WO 01/77174

PCT/US01/11206

erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

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.

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

WO 01/77174

PCT/US01/11206

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 technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030; and Cole, S.P. et al. (1984) *Mol. Cell Biol.* 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) *Proc.*

WO 01/77174

PCT/US01/11206

Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce TRICH-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotype composition, may be
 5 generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA
 10 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 the F(ab)₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and
 15 easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such
 20 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*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay
 25 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 affinities for multiple TRICH epitopes, represents the average affinity, or avidity, of the antibodies
 30 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 10^6 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
 35 immunopurification and similar procedures which ultimately require dissociation of TRICH.

WO 01/77174

PCT/US01/11206

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
 5 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.,
 10 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
 15 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
 20 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.B. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-473; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral
 25 vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, supra; Uekert, 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.*
 30 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
 35 immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency

WO 01/77174

PCT/US01/11206

- (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. Sumia (1997) *Nature* 339:239-242)), (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.
- 15 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. Récipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).
- 20 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 vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSII/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. 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 ecodyson-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding TRICH from a normal individual.

WO 01/77174

PCT/US01/11206

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 Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding 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 (Csote, M.E. et al. (1995) *Transplantation* 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy").

WO 01/77174

PCT/US01/11206

hereby incorporated by reference. For adenoviral vectors, see also Antinuzzi, P.A. et al. (1999) *Annu. Rev. Nutr.* 19:511-544 and Verma, I.M. and N. Somia (1997) *Nature* 389:239-242, both incorporated by reference herein.

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

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding 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, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for 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 (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy

WO 01/77174

PCT/US01/11206

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 Hoher, 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 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.

WO 01/77174

PCT/US01/1206

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
 5 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
 10 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
 15 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
 20 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
 25 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
 30 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 forming the basis for a comparison of the expression of the polynucleotide both with and without

WO 01/77174

PCT/US01/11206

exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.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 sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

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

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

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

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

WO 01/77174

PCT/US01/11206

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 ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

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

WO 01/77174

PCT/US01/11206

response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

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

DIAGNOSTICS

In another embodiment, antibodies which specifically bind 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 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

WO 01/77174

PCT/US01/11206

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.

5 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:16-30 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
10 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 ³²P or ³⁵S, or by enzymatic labels,
15 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
20 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,
25 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, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, cystinuria, iminoglycinuria, Hartup
30 disease, and Fanconi disease; a neurological disorder such as epilepsy, ischemic cerebrovascular

WO 01/77174

PCT/US01/1206

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, 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, and acid maltase deficiency (AMD, also known as Pompe's disease); 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 polyendocrinopathy-candidiasis-ectodermal dystrophy (APHCED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis,

WO 01/77174

PCT/US01/11206

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, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. 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 amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated,

WO 01/77174

PCT/US01/11206

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.

5 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
10 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
15 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
20 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,
25 from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplicons in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis
30 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
35 using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

WO 01/77174

PCT/US01/1206

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; Duplau, 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 Number 5,840,434, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

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*,

WO 01/77174

PCT/US01/1206

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

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

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

WO 01/77174

PCT/US01/11206

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 (Laeking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendez, L.G. et al. (1999) *Biotecniques* 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. Seifamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

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

WO 01/77174

PCT/US01/11206

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 with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent *in situ* hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulich, 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

WO 01/77174

PCT/US01/11206

(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. 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.

WO 01/77174

PCT/US01/11206

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.

5 The disclosures of all patents, applications and publications, mentioned above and below, in particular U.S. Ser. No. 60/195,595, U.S. Ser. No. 60/196,872, U.S. Ser. No. 60/199,020, U.S. Ser. No. 60/200,552, U.S. Ser. No. 60/202,348, and U.S. Ser. No. 60/203,495, are expressly incorporated by reference herein.

10 EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a
15 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
20 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
25 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
30 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.,
35 PBLUESCRIPT plasmid (Stratagene), PSPORI1 plasmid (Life Technologies), pCDNA2.1 plasmid

WO 01/77174

PCT/US01/11206

(Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRE, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

5

WO 01/77174

PCT/US01/11206

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
 5 AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4 °C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a
 10 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
 20 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
 25 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
 30 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
 35 BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family

WO 01/77174

PCT/US01/11206

databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Fddy, 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.

5 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 assemblies to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblies were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive
10 the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length
15 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.

20 Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column
25 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
30 ID NO:16-30. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 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., gbpri and gbhtg). Genscan is
35 a general-purpose gene identification program which analyzes genomic DNA sequences from a

WO 01/77174

PCT/US01/11206

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 *gbpri* public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from *genpept* to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

20 V. Assembly of Genomic Sequence Data with cDNA Sequence Data "Stitched" Sequences

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

35 Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or

WO 01/77174

PCT/US01/11206

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 gbprf public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

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

VI. Chromosomal Mapping of TRICH Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:16-30 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:16-30 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for 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

WO 01/77174

PCT/US01/11206

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

5 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.)

10 Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LJJHSEQ (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:

$$15 \quad \frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum}(\text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}))}$$

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
20 calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate
25 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%
30 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 Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is
35 classified into one of the following organ/tissue categories: cardiovascular system; connective tissue;

WO 01/77174

PCT/US01/11206

digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding TRICH. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte 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.

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), ELONGASH enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II

WO 01/77174

PCT/US01/11206

(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.

5 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
10 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.

15 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
20 recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or
25 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:16-30 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base
30 pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a
35 SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech).

WO 01/77174

PCT/US01/11206

An aliquot containing 10^5 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 (Nytan 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, $0.1 \times$ saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

10 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 (Schemm (1999), *supra*). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schemm, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.* 16:27-31.)

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

WO 01/77174

PCT/US01/11206

reverse transcribed using MMLV reverse-transcriptase, 0.05 µg/µl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/µl RNase inhibitor, 500 µM dATP, 500 µM dGTP, 500 µM dTTP, 40 µM dCTP, 40 µM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 µl volume containing 200 ng poly(A)⁺ RNA with

5 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 µl 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.

10 (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 µl 5X SSC/0.2% SDS.

Microarray Preparation

15 Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 µg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia

20 Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water,

25 and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 µl of the array element DNA, at an average concentration of 100 ng/µl, is loaded into the open capillary printing element by a high-speed robotic

30 apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

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

WO 01/77174

PCT/US01/11206

0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μ l of sample mixture consisting of 0.2 μ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

Detection

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

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

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

WO 01/77174

PCT/US01/11206

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

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

XI. Complementary Polynucleotides

Sequences complementary to the 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 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

Expression and purification of TRICH is achieved using bacterial or virus-based expression 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

WO 01/77174

PCT/US01/11206

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.)

- 5 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
10 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,
15 *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
20 vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a
25 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
30 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;
35 alterations in expression of cell surface and intracellular proteins as measured by reactivity with

WO 01/77174

PCT/US01/11206

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
 5 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
 10 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
 15 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
 20 described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. 11.)

- Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the
 25 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.

- 35 Media containing TRICH are passed over the immunoaffinity column, and the column is

WO 01/77174

PCT/US01/11206

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.

5 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. Sheng (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 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.

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

WO 01/77174

PCT/US01/11206

transformation under conditions appropriate for the cell line to allow expression and accumulation of TRICH and β -galactosidase.

Transformed cells expressing β -galactosidase are stained blue when a suitable colorimetric substrate is added to the culture media under conditions that are well known in the art. Stained cells 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 (Ishi 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 pBH₁, 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-glucuronate, 4 mM KCl, and 10 mM Hepes (pH 7.2). The intracellular solution is supplemented with varying concentrations of the TRICH mediator, such as cAMP, cGMP, or Ca²⁺ (in the form of CaCl₂), 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.

Transport activity of TRICH is assayed by measuring uptake of labeled substrates into *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 CaCl₂, 1mM MgCl₂, 1mM Na₂HPO₄, 5 mM Hepes, 3.8 mM NaOH, 50 μ g/ml gentamycin, pH 7.8) to allow expression of TRICH. Oocytes are then transferred to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM CaCl₂, 1mM MgCl₂, 10 mM Hepes/Tris pH 7.5). Uptake of various substrates (e.g., 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.

WO 01/77174

PCT/US01/11206

ATPase activity associated with TRICH can be measured by hydrolysis of radiolabeled ATP-
 [γ-³²P], separation of the hydrolysis products by chromatographic methods, and quantitation of the
 recovered ³²P using a scintillation counter. The reaction mixture contains ATP-[γ-³²P] and varying
 amounts of TRICH in a suitable buffer incubated at 37°C for a suitable period of time. The reaction
 5 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 ³²P liberated is counted in a scintillation counter. The
 amount of radioactivity recovered is proportional to the ATPase activity of TRICH in the assay.

XVIII. Identification of TRICH Agonists and Antagonists

10 TRICH is expressed in a eukaryotic cell line such as CHO (Chinese Hamster Ovary) or HEK
 (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
 15 membrane (Velicofeli, 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 Ca²⁺ 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
 20 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
 25 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
 30 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 01/77174

PCT/US01/11206

Table 1

Incyte Product ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
1784715	1	1784715CD1	16	1784715CE1
1784715	2	1784715CD2	17	1784715CE2
1784715	3	1784715CD3	18	1784715CE3
1784715	4	1784715CD4	19	1784715CE4
1784715	5	1784715CD5	20	1784715CE5
1784715	6	1784715CD6	21	1784715CE6
1784715	7	1784715CD7	22	1784715CE7
1784715	8	1784715CD8	23	1784715CE8
1784715	9	1784715CD9	24	1784715CE9
1784715	10	1784715CD10	25	1784715CE10
1784715	11	1784715CD11	26	1784715CE11
1784715	12	1784715CD12	27	1784715CE12
1784715	13	1784715CD13	28	1784715CE13
1784715	14	1784715CD14	29	1784715CE14
1784715	15	1784715CD15	30	1784715CE15

WO 01/77174

PCT/US01/11266

Table 2

Polypeptide Seq ID No.	Incyto Polypeptide ID	Genbank ID No.	Probability score	Genbank Homolog
1	1784775CD1	g4302730	4.00E-94	[Homo sapiens] multidrug resistance-associated protein 7
2	7472034CD1	g2317274	9.00E-168	[Homo sapiens] Aquaporin AQP9 Kuriyama, H. et al. (1997) Biochem. Biophys. Res. Commun. 241, 53-58
3	1878561CD1	g545998	1.10E-164	[Rattus sp.] tricarboxylate carrier rate, liver, Peptide Mitochondrial Partial, 357 aa Azzi, A. et al. (1993) J. Bioenerg. Biomembr. 25, 513-524
4	2246292CD1	g6045150	0	[Rattus norvegicus] TAP-like ABC transporter Kamaguchi, S. et al. (1999) FEBS Lett. 457, 231-236
5	5151730CD1	g475542	0	[Rattus norvegicus] glutamate receptor delta-1 subunit
6	7472384CD1	g220416	0	[Rus musculus] glutamate receptor channel subunit delta-1 Kawazaki, M. et al. (1997) Biochem. Biophys. Res. Commun. 193:886-892
7	7472384CD1	g7546839	0	[Cavia Porcellus] potassium channel beta-1 subunit Kawazaki, M. et al. (2000) J. Biol. Chem. 275:16550-16557
8	7473422CD1	g2687858	4.30E-179	[Pseudomonas aeruginosa] renal glutamate transporter Wolff, M.A. et al. (1997) Expression cloning and characterization of a renal organic anion transporter from winter flounder. FEBS Lett. 417:287-291
		g9950945	0	[Pseudomonas aeruginosa] iron (III)-transport system permease HBB Stover, C.K. et al. (2000) Nature 406:959-964

WO 01/77174

PCT/US01/11206

Table 2 (cont.)

Polypeptide Seq ID NO.	Protein Accession ID	GenBank ID NO.	Probability Score	GenBank Homology
9	2464715CD1	9531469	0	[Rattus norvegicus] renal osmotic stress-induced Na-Cl organic solute cotransporter
10	1734724CD1	93347922	5.00E-190	[Mus musculus] oribat transporter isoform AL2 Rash, S.R., et al. (1998) Cloning, gene structure and genomic localization of an oribat transporter from mouse kidney with 81% alternatively-spliced isoforms. <i>Receptors Channels</i> 6:113-128
11	1563237CD1	93980315	5.70E-46	[Oryctolagus cuniculus] hepatic gamma-glutamyl bile acid transporter
12	7473443CD1	94186073	3.20E-280	[Rattus norvegicus] peptide/histidine transporter Yasuda, T. et al. (1997) Cloning and functional expression of a brain peptide/histidine transporter. <i>J. Biol. Chem.</i> 272:10205-10211
13	7473438CD1	92465542	7.60E-84	[Homo sapiens] peptide transporter 3 [Mus musculus] calcium channel alpha-2-delta-C subunit [Homo sapiens] TWIK-related acid-sensitive K ⁺ channel Duprat, F. et al. (1997) hSK, a human background K ⁺ channel to sense external pH variations near physiological pH. <i>EMBO J.</i> 16:5464-5471
		91128686	0	[Homo sapiens] two pore potassium channel K _{ATP3.3}

WO 01/77174

PCT/US01/11206

Table 2 (cont.)

Polypeptide SEQ ID NO:	Triyle Polypeptide ID	GenBank ID NO:	Probability Score	GenBank Homolog
14	747728CUI	g30E473	1.90E-39	[Homo sapiens] calcium channel gamma subunit Rowers, P.A. et al. (1993) Molecular characterization of the gene encoding the gamma subunit of the human skeletal muscle 1,4- dihydropyridine-sensitive Ca ²⁺ channel (CACNG3), cDNA sequence, gene structure, and chromosomal location, J. Biol. Chem. 268:3275- 3279
15	747258CUI	g11177514	C	[Homo sapiens] tender pore domain potassium channel TRK-2 Rajan, S. et al. (2001) J. Biol. Chem. 276:7302-7311

Table 3

SWISS-PROT ID	Trypsinolytic polypeptide ID	Minimal Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
P17847	P17847:CD1	520	Y297, T11, S146, T192, T194, S252, S10		Transmembrane domain: 1-146; 127-171; 179-192; 193-252; 253-304; 305-346; 347-389; 390-429; 430-469; 470-509; 510-549; 550-589; 590-629; 630-669; 670-709; 710-749; 750-789; 790-829; 830-869; 870-909; 910-949; 950-989; 990-1029; 1030-1069; 1070-1109; 1110-1149; 1150-1189; 1190-1229; 1230-1269; 1270-1309; 1310-1349; 1350-1389; 1390-1429; 1430-1469; 1470-1509; 1510-1549; 1550-1589; 1590-1629; 1630-1669; 1670-1709; 1710-1749; 1750-1789; 1790-1829; 1830-1869; 1870-1909; 1910-1949; 1950-1989; 1990-2029; 2030-2069; 2070-2109; 2110-2149; 2150-2189; 2190-2229; 2230-2269; 2270-2309; 2310-2349; 2350-2389; 2390-2429; 2430-2469; 2470-2509; 2510-2549; 2550-2589; 2590-2629; 2630-2669; 2670-2709; 2710-2749; 2750-2789; 2790-2829; 2830-2869; 2870-2909; 2910-2949; 2950-2989; 2990-3029; 3030-3069; 3070-3109; 3110-3149; 3150-3189; 3190-3229; 3230-3269; 3270-3309; 3310-3349; 3350-3389; 3390-3429; 3430-3469; 3470-3509; 3510-3549; 3550-3589; 3590-3629; 3630-3669; 3670-3709; 3710-3749; 3750-3789; 3790-3829; 3830-3869; 3870-3909; 3910-3949; 3950-3989; 3990-4029; 4030-4069; 4070-4109; 4110-4149; 4150-4189; 4190-4229; 4230-4269; 4270-4309; 4310-4349; 4350-4389; 4390-4429; 4430-4469; 4470-4509; 4510-4549; 4550-4589; 4590-4629; 4630-4669; 4670-4709; 4710-4749; 4750-4789; 4790-4829; 4830-4869; 4870-4909; 4910-4949; 4950-4989; 4990-5029; 5030-5069; 5070-5109; 5110-5149; 5150-5189; 5190-5229; 5230-5269; 5270-5309; 5310-5349; 5350-5389; 5390-5429; 5430-5469; 5470-5509; 5510-5549; 5550-5589; 5590-5629; 5630-5669; 5670-5709; 5710-5749; 5750-5789; 5790-5829; 5830-5869; 5870-5909; 5910-5949; 5950-5989; 5990-6029; 6030-6069; 6070-6109; 6110-6149; 6150-6189; 6190-6229; 6230-6269; 6270-6309; 6310-6349; 6350-6389; 6390-6429; 6430-6469; 6470-6509; 6510-6549; 6550-6589; 6590-6629; 6630-6669; 6670-6709; 6710-6749; 6750-6789; 6790-6829; 6830-6869; 6870-6909; 6910-6949; 6950-6989; 6990-7029; 7030-7069; 7070-7109; 7110-7149; 7150-7189; 7190-7229; 7230-7269; 7270-7309; 7310-7349; 7350-7389; 7390-7429; 7430-7469; 7470-7509; 7510-7549; 7550-7589; 7590-7629; 7630-7669; 7670-7709; 7710-7749; 7750-7789; 7790-7829; 7830-7869; 7870-7909; 7910-7949; 7950-7989; 7990-8029; 8030-8069; 8070-8109; 8110-8149; 8150-8189; 8190-8229; 8230-8269; 8270-8309; 8310-8349; 8350-8389; 8390-8429; 8430-8469; 8470-8509; 8510-8549; 8550-8589; 8590-8629; 8630-8669; 8670-8709; 8710-8749; 8750-8789; 8790-8829; 8830-8869; 8870-8909; 8910-8949; 8950-8989; 8990-9029; 9030-9069; 9070-9109; 9110-9149; 9150-9189; 9190-9229; 9230-9269; 9270-9309; 9310-9349; 9350-9389; 9390-9429; 9430-9469; 9470-9509; 9510-9549; 9550-9589; 9590-9629; 9630-9669; 9670-9709; 9710-9749; 9750-9789; 9790-9829; 9830-9869; 9870-9909; 9910-9949; 9950-9989; 9990-10029; 10030-10069; 10070-10109; 10110-10149; 10150-10189; 10190-10229; 10230-10269; 10270-10309; 10310-10349; 10350-10389; 10390-10429; 10430-10469; 10470-10509; 10510-10549; 10550-10589; 10590-10629; 10630-10669; 10670-10709; 10710-10749; 10750-10789; 10790-10829; 10830-10869; 10870-10909; 10910-10949; 10950-10989; 10990-11029; 11030-11069; 11070-11109; 11110-11149; 11150-11189; 11190-11229; 11230-11269; 11270-11309; 11310-11349; 11350-11389; 11390-11429; 11430-11469; 11470-11509; 11510-11549; 11550-11589; 11590-11629; 11630-11669; 11670-11709; 11710-11749; 11750-11789; 11790-11829; 11830-11869; 11870-11909; 11910-11949; 11950-11989; 11990-12029; 12030-12069; 12070-12109; 12110-12149; 12150-12189; 12190-12229; 12230-12269; 12270-12309; 12310-12349; 12350-12389; 12390-12429; 12430-12469; 12470-12509; 12510-12549; 12550-12589; 12590-12629; 12630-12669; 12670-12709; 12710-12749; 12750-12789; 12790-12829; 12830-12869; 12870-12909; 12910-12949; 12950-12989; 12990-13029; 13030-13069; 13070-13109; 13110-13149; 13150-13189; 13190-13229; 13230-13269; 13270-13309; 13310-13349; 13350-13389; 13390-13429; 13430-13469; 13470-13509; 13510-13549; 13550-13589; 13590-13629; 13630-13669; 13670-13709; 13710-13749; 13750-13789; 13790-13829; 13830-13869; 13870-13909; 13910-13949; 13950-13989; 13990-14029; 14030-14069; 14070-14109; 14110-14149; 14150-14189; 14190-14229; 14230-14269; 14270-14309; 14310-14349; 14350-14389; 14390-14429; 14430-14469; 14470-14509; 14510-14549; 14550-14589; 14590-14629; 14630-14669; 14670-14709; 14710-14749; 14750-14789; 14790-14829; 14830-14869; 14870-14909; 14910-14949; 14950-14989; 14990-15029; 15030-15069; 15070-15109; 15110-15149; 15150-15189; 15190-15229; 15230-15269; 15270-15309; 15310-15349; 15350-15389; 15390-15429; 15430-15469; 15470-15509; 15510-15549; 15550-15589; 15590-15629; 15630-15669; 15670-15709; 15710-15749; 15750-15789; 15790-15829; 15830-15869; 15870-15909; 15910-15949; 15950-15989; 15990-16029; 16030-16069; 16070-16109; 16110-16149; 16150-16189; 16190-16229; 16230-16269; 16270-16309; 16310-16349; 16350-16389; 16390-16429; 16430-16469; 16470-16509; 16510-16549; 16550-16589; 16590-16629; 16630-16669; 16670-16709; 16710-16749; 16750-16789; 16790-16829; 16830-16869; 16870-16909; 16910-16949; 16950-16989; 16990-17029; 17030-17069; 17070-17109; 17110-17149; 17150-17189; 17190-17229; 17230-17269; 17270-17309; 17310-17349; 17350-17389; 17390-17429; 17430-17469; 17470-17509; 17510-17549; 17550-17589; 17590-17629; 17630-17669; 17670-17709; 17710-17749; 17750-17789; 17790-17829; 17830-17869; 17870-17909; 17910-17949; 17950-17989; 17990-18029; 18030-18069; 18070-18109; 18110-18149; 18150-18189; 18190-18229; 18230-18269; 18270-18309; 18310-18349; 18350-18389; 18390-18429; 18430-18469; 18470-18509; 18510-18549; 18550-18589; 18590-18629; 18630-18669; 18670-18709; 18710-18749; 18750-18789; 18790-18829; 18830-18869; 18870-18909; 18910-18949; 18950-18989; 18990-19029; 19030-19069; 19070-19109; 19110-19149; 19150-19189; 19190-19229; 19230-19269; 19270-19309; 19310-19349; 19350-19389; 19390-19429; 19430-19469; 19470-19509; 19510-19549; 19550-19589; 19590-19629; 19630-19669; 19670-19709; 19710-19749; 19750-19789; 19790-19829; 19830-19869; 19870-19909; 19910-19949; 19950-19989; 19990-20029; 20030-20069; 20070-20109; 20110-20149; 20150-20189; 20190-20229; 20230-20269; 20270-20309; 20310-20349; 20350-20389; 20390-20429; 20430-20469; 20470-20509; 20510-20549; 20550-20589; 20590-20629; 20630-20669; 20670-20709; 20710-20749; 20750-20789; 20790-20829; 20830-20869; 20870-20909; 20910-20949; 20950-20989; 20990-21029; 21030-21069; 21070-21109; 21110-21149; 21150-21189; 21190-21229; 21230-21269; 21270-21309; 21310-21349; 21350-21389; 21390-21429; 21430-21469; 21470-21509; 21510-21549; 21550-21589; 21590-21629; 21630-21669; 21670-21709; 21710-21749; 21750-21789; 21790-21829; 21830-21869; 21870-21909; 21910-21949; 21950-21989; 21990-22029; 22030-22069; 22070-22109; 22110-22149; 22150-22189; 22190-22229; 22230-22269; 22270-22309; 22310-22349; 22350-22389; 22390-22429; 22430-22469; 22470-22509; 22510-22549; 22550-22589; 22590-22629; 22630-22669; 22670-22709; 22710-22749; 22750-22789; 22790-22829; 22830-22869; 22870-22909; 22910-22949; 22950-22989; 22990-23029; 23030-23069; 23070-23109; 23110-23149; 23150-23189; 23190-23229; 23230-23269; 23270-23309; 23310-23349; 23350-23389; 23390-23429; 23430-23469; 23470-23509; 23510-23549; 23550-23589; 23590-23629; 23630-23669; 23670-23709; 23710-23749; 23750-23789; 23790-23829; 23830-23869; 23870-23909; 23910-23949; 23950-23989; 23990-24029; 24030-24069; 24070-24109; 24110-24149; 24150-24189; 24190-24229; 24230-24269; 24270-24309; 24310-24349; 24350-24389; 24390-24429; 24430-24469; 24470-24509; 24510-24549; 24550-24589; 24590-24629; 24630-24669; 24670-24709; 24710-24749; 24750-24789; 24790-24829; 24830-24869; 24870-24909; 24910-24949; 24950-24989; 24990-25029; 25030-25069; 25070-25109; 25110-25149; 25150-25189; 25190-25229; 25230-25269; 25270-25309; 25310-25349; 25350-25389; 25390-25429; 25430-25469; 25470-25509; 25510-25549; 25550-25589; 25590-25629; 25630-25669; 25670-25709; 25710-25749; 25750-25789; 25790-25829; 25830-25869; 25870-25909; 25910-25949; 25950-25989; 25990-26029; 26030-26069; 26070-26109; 26110-26149; 26150-26189; 26190-26229; 26230-26269; 26270-26309; 26310-26349; 26350-26389; 26390-26429; 26430-26469; 26470-26509; 26510-26549; 26550-26589; 26590-26629; 26630-26669; 26670-26709; 26710-26749; 26750-26789; 26790-26829; 26830-26869; 26870-26909; 26910-26949; 26950-26989; 26990-27029; 27030-27069; 27070-27109; 27110-27149; 27150-27189; 27190-27229; 27230-27269; 27270-27309; 27310-27349; 27350-27389; 27390-27429; 27430-27469; 27470-27509; 27510-27549; 27550-27589; 27590-27629; 27630-27669; 27670-27709; 27710-27749; 27750-27789; 27790-27829; 27830-27869; 27870-27909; 27910-27949; 27950-27989; 27990-28029; 28030-28069; 28070-28109; 28110-28149; 28150-28189; 28190-28229; 28230-28269; 28270-28309; 28310-28349; 28350-28389; 28390-28429; 28430-28469; 28470-28509; 28510-28549; 28550-28589; 28590-28629; 28630-28669; 28670-28709; 28710-28749; 28750-28789; 28790-28829; 28830-28869; 28870-28909; 28910-28949; 28950-28989; 28990-29029; 29030-29069; 29070-29109; 29110-29149; 29150-29189; 29190-29229; 29230-29269; 29270-29309; 29310-29349; 29350-29389; 29390-29429; 29430-29469; 29470-29509; 29510-29549; 29550-29589; 29590-29629; 29630-29669; 29670-29709; 29710-29749; 29750-29789; 29790-29829; 29830-29869; 29870-29909; 29910-29949; 29950-29989; 29990-30029; 30030-30069; 30070-30109; 30110-30149; 30150-30189; 30190-30229; 30230-30269; 30270-30309; 30310-30349; 30350-30389; 30390-30429; 30430-30469; 30470-30509; 30510-30549; 30550-30589; 30590-30629; 30630-30669; 30670-30709; 30710-30749; 30750-30789; 30790-30829; 30830-30869; 30870-30909; 30910-30949; 30950-30989; 30990-31029; 31030-31069; 31070-31109; 31110-31149; 31150-31189; 31190-31229; 31230-31269; 31270-31309; 31310-31349; 31350-31389; 31390-31429; 31430-31469; 31470-31509; 31510-31549; 31550-31589; 31590-31629; 31630-31669; 31670-31709; 31710-31749; 31750-31789; 31790-31829; 31830-31869; 31870-31909; 31910-31949; 31950-31989; 31990-32029; 32030-32069; 32070-32109; 32110-32149; 32150-32189; 32190-32229; 32230-32269; 32270-32309; 32310-32349; 32350-32389; 32390-32429; 32430-32469; 32470-32509; 32510-32549; 32550-32589; 32590-32629; 32630-32669; 32670-32709; 32710-32749; 32750-32789; 32790-32829; 32830-32869; 32870-32909; 32910-32949; 32950-32989; 32990-33029; 33030-33069; 33070-33109; 33110-33149; 33150-33189; 33190-33229; 33230-33269; 33270-33309; 33310-33349; 33350-33389; 33390-33429; 33430-33469; 33470-33509; 33510-33549; 33550-33589; 33590-33629; 33630-33669; 33670-33709; 33710-33749; 33750-33789; 33790-33829; 33830-33869; 33870-33909; 33910-33949; 33950-33989; 33990-34029; 34030-34069; 34070-34109; 34110-34149; 34150-34189; 34190-34229; 34230-34269; 34270-34309; 34310-34349; 34350-34389; 34390-34429; 34430-34469; 34470-34509; 34510-34549; 34550-34589; 34590-34629; 34630-34669; 34670-34709; 34710-34749; 34750-34789; 34790-34829; 34830-34869; 34870-34909; 34910-34949; 34950-34989; 34990-35029; 35030-35069; 35070-35109; 35110-35149; 35150-35189; 35190-35229; 35230-35269; 35270-35309; 35310-35349; 35350-35389; 35390-35429; 35430-35469; 35470-35509; 35510-35549; 35550-35589; 35590-35629; 35630-35669; 35670-35709; 35710-35749; 35750-35789; 35790-35829; 35830-35869; 35870-35909; 35910-35949; 35950-35989; 35990-36029; 36030-36069; 36070-36109; 36110-36149; 36150-36189; 36190-36229; 36230-36269; 36270-36309; 36310-36349; 36350-36389; 36390-36429; 36430-36469; 36470-36509; 36510-36549; 36550-36589; 36590-36629; 36630-36669; 36670-36709; 36710-36749; 36750-36789; 36790-36829; 36830-36869; 36870-36909; 36910-36949; 36950-36989; 36990-37029; 37030-37069; 37070-37109; 37110-37149; 37150-37189; 37190-37229; 37230-37269; 37270-37309; 37310-37349; 37350-37389; 37390-37429; 37430-37469; 37470-37509; 37510-37549; 37550-37589; 37590-37629; 37630-37669; 37670-37709; 37710-37749; 37750-37789; 37790-37829; 37830-37869; 37870-37909; 37910-37949; 37950-37989; 37990-38029; 38030-38069; 38070-38109; 38110-38149; 38150-38189; 38190-38229; 38230-38269; 38270-38309; 38310-38349; 38350-38389; 38390-38429; 38430-38469; 38470-38509; 38510-38549; 38550-38589; 38590-38629; 38630-38669; 38670-38709; 38710-38749; 38750-38789; 38790-38829; 38830-38869; 38870-38909; 38910-38949; 38950-38989; 38990-39029; 39030-39069; 39070-39109; 39110-39149; 39150-39189; 39190-39229; 39230-39269; 39270-39309; 39310-39349;	

WO 01/77174

PCT/US01/11206

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
3	1878591C01	322	T30 T64 T142 T296 S309 S46 T134	N100 N124 N132 N135	PROTEIN TRANSMEMBRANE CHROMOSOME PUTATIVE TRANSPORTER P000696:L5-L251 ABC Transporter: L600-L614	BLAST_PRODOM
4	3246292C01	723	T139 S161 T209 T311 T377 T469 S300 S720 S154 Y359 S28 S31 S46 T103 T181 S275 T367 S528 S552 S628 S659	N280 N405 N441 N556 N716	ATP/GTP-binding site motif A: G496-S503 transmembrane domain V85-P104, V185-P294, L329- G347 ABC transporter: G689-G673 ABC transporter transmembrane region: L188-G450 ABC transporters family SL00211A:K454-C505 SL00211B:L600-D631 signature: L582-S631 ATP-BINDING TRANSPORT PROTEIN TRANSMEMBRANE GLYCOPROTEIN TRANSPORTER: P000130:V229- A46 M22 SNOWIN DMG7894 Q05181 1-695:L396-V516 DMG7894 S255711-703:V422-X698 ABC TRANSPORTERS FAMILY DMG0008 Q05181 502-717:V461- G673 DMG0008 S11426 479-695:V461- G673	MOTIFS IMMER EXMER_FFAM BLIMPS_BLOCKS PROFLIMSCAN BLAST_PRODOM BLAST_DOMO

WO 01/77174

PCT/US01/11206

Table 3 (cont.)

SEQ ID NO.	Invertebrate Polypeptide*	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Platforms
5	S151730C01	1005	S17, S49, S97, S123, S171, S180, T211, T223, S462, S287, S335, T336, S366, T338, S400, S417, S428, T443, T438, S510, S528, S533, S559, T567, T681, T704, T724, S765, T752, S765, S783, S825, S857, T864, S882, S901, T939, Y220, Y679	S200, S422, N698	catecholergic glucuronate receptor: P60099JC; S518-S533 L-methyl-D-aspartate (NMDA) receptor signature: P600177D; I632-S659 HYDROXYTRYPTAMINE 5A RECEPTOR (5-HT-2A) / (5-HT-2B) RECEPTOR (5-HT-2); P600177D; I532-S659 GLUTAMATE RECEPTOR SUBUNIT DELTA2 DELTA CHANNEL SUBTYPE DELTA2 P000864; K83-L437 P000864; K83-L437 P000500; S565-T949, P000273; L437-S550 GLUTAMATE RECEPTOR: TM08727 S28857 L-421-M1-S422, TM00747 S28857 S27-697-S677-1998; DM08722 S28858 2-425-A20-M422; DM00247 S28858 528-900-S627-1998 Signal peptide: M1-F20 Transmembrane Domain (transmembrane domain): L557-L583, S810-1847 Receptor family ligand binding region (ANF receptor): I12-S424	BLUES-TRINIS

Table 3 (cont.)

REF ID	Protein Name	Amino acid residues	Potential phosphorylation sites	Potential glycosylation sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
6	747258/CD1	374	S21, S55, S127, S179, S251, S319, S331, S341, S360, S373.	N53	TASK K ⁺ channel. Signature: P801095A: V6-V25, P801095H: F26-F51, P801095C: S63-Q77, P801095D: Q136-R-45, P801095B: C146-T161, P801095F: Q208-V210, P801095G: G236-E254, P801095H: S283-D366, P801095K: F348-V374.	BLAST-FASTA
7	747253/CD1	519	S94, T121, T176, S192, S246, S308, S362, S376, T433, T513, S563, S564	N54, N74, N99, N96, N131, N149	CHANNEL IONIC TWIRRELATED ACIDSENSITIVE K ⁺ CHANNEL F340G.3 PROTEIN PUTATIVE POTASSIUM: P001302G: N1-N74 Signal Peptide (Signal peptide): N1-V75 Transmembrane Domain (Transmembrane domain): F225-V243 TASK K ⁺ channel (TASK channel): N1-V374 Ion transport Protein (Ion trans): F34-L344 Sugar transport Proteins: P8001351: S58142-145: 478: R210-E249 (p = 6.1e-06) Transmembrane domain: G460-M425, T428-E447 Sugar (and other) transporter: S133-A553	BLAST-DMO HMMER-TEMSMEMBRANE HMMER-PPAN
					Transporter: organic cation Transport protein: putative ion -small cationic Transmembrane protein: P0003661: N31-M-25	BLAST-PRUXM

WO 01/77174

PCT/US01/11206

Table 3 (cont.)

SEQ ID NO.	Protein Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
8	7973422CD1	549	S2, S19, S49, S55, T182, T188, S189, S193, T449, T541		Binding-protein-dependent transport signature: BPD Transp: F182-E252	HMMER-SPAM
					Binding-protein-dependent transport systems inner membrane component: Bpd_transp_inn_membr.pri: L432-E231	PROFILERSCAN
					Protein: Iron Ion III transport system, permease transport, transmembrane inner membrane F178 F203L554; K46-L483	BLAST-PRODM
					LOWAIN, SPID, LKON, TRANSPORT, PERITPLASMIC, UNQ5963/P64048/210-473; A316-L540	BLAST-DMO
					BPD Transporter Integral Membrane protein signature: L183-E214, L443-5471	NOTIFS
					Transmembrane domain: E48-V70; N97-A117; M262-L279; V315-S334, V356-V375	HMMER-TRANSMEMBRANE

WO 01/77174

PCT/US01/11206

Table 3 (cont.)

SSQ ID No.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Database
9	2864715cd1	636	S17, S30, S100, S143, S170, S186, T212, T219, T243, T267, T310, S283, T233, S261, S273	N158 N162 N258 N334 N368	SODIUM/NEUROTRANSMITTER SYNAPTOR FAMILY: TM00772 850998 19-616:831-6614 Sodium neurotransmitter: BL00610A: Q40-E99; E-00610B: W104-W153; R100610C: W193- G246; BL00610D: E261-T313; BL00610E: A405-V448; BL00610F: R502-Q850; BL00610G: L576-P598 Sodium neurotransmitter synaptoxin family signature: neurotransm_c1nasp_1: D36-P90 SODIUM/NEUROTRANSMITTER: S00016A: Q0016A; R10175B: A59-186; P00176C: G10176D: R00176D: P232-T339; P00176E: V304-V324; R600175F: M111- T429; P00176G: S491-V311; P00176H: Q531-V551 TRANSPORTER NEUROTRANSMITTER TRANSPORT TRANSMEMBRANE SYNAPTOR GLYCOPROTEIN SODIUM CHLORIDE/DEPENDENT SODIUM/DEPENDENT GABA: P0000448:C383-1606 Transmembrane Domain: Y70-L101, F123-E-41, I305-V124, S417- V440, L402-V480, T531-P550 Sodium neurotransmitter: synaptoxin family (SNP): R32- R332, L378-N608	BLAST-DBD BLIMPS-BLOCKS PROFILES-AN BLIMPS-PRINTS BLAST-PRODOM HOMER- TRANSMEMBRANE HOMER-PPM

WO 01/77174

PCT/US01/11206

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
10	173474CD1	491	T227, S233, S439, T475	N5, N18, N24, N180	SODIUM; ACTD; BILE; TRANSPORTER DOMAIN; P003972;138655[8-318;S251-Y439] Transmembrane domain: V241-L261, L288-F307, W325-4243, K416-1435 Sodium Bile acid symporter family (SMP): L113-E145 (score=6.0, e-value=0.00017) ACID COTRANSPORTING POLYPEPTIDE; SODIUM/BILE COTRANSPORTER; Na+/BILE SODIUM/TAUROCHOLATE TRANSMEMBRANE TRANSPORT SYMPORTER: P00753;13742-Y439 PDR2 family (putative protein Smp28.1, P00074190; BL01022A:262-3160; BL01022B:A72-1117; BL01022C:G164-Y187; BL01022D:2199-Y211; BL01022E:2416-2451) Glucose transporter signature: PR00172C: V241-L261.	SLAST-DOM HMMER HMMER-PPAM SLAST-BROADOM
						BLIPS-BLOCKS
						BLIPS-PRINTS

WO 01/77174

PCT/US01/11206

Table 3 (cont.)

SEQ ID NO	Incyto Polypeptide NO	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
11	1563237CD1	525	S281, S302, S367, T195, S174, S371, S28-, S371	N61, N66, N178, N223, N356, N383	PEPTIDE OLIGOPEPTIDE TRANSFER; PROTEIN SYMPORT ISOFORM H/PEPTIDE COTRANSFER; PD001559; Y101-S363	BLAST-PRODOM
12	747443CD1	1310	S52, T81, T96, S108, S109, S113, S176, S223, S138, S272, S480, S422, S456, S511, S556, S501, T725, S748, S761, T738, T780, S897, T913, S976, S1026, S1052, S1062, T1125, S1126, T1150, S1160, T1170, T1235, T1239, T1245, T1287	N16, N457, N661, N979, T124, T206	PTR? FAMILY PROTON/OLIGOPEPTIDE SYMPORTERS: M01990/P46032 46-551:542-L269 POT family (putative protein transporter) PTR3: Y101-S259; PTR2: N396-S447 Similarity to calcium channels: D806895/P54289, 461-591:S759-L1187 CHANNEL CALCTUM PRMUNSON; TONIC TRANSMEMBRANE ION TRANSFERT; VOLTAGE-GATED DIHYDROPIRINDONESENSITIVE LITIFB: P0013637; I631-K750	BLAST-PRODOM BLAST-DOCK BLAST-PRODOM

WO 01/77174

PCT/US01/11206

Table 3 (cont.)

Protein No.	Protein Name	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
13	7472436CD1	400	S75 S101 S125 T162 S197 S249 S7 S105 S175 S312		Transmembrane domains: S75-F96; F295-V313 TASK K ⁺ channel domain: W71-L50C TASK K ⁺ channel signature: PRO1095; V76-V95; F96-K121; S133-Q147; Q196-C215; Q216-V231; Q279-L400; G306-P324 CHANNEL PROTEIN IONIC POTASSIUM SUBUNIT K ⁺ POSITIVE SUBFAMILY K MEMBER P0021430: L224-V312	HMMER HMMER-PPAM BLIPS-PRINTS BLAST-PRODOM
14	7472436CD1	260	T121 T133 T67 T66 T132 S195	M72 N113	Signal peptide: M1-L53 Transmembrane domains: V146-S164; F165-V187; V235-Q242 CALCIUM CHANNEL GAMMA SUBUNIT ETHYLOPYRIDINESENSITIVE SKELTAL MUSCLE IONIC TRANSMEMBRANE P015649: L31-C223	SPSCAN HMMER BLAST-PRODOM
15	7472538CD1	489	S257 S373 S424 S58 S62 S260 T295 S416 S429 S2 S5 S13 T36 W75 T149 T156 T252 S436	N78	Transmembrane domains: F204-F223; F476-M293; F340-L359 TASK K ⁺ channel domain: L42-F484 (Score = -115.2; E-value = 6.0078)	HMMER HMMER-PPAM

WO 01/77174

PCT/US01/11206

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment (s)	Sequence Fragments	5' Position	3' Position
16	1784775CB1	1735	233-901, 1- 59, 1252- 1513	60118407D2 60209.0201 675053J1 (SINIFER02) 6431853k1 (LUNG00007) 601.96550429.000023. Easta.edic.toup Easta.edic	948 1053 1057 1021 1 1	1303 1416 1735 1021 771 1041
17	7473334CB2	1041	158-380	694.96742097.000003. Easta.edic	1	1041
18	1878581CB1	2367	1-42, 1275- 2367	7145674RV1 71461790V1 71465463V1 71464816V1 71465977V1 71456476V1 71466268V1 4202150H1 (BEATUT029) 5511403F6 (HEARET03) 50111359D2 95596013 38-1252F6 (CONTUT01) 2246292R6 (HIFON002) 5890917J1 (BEATUT03) 7168417H1 (MCLRM001) 4651312T6 (BRAKOT02) 5120730H1 (BRAKON05) 3334157H1 (BRALF010) 4651112F6 (BEAENC02) 5015994R1 (BEAENC03) 2395385F6 (THP1A2T01)	238 1602 300 1499 807 1723 902 1 1747 1 2884 1387 2978 650 100 2689 2365 549 1959 1713 1026	852 2235 263 2193 1533 2367 1564 280 2162 473 3333 1293 648 3320 3044 918 2552 1968 1450
19	2246292CB1	3343	1-867, 2606-2809, 1501-1883, 960-1013			

WO 01/77174

PCT/US01/11206

Table 4 (cont.)

Polynucleotide Seq ID NO:	Inbyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
20	5151730CB1	357	863-1649, 1-113, 2288-3339	688590J71 (BRANDT001) 5151740F6 (HARRIS01) 1675978T6 (BLAND003) 3967290H1 (LUNGS023) 4466394H1 (ROSDIN01) 6974079F8 (HARRIS06) 7288330H1 (HARRIS06) 6250428F8 (LUNPLU02) 7356591H2 (SPURIN01) 7174088H1 (BRISTIN01) 3486391T6 (EPICIN01) 6891842H1 (BRATID03) 463149H1 (TILYANC02) 6211961G4987-row 92001843	2374 504 2973 3381 2229 1855 1855 183 804 1082 2021 1859 959 1 946	2932 527 3502 3517 2495 582 2316 456 1409 1683 3440 1993 1128 1248
21	7472584CB1	1248	751-945, 96-219	7693385H2 (MODTUE01) 7693385J2 (MODTUE01) GNG_45776606_000002_002 edit	1 253 1 1	743 879 1770
22	7472536CB1	1770	1-484, 1136-1770, 573-1030	GNG_45523050_000002_006 506	1	2544
23	7473422CB1	2544	1-2544	70870289V1 70520541V1	1771 550	2335 1154
24	2864715CB1	2871	1-32, 750- 1421, 1950- 2871	6793918H1 (SINITMC01) 6829267H1 (SINITMC01) 70792499V1 6753311H1 (SINITMC01) 6753435H2 (SINITMC01) 1754724F6 (COLINOT22) 60123116D1	1254 1088 2213 2261 1710 1963 502	1738 1615 2871 2261 466 1704 987
25	1734724CB1	2141	658-817	4222214H1 (PANCOTC7) 3495639T7 (ASREPT07) 3495639F6 (ASREPT07) GNG_47344269_001021_002 6107668H1 (BRANDT03)	1942 2525 963 69 1	2141 1520 1551 1544 545

WO 01/77174

PCT/US01/11206

Table 4 (cont.)

Polynucleotide Seq ID NO:	Inbyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
26	1561337CB1	1902	1-1283	7147258H1 (CDLNIVL) 2207658F6 (SINTPEK3) 1563237G6 (SFLNRCV01) 2118671T6 (SRSPCTU02) 2118671T6 (SRSPCTU02) 7679862H1 (BRAPFUE01) 71152376V2 FL722155_00001.raw 7139553V1 GHL-Q3810573.raw 71151715V1 71158575V1 GEL-Q3810573.ed1c 5758248H1 (HEANOK01) 220573F6 (PROSNOT16) FL767199_00001 CP2_991027_R15_masked fa.Cont-1052563	689 8-8 1312 1410 1 3186 1864 3713 1 3338 2741 634 1907 1541 722 1 1	1258 1379 1899 1902 722 3857 2161 4125 804 3852 3345 3059 2460 2140 2974 1162
27	7473443CB1	4125	1-41, 3763- 4125, 407- 435, 491- 2466	GNN-Q7523629_000001_002 7259171H1 (ESAIFPK06) GNN-Q6560849_000001_002	1 1 285 285	896 599 2080
28	7473438CB1	2460	1528-1717, 1-374, 998- 1020, 515- 540, 2181- 2460, 1182- 1278			
29	7474286CB1	896	86-487			
30	7472589CB1	2080	1-492, 1002-1177, 618-691			

WO 01/77174

PCT/US01/11206

Table 5

Polynucleotide Seq. ID No.	Project ID	Representative Library
18	784775CB1	BRADIT010
19	2878581CB1	SPLARET02
20	2246292CB1	HIPOM002
21	5151730CB1	PROST007
22	7472584CB1	TYMOT02
23	7472536CB1	LMOT001
24	2864715CB1	SIMTUT03
25	1734724CB1	ADRETT07
26	1553237CB1	DNGMOT01
27	7473443CB1	THPLAZS08
28	7473434CB1	PROSMUT16
29	7474486CB1	PMVMUT03
30	7472589CB1	BRATFER06

Table 6

Library	Vector	Library Description
BR007001	pINCY	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 43-year-old Caucasian female during a unilateral adrenalectomy. Pathology indicated pheochromocytoma.
BR018006	pCMV2.1	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.
BR007010	pINCY	Library was constructed using RNA isolated from diseased cerebellar tissue removed from the brain of a 74-year-old Caucasian male, who died from Alzheimer's disease.
UNGL0001	pINCY	Library was constructed using RNA isolated from dorsal root ganglion tissue removed from the cervical spine of a 32-year-old Caucasian male who died from acute pulmonary edema and bronchopneumonia. Bilateral pleural and peritoneal effusions, and malignant lymphoma (natural killer cell type). Patient history included probable cytomegalovirus infection, hepatic congestion and steatosis, splenomegaly, hemorrhagic cystitis, thyroid hemorrhage, and Bell's palsy. Surgeries included colonoscopy, large intestine biopsy, adenotomectomy, and nasopharyngeal endoscopy and biopsy. Treatment included radiation therapy.
HY000002	pSPORT1	This normalised hippocampus library was constructed from 1.13M independent clones from a hippocampus tissue library. RNA was isolated from the hippocampus tissue of a 72-year-old Caucasian female who died from an intracranial bleed. Patient history included nose cancer, hypertension, and arthritis. The normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994), 91, 9228).
LR000701	pCMV2.1	This 5' biased random primed library was constructed using RNA isolated from lymph node tumor tissue removed from a 67-year-old Caucasian male during regional lymph node excision, open biopsy of tongue, and partial glossectomy. Pathology indicated metastatic basaloid squamous cell carcinoma. Pathology for the associated tumor tissue indicated basaloid squamous cell carcinoma forming an ulcerated mass in the base of the tongue. Tumor deeply invaded the underlying skeletal muscle. The patient presented with tongue cancer. Patient history included benign hypertension and alcohol abuse. The patient was not taking any medications. Family history included diabetes in the son.

WO 01/77174

PCT/US01/11266

Table 6 (cont.)

Library	Vector	Library Description
PROST016	pINCY	Library was constructed using RNA isolated from diseased prostate tissue removed from a 68-year-old Caucasian male during a radical prostatectomy. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 3+4). The patient presented with elevated prostate specific antigen (PSA) during this hospitalization, the patient was diagnosed with metastatic prostate cancer. Patient history included osteoarthritis, and type II diabetes. Family history included benign hypertension, acute myocardial infarction, hyperlipidemia, and arteriosclerotic coronary artery disease.
PROST007	pINCY	The library was constructed using RNA isolated from diseased prostate tissue removed from a 73-year-old Caucasian male during radical prostatectomy. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated adenocarcinoma, Gleason 3+3, involving the left side peripherally and anteriorly. The tumor perforated the capsule to involve periprostatic tissue. The tumor margins on the left. The patient presented with elevated prostate specific antigen. Patient history included bladder cancer, speech disturbance and acquired spondylolisthesis. Family history included benign hypertension and cerebrovascular disease.
PTUT003	pINCY	Library was constructed using RNA isolated from the left parathyroid tissue of a 69-year-old Caucasian female during a partial parathyroidectomy. Pathology indicated hyperplasia. The patient presented with primary hyperparathyroidism.
SINF0033	pINCY	Library was constructed using RNA isolated from ileal tumor tissue obtained from a 49-year-old Caucasian female during destruction of peritoneal tissue, peritoneal adhesiolysis, ileum resection, and permanent colostomy. Pathology indicated grade 4 adenocarcinoma. Patient history included benign hypertension. Previous surgeries included total abdominal hysterectomy, bilateral salpingo-oophorectomy, regional lymph node excision, an incidental appendectomy, and dilation and curettage. Family history included benign hypertension, and dilation and curettage. Family history included benign hypertension, cerebrovascular disease, hyperlipidemia, atherosclerotic coronary artery disease, hyperlipidemia, type II diabetes, and stomach cancer.
SPLATE002	pINCY	Library was constructed using RNA isolated from spleen tissue removed from a Caucasian male fetus, who died at 23 weeks' gestation.

WO 01/77174

PCT/US01/11206

Table 6 (cont.)

Library	Vector	Library Description
THEJAZS08	ESDRT1	This sub-actd THP-1 promonocyte cell line library was constructed using 5.76 x 10 ⁵ clones from a 5-aza-2'-deoxycytidine (AZ) treated THP-1 cell library. Starting RNA was made from THP-1 promonocyte cells treated for three days with 0.8 micromolar AZ. The donor had acute monocytic leukemia. The hybridization probe for subtraction was derived from a similarly constructed library, made from 1 microgram of polyA RNA isolated from untreated THP-1 cells. 5.76 million clones from the AZ-treated THP-1 cell library were then subjected to two rounds of subtractive hybridization with 5 million clones from the untreated THP-1 cell library. Subtractive hybridization conditions were based on the methodologies of Seacoop et al., <i>MGR</i> (1991) 19:1934, and Ronaldo et al., <i>Genome Research</i> (1996) 6:791.
TELYNCT02	Lambda UNIZAP	Library was constructed using RNA isolated from non-adherent peripheral blood mononuclear cells. The blood was obtained from unrelated male and female donors and treated with LPS for 6 hours. Cells from each donor were purified on Ficoll Hypaque, then harvested by centrifugation, lysed in a buffer containing DNase, and spun through a QIAGEN RNeasy spin column for RNA extraction. Total RNA was isolated using a QIAGEN RNeasy spin column. PolyA RNA was isolated using an XhoI-clamp (7) primer. Double-stranded cDNA was blunt-ended, ligated to 3' end adaptors, digested with XhoI, size selected, and cloned into the XhoI and EcoRI sites of the lambda UNIZAP vector.

WO 01/77174

PCT/US01/11206

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABIPARACEL.PDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Perceci Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESYs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA compares at least five functions: fasta, fasta, fastx, fastx, and ssearch.	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.	ESYs: fasta E value=1.0E-6 Assembled ESYs: 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=100 or greater
BLUMPS	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. 235:1301-1332; Sammariner, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View in a Nuclei, Cambridge Univ. Press, pp. 1-350.	PFAM hit: Probability value= 1.0E-5 or less Signal peptide hits: Score= 0 or greater

WO 01/77174

PCT/US01/11206

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Rainieri, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score-GOG-specified "TRGI" value for that particular Prosite motif. Generally, score=1.42.1.
Phred	A base-calling algorithm that examines unsubmitted sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phil's Revised Assembly Program including SWAT and GrassMatch programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Scores=120 or greater; Match length=56 or greater
Conseq	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nelson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Scores=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:365-371.	
TMMIMER	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.	
Modis	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Ramach, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 01/77174

PCT/US01/11206

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence selected from the group consisting of
 - 5 SEQ ID NO:1-15,
 - b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:4 and SEQ ID NO:6-15,
 - c) a naturally occurring polypeptide comprising an amino acid sequence at least 92% identical to the amino acid sequence of SEQ ID NO:2,
 - 10 d) a naturally occurring polypeptide comprising an amino acid sequence at least 96% identical to the amino acid sequence of SEQ ID NO:3,
 - e) a naturally occurring polypeptide comprising an amino acid sequence at least 99% identical to the amino acid sequence of SEQ ID NO:5,
 - 15 f) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, and
 - g) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.
- 20 2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-15.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
- 25 4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:16-30.
- 30 6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
7. A cell transformed with a recombinant polynucleotide of claim 6.

WO 01/77174

PCT/US01/11206

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method for producing a polypeptide of claim 1, the method comprising:
- a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said
5 cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide
comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of
claim 1, and
- b) recovering the polypeptide so expressed.
10. An isolated antibody which specifically binds to a polypeptide of claim 1.
11. An isolated polynucleotide selected from the group consisting of:
- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting
of SEQ ID NO:16-30,
- 15 b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90%
identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:16-30,
- c) a polynucleotide complementary to the polynucleotide of a),
- d) a polynucleotide complementary to the polynucleotide of b), and
- 20 e) an RNA equivalent of a)-d).
12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a
polynucleotide of claim 11.
13. A method for detecting a target polynucleotide in a sample, said target polynucleotide
25 having a sequence of a polynucleotide of claim 11, the method comprising:
- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides
comprising a sequence complementary to said target polynucleotide in the sample, and which probe
specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization
complex is formed between said probe and said target polynucleotide or fragments thereof, and
- 30 b) detecting the presence or absence of said hybridization complex, and, optionally, if
present, the amount thereof.
14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

WO 01/77174

PCT/US01/11206

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
- 5 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

10

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.

18. 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 16.

15

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- 20 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting agonist activity in the sample.

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

25

21. 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 claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- 30 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

WO 01/77174

PCT/US01/11206

23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional TRICH, comprising administering to a patient in need of such treatment a composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

- a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

28. A method for assessing toxicity of a test compound, said method comprising:

- a) treating a biological sample containing nucleic acids with the test compound;

WO 01/77174

PCT/US01/11206

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;

c) quantifying the amount of hybridization complex; and

d) comparing the amount of hybridization complex in the treated biological sample with the 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.

29. A diagnostic test for a condition or disease associated with the expression of TRICH in a biological sample comprising the steps of:

a) combining the biological sample with an antibody of claim 10, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and

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

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

a) a chimeric antibody,

b) a single chain antibody,

c) a Fab fragment,

d) a F(ab')₂ fragment, or

e) a humanized antibody.

31. A composition comprising an antibody of claim 10 and an acceptable excipient.

32. A method of diagnosing a condition or disease associated with the expression of TRICH in a subject, comprising administering to said subject an effective amount of the composition of claim 31.

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

34. A method of diagnosing a condition or disease associated with the expression of TRICH

WO 01/77174

PCT/US01/11206

in a subject, comprising administering to said subject an effective amount of the composition of claim 33.

35. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 5 10 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibodies from said animal; and

10 c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.

36. An antibody produced by a method of claim 35.

15

37. A composition comprising the antibody of claim 36 and a suitable carrier.

38. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising:

20 a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibody producing cells from the animal;

25 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 having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.

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

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

41. The antibody of claim 10, wherein the antibody is produced by screening a library

WO 01/77174

PCT/US01/11206

expression library.

42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant immunoglobulin library.

5

43. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15 in a sample, comprising the steps of:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

10 b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15 in the sample.

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15 from a sample, the method comprising:

15 a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-15.

20

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

46. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

25 47. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

48. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

49. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

30

50. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

51. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

WO 01/77174

PCT/US01/11206

52. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.
53. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.
54. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.
55. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.
56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.
60. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:16.
61. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:17.
62. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:18.
63. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:19.
64. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:20.
65. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:21.

WO 01/77174

PCT/US01/11206

66. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:22.

67. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
3 NO:23.

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

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

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

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

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

73. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:29.

74. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
25 NO:30.

WO 01/77174

PCT/US01/11206

```

<110> INCYTE GENOMICS, INC.
      REDDY, Roopa
      THORNTON, Michael
      BOROWSKY, Mark L.
      TANG, Y. Tom
      KHAN, Farrah A.
      TREBOULEY, Catherine M.
      GANDHI, Sreera R.
      YAO, Monique G.
      SARJANWALA, Raghu S.
      BAUGHN, Mariah R.
      NGUYEN, Daniel B.
      POLICKY, Jennifer L.
      YUE, Henry
      SEILHAMER, Jeffrey J.
      SALLA, Narinder K.
      LAL, Preeti
      KEARNEY, Liam
      WALSH, Roderick T.
      JO, Dyang Aina M.
      LO, Yan
      GREENE, Barrie D.
      RAUHMANN, Brigitte E.
      PATTERSON, Chandra

<120> TRANSPORTERS AND ION CHANNELS

<130> PI-0067 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/195,595; 60/196,873; 60/199,030; 60/200,553; 60/202,348;
      60/203,495
<151> 2000-04-06; 2000-04-12; 2000-04-20; 2000-04-28; 2000-05-05;
      2000-05-11

<160> 30
<170> PERL Program

<210> 1
<211> 520
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<233> Incyte ID No: 1784775CD1

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

```

WO 01/77174

PCT/US01/11206

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

<210> 2
 <211> 346
 <212> PPT
 <213> Homo sapiens

WO 01/77174

PCT/US01/11206

```

<220>
<221> misc_feature
<223> Incyte ID No: 7473034CD1

<400> 2
Met Val Gln Ala Ser Gly His Arg Arg Ser Thr Arg Gly Ser Lys
1      5      10      15
Met Val Ser Trp Ser Val Ile Ala Lys Ile Gln Glu Ile Trp Cys
20      25      30
Glu Glu Asp Glu Arg Lys Met Val Arg Glu Phe Leu Ala Glu Phe
35      40      45
Met Ser Thr Tyr Val Met Met Val Phe Gly Leu Gly Ser Val Ala
50      55      60
His Met Val Leu Asn Lys Thr Tyr Gly Ser Tyr Leu Gly Val Asn
65      70      75
Leu Gly Phe Gly Phe Gly Val Thr Met Gly Val His Val Ala Gly
80      85      90
Arg Ile Ser Gly Ala His Met Asn Ala Ala Val Thr Phe Thr Asn
95      100      105
Cys Ala Leu Gly Arg Val Pro Trp Arg Lys Phe Pro Val His Val
110      115      120
Leu Gly Gln Pro Leu Gly Ser Phe Leu Ala Ala Thr Ile Tyr
125      130      135
Ser Leu Phe Tyr Ser Ala Ile Leu His Phe Ser Gly Gly Glu Leu
140      145      150
Met Val Thr Gly Pro Phe Ala Thr Ala Gly Ile Phe Ala Thr Tyr
155      160      165
Leu Pro Asp His Met Thr Leu Trp Arg Gly Phe Leu Asn Glu Glu
170      175      180
Trp Leu Thr Arg Met Leu Gln Leu Cys Leu Phe Thr Ile Thr Asp
185      190      195
Gln Glu Asn Asn Pro Ala Leu Pro Gly Thr His Ala Leu Val Ile
200      205      210
Ser Ile Leu Val Val Ile Ile Arg Val Ser His Gly Ile Asn Thr
215      220      225
Gly Tyr Ala Ile Asn Pro Ser Arg Asp Pro Pro Pro Ser Ile Phe
230      235      240
Thr Phe Ile Ala Gly Trp Gly Lys Gln Val Phe Ser Asp Gly Glu
245      250      255
Asn Trp Trp Trp Val Pro Val Val Ala Pro Leu Leu Gly Ala Ser
260      265      270
Leu Gly Gly Ile Ile Tyr Leu Val Phe Ile Gly Ser Thr Ile Pro
275      280      285
Arg Glu Pro Leu Lys Leu Glu Asp Ser Val Ala Tyr Glu Asp His
290      295      300
Gly Ile Thr Val Leu Pro Lys Met Gly Ser His Glu Pro Met Ile
305      310      315
Ser Pro Leu Thr Leu Ile Ser Val Ser Leu Ala Asn Arg Ser Ser
320      325      330
Val His Ser Ala Pro Pro Leu His Glu Ser Met Ala Leu Glu His
335      340      345
Phe

```

```

<210> 3
<211> 322
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1878591CD1

```

WO 01/77174

PCT/US01/11206

```

<400> 3
Met Ser Gly Glu Leu Pro Pro Asn Ile Asn Ile Lys Glu Pro Arg
1      5      10      15
Trp Asp Gln Ser Thr Phe Ile Gly Arg Ala Asn His Phe Phe Thr
20     25     30
Val Thr Asp Pro Arg Asn Ile Leu Leu Thr Asn Glu Gln Leu Glu
35     40     45
Ser Ala Arg Lys Ile Val His Asp Tyr Arg Gln Gly Ile Val Pro
50     55     60
Pro Gly Leu Thr Glu Asn Glu Leu Trp Arg Ala Lys Tyr Ile Tyr
65     70     75
Asp Ser Ala Phe His Pro Asp Thr Gly Glu Lys Met Ile Leu Ile
80     85     90
Gly Arg Met Ser Ala Gln Val Pro Met Asn Met Thr Ile Thr Gly
95     100    105
Cys Met Met Thr Phe Tyr Arg Thr Thr Pro Ala Val Leu Phe Trp
110    115    120
Gln Trp Ile Asn Gln Ser Phe Asn Ala Val Val Asn Tyr Thr Asn
125    130    135
Arg Ser Gly Asp Ala Pro Leu Thr Val Asn Glu Leu Gly Thr Ala
140    145    150
Tyr Val Ser Ala Thr Thr Gly Ala Val Ala Thr Ala Leu Gly Leu
155    160    165
Asn Ala Leu Thr Lys His Val Ser Pro Leu Ile Gly Arg Phe Val
170    175    180
Pro Phe Ala Ala Val Ala Ala Ala Asn Cys Ile Asn Ile Pro Leu
185    190    195
Met Arg Gln Arg Glu Leu Lys Val Gly Ile Pro Val Thr Asp Glu
200    205    210
Asn Gly Asn Arg Leu Gly Glu Ser Ala Asn Ala Ala Lys Gln Ala
215    220    225
Ile Thr Gln Val Val Val Ser Arg Ile Leu Met Ala Ala Pro Gly
230    235    240
Met Ala Ile Pro Pro Phe Ile Met Asn Thr Leu Glu Lys Lys Ala
245    250    255
Phe Leu Lys Arg Phe Pro Trp Met Ser Ala Pro Ile Gln Val Gly
260    265    270
Leu Val Gly Phe Cys Leu Val Phe Ala Thr Pro Leu Cys Cys Ala
275    280    285
Leu Phe Pro Gln Lys Ser Ser Met Ser Val Thr Ser Leu Glu Ala
290    295    300
Glu Leu Gln Ala Lys Ile Gln Glu Ser His Pro Glu Leu Arg Arg
305    310    315
Val Tyr Phe Asn Lys Gly Leu
320

```

<210> 4
 <211> 723
 <212> FRT
 <213> Homo sapiens

 <220>
 <221> misc_feature
 <223> Incyte ID No: 2246992CD1

 <400> 4
 Met Arg Leu Trp Lys Ala Val Val Val Thr Leu Ala Phe Met Ser
 1 5 10 15
 Val Asp Ile Cys Val Thr Thr Ala Ile Tyr Val Phe Ser His Leu
 20 25 30
 Asp Arg Ser Leu Leu Glu Asp Ile Arg His Phe Asn Ile Phe Asp
 35 40 45
 Ser Val Leu Asp Leu Trp Ala Ala Cys Leu Tyr Arg Ser Cys Leu

WO 01/77174

PCT/US01/11206

50	55	60
Leu Leu Gly Ala Tyr Ile Gly Val Ala Lys Asn Ser Ala Leu Gly		
65	70	75
Pro Arg Arg Leu Arg Ala Ser Trp Leu Val Ile Ser Leu Val Cys		
80	85	90
Leu Phe Val Gly Ile Tyr Ala Met Val Lys Leu Leu Leu Phe Ser		
95	100	105
Glu Val Arg Arg Pro Ile Arg Asp Pro Trp Phe Trp Ala Leu Phe		
110	115	120
Val Trp Thr Tyr Ile Ser Leu Gly Ala Ser Phe Leu Leu Trp Trp		
125	130	135
Leu Leu Ser Thr Val Arg Pro Gly Thr Gln Ala Leu Glu Pro Gly		
140	145	150
Ala Ala Thr Glu Ala Glu Gly Phe Pro Gly Ser Gly Arg Pro Pro		
155	160	165
Pro Glu Gln Ala Ser Gly Ala Thr Leu Gln Lys Leu Leu Ser Tyr		
170	175	180
Thr Lys Pro Asp Val Ala Phe Leu Val Ala Ala Ser Phe Phe Leu		
185	190	195
Ile Val Ala Ala Leu Gly Glu Thr Phe Leu Pro Tyr Tyr Thr Gly		
200	205	210
Arg Ala Ile Asp Gly Ile Val Ile Gln Lys Ser Met Asp Gln Phe		
215	220	225
Ser Thr Ala Val Val Ile Val Cys Leu Leu Ala Ile Gly Ser Ser		
230	235	240
Phe Ala Ala Gly Ile Arg Gly Gly Ile Phe Thr Leu Ile Phe Ala		
245	250	255
Arg Leu Asn Ile Arg Leu Arg Asn Cys Leu Phe Arg Ser Leu Val		
260	265	270
Ser Gln Glu Thr Ser Phe Phe Asp Gly Asn Arg Thr Gly Asp Leu		
275	280	285
Ile Ser Arg Leu Thr Ser Asp Thr Thr Met Val Ser Asp Leu Val		
290	295	300
Ser Gln Asn Ile Asn Val Phe Leu Arg Asn Thr Val Lys Val Thr		
305	310	315
Gly Val Val Val Phe Met Phe Ser Leu Ser Trp Gln Leu Ser Leu		
320	325	330
Val Thr Phe Met Gly Phe Pro Ile Ile Met Met Val Ser Asn Ile		
335	340	345
Tyr Gly Lys Tyr Tyr Lys Arg Leu Ser Lys Glu Val Gln Asn Ala		
350	355	360
Leu Ala Arg Ala Ser Asn Thr Ala Glu Glu Thr Ile Ser Ala Met		
365	370	375
Lys Thr Val Arg Ser Phe Ala Asn Glu Glu Glu Glu Ala Glu Val		
380	385	390
Tyr Leu Arg Lys Leu Gln Gln Val Tyr Lys Leu Asn Arg Lys Glu		
395	400	405
Ala Ala Ala Tyr Met Tyr Tyr Val Trp Gly Ser Gly Ser Val Gly		
410	415	420
Ser Val Tyr Ser Gly Leu Met Gln Gly Val Gly Ala Ala Glu Lys		
425	430	435
Val Phe Glu Phe Ile Asp Arg Gln Pro Thr Met Val His Asp Gly		
440	445	450
Ser Leu Ala Pro Asp His Leu Glu Gly Arg Val Asp Phe Glu Asn		
455	460	465
Val Thr Phe Thr Tyr Arg Thr Arg Pro His Thr Gln Val Leu Gln		
470	475	480
Asn Val Ser Phe Ser Leu Ser Pro Gly Lys Val Thr Ala Leu Val		
485	490	495
Gly Pro Ser Gly Ser Gly Lys Ser Ser Cys Val Asn Ile Leu Glu		
500	505	510
Asn Phe Tyr Pro Leu Glu Gly Gly Arg Val Leu Leu Asp Gly Lys		
515	520	525

WO 01/77174

PCT/US01/11206

```

Pro Ile Ser Ala Tyr Asp His Lys Tyr Leu His Arg Val Ile Ser
530 535
Leu Val Ser Gln Glu Pro Val Leu Phe Ala Arg Ser Ile Thr Asp
545 550
Asn Ile Ser Tyr Gly Leu Pro Thr Val Pro Phe Glu Met Val Val
560 565
Glu Ala Ala Gln Lys Ala Asn Ala His Gly Phe Ile Met Glu Leu
575 580
Gln Asp Gly Tyr Ser Thr Glu Thr Gly Glu Lys Gly Ala Gln Leu
590 595
Ser Gly Gly Gln Lys Gln Arg Val Ala Met Ala Arg Ala Leu Val
605 610
Arg Asn Pro Pro Val Leu Ile Leu Asp Glu Ala Thr Ser Ala Leu
620 625
Asp Ala Glu Ser Glu Tyr Leu Ile Gln Gln Ala Ile His Gly Asn
635 640
Leu Gln Lys His Thr Val Leu Ile Ile Ala His Arg Leu Ser Thr
650 655
Val Glu His Ala His Leu Ile Val Val Leu Asp Lys Gly Arg Val
665 670
Val Gln Gln Gly Thr His Gln Gln Leu Leu Ala Gln Gly Gly Leu
680 685
Tyr Ala Lys Leu Val Gln Arg Gln Met Leu Gly Leu Gln Pro Ala
695 700
Ala Asp Phe Thr Ala Gly His Asn Glu Pro Val Ala Asn Gly Ser
710 715
His Lys Ala
720

```

<210> 5

<211> 1009

<212> PRT

<213> Homo sapiens

<230>

<231> misc_feature

<232> Incyte ID No: 5151730CD1

<400> 5

```

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

```

WO 01/77174

PCT/US01/11266

```

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

```

WO 01/77174

PCT/US01/11206

```

650      655      660
Met Asp Asn Pro Ile Arg Thr Phe Gln Asp Leu Ser Lys Gln Val
665      670      675
Glu Met Ser Tyr Gly Thr Val Arg Asp Ser Ala Val Tyr Glu Tyr
680      685      690
Phe Arg Ala Lys Gly Thr Asn Pro Leu Glu Gln Asp Ser Thr Phe
695      700      705
Ala Glu Leu Trp Arg Thr Ile Ser Lys Asn Gly Gly Ala Asp Asn
710      715      720
Cys Val Ser Ser Pro Ser Glu Gly Ile Arg Lys Ala Lys Lys Gly
725      730      735
Asn Tyr Ala Phe Leu Trp Asp Val Ala Val Val Glu Tyr Ala Ser
740      745      750
Leu Thr Asp Asp Asp Cys Ser Val Thr Val Ile Gly Asn Ser Ile
755      760      765
Ser Ser Lys Gly Tyr Gly Ile Ala Leu Gln His Gly Ser Pro Tyr
770      775      780
Arg Asp Leu Phe Ser Gln Arg Ile Leu Glu Leu Gln Asp Thr Gly
785      790      795
Asp Leu Asp Val Leu Lys Gln Lys Trp Trp Pro His Met Gly Arg
800      805      810
Cys Asp Leu Thr Ser His Ala Ser Ala Gln Ala Asp Gly Lys Ser
815      820      825
Leu Lys Leu His Ser Phe Ala Gly Val Phe Cys Ile Leu Ala Ile
830      835      840
Gly Leu Leu Leu Ala Cys Leu Val Ala Ala Leu Glu Leu Trp Trp
845      850      855
Asn Ser Asn Arg Cys His Gln Glu Thr Pro Lys Glu Asp Lys Glu
860      865      870
Val Asn Leu Glu Gln Val His Arg Arg Met Asn Ser Leu Met Asp
875      880      885
Glu Asp Ile Ala His Lys Gln Ile Ser Pro Ala Ser Ile Glu Leu
890      895      900
Ser Ala Leu Glu Met Gly Gly Leu Ala Pro Thr Gln Thr Leu Glu
905      910      915
Pro Thr Arg Glu Tyr Gln Asn Thr Gln Leu Ser Val Ser Thr Phe
920      925      930
Leu Pro Glu Gln Ser Ser His Gly Thr Ser Arg Thr Leu Ser Ser
935      940      945
Gly Pro Ser Ser Asn Leu Pro Leu Pro Leu Ser Ser Ser Ala Thr
950      955      960
Met Pro Ser Met Gln Cys Lys His Arg Ser Pro Asn Gly Gly Leu
965      970      975
Phe Arg Gln Ser Pro Val Lys Thr Pro Ile Pro Met Ser Phe Gln
980      985      990
Pro Val Pro Gly Gly Val Leu Pro Glu Ala Leu Asp Thr Ser His
995      1000      1005
Gly Thr Ser Ile

```

```

<210> 6
<211> 374
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<233> Inverte ID No: 7472584CD1

```

```

<400> 6
Met Lys Arg Gln Asn Val Arg Thr Leu Ser Leu Ile Val Cys Thr
1      5      10      15
Phe Thr Tyr Leu Leu Val Gly Ala Ala Val Phe Asp Ala Leu Glu

```


WO 01/77174

PCT/US01/11206

```

      20      25      30
Ser Asp His Glu Met Arg Glu Glu Glu Lys Leu Lys Ala Glu Glu
      35      40      45
Ile Arg Ile Lys Gly Lys Tyr Asn Ile Ser Ser Glu Asp Tyr Arg
      50      55      60
Gln Leu Glu Leu Val Ile Leu Gln Ser Glu Pro His Arg Ala Gly
      65      70      75
Val Gln Trp Lys Phe Ala Gly Ser Phe Tyr Phe Ala Ile Thr Val
      80      85      90
Ile Thr Thr Ile Gly Tyr Gly His Ala Ala Pro Gly Thr Asp Ala
      95     100     105
Gly Lys Ala Phe Cys Met Phe Tyr Ala Val Leu Gly Ile Pro Leu
     110     115     120
Thr Leu Val Met Phe Gln Ser Leu Gly Glu Arg Met Asn Thr Phe
     125     130     135
Val Arg Tyr Leu Leu Lys Arg Ile Lys Lys Cys Cys Gly Met Arg
     140     145     150
Asn Thr Asp Val Ser Met Glu Asn Met Val Thr Val Gly Phe Phe
     155     160     165
Ser Cys Met Gly Thr Leu Cys Ile Gly Ala Ala Ala Phe Ser Gln
     170     175     180
Cys Glu Glu Trp Ser Phe Phe His Ala Tyr Tyr Cys Phe Ile
     185     190     195
Thr Leu Thr Thr Ile Gly Phe Gly Asp Tyr Val Ala Leu Gln Thr
     200     205     210
Lys Gly Ala Leu Gln Lys Lys Pro Leu Tyr Val Ala Phe Ser Phe
     215     220     225
Met Tyr Ile Leu Val Gly Leu Thr Val Ile Gly Ala Phe Leu Asn
     230     235     240
Leu Val Val Leu Arg Phe Leu Thr Met Asn Ser Glu Asp Glu Arg
     245     250     255
Arg Asp Ala Glu Glu Arg Ala Ser Leu Ala Gly Asn Arg Asn Ser
     260     265     270
Met Val Ile His Ile Pro Glu Glu Pro Arg Pro Ser Arg Pro Arg
     275     280     285
Tyr Lys Ala Asp Val Pro Asp Leu Gln Ser Val Cys Ser Cys Thr
     290     295     300
Cys Tyr Arg Ser Gln Asp Tyr Gly Gly Arg Ser Val Ala Pro Gln
     305     310     315
Asn Ser Phe Ser Ala Lys Leu Ala Pro His Tyr Phe His Ser Ile
     320     325     330
Ser Tyr Lys Ile Glu Glu Ile Ser Pro Ser Thr Leu Lys Asn Ser
     335     340     345
Leu Phe Pro Ser Pro Ile Ser Ser Ile Ser Pro Gly Leu His Ser
     350     355     360
Phe Thr Asp His Gln Arg Leu Met Lys Arg Arg Lys Ser Val
     365     370

```

<210> 7
 <211> 589
 <212> PRT
 <213> Homo sapiens

 <210>
 <211> misc_feature
 <223> Incyte ID No: 7472536CD1

 <400> 7
 Met Ser Ala Val Leu Thr Pro Gly Leu Phe Leu Pro Leu Pro Gly
 1 5 10 15
 Pro Leu Pro Ala Ser Leu His Lys Ala Gly Gly Thr Gly Pro Gln
 20 25 30
 Val Arg Pro Met Ala Met Ala Phe Thr Asp Leu Leu Asp Ala Leu

WO 01/77174

PCT/US01/11206

Gly	Ser	Met	Gly	35	Phe	Gln	Leu	Asn	His	Thr	Ala	Leu	Leu	Leu	45
			Arg	50					55						60
Leu	Pro	Cys	Gly	65	Leu	Leu	Ala	Cys	His	Asn	Phe	Leu	Gln	Asn	Phe
				70					75						80
Thr	Ala	Ala	Val	85	Pro	Pro	His	His	Cys	Arg	Gly	Pro	Ala	Asn	His
				90					95						100
Thr	Glu	Ala	Ser	105	Thr	Asn	Asp	Ser	Gly	Ala	Trp	Leu	Arg	Ala	Thr
				110					115						120
Ile	Pro	Leu	Asp	125	Gln	Leu	Gly	Ala	Pro	Glu	Pro	Cys	Arg	Arg	Phe
				130					135						140
Thr	Iys	Pro	Gln	145	Trp	Ala	Leu	Leu	Ser	Pro	Asn	Ser	Ser	Ile	Pro
				150					155						160
Gly	Ala	Ala	Thr	165	Glu	Gly	Cys	Lys	Asp	Gly	Trp	Val	Tyr	Asn	Arg
				170					175						180
Ser	Val	Phe	Pro	185	Ser	Thr	Ile	Val	Met	Glu	Gln	Trp	Asp	Leu	Val
				190					195						200
Cys	Glu	Ala	Arg	205	Thr	Leu	Arg	Asp	Leu	Ala	Gln	Ser	Val	Tyr	Ile
				210					215						220
Ala	Gly	Val	Leu	225	Val	Gly	Ala	Ala	Val	Phe	Gly	Ser	Leu	Ala	Asp
				230					235						240
Arg	Leu	Gly	Cys	245	Lys	Gly	Pro	Leu	Val	Trp	Ser	Tyr	Leu	Gln	Leu
				250					255						260
Ala	Ala	Ser	Gly	265	Ala	Ala	Thr	Ala	Tyr	Phe	Ser	Ser	Phe	Ser	Ala
				270					275						280
Tyr	Cys	Val	Phe	285	Arg	Phe	Leu	Met	Gly	Met	Thr	Phe	Ser	Gly	Ile
				290					295						300
Ile	Leu	Asn	Ser	305	Val	Ser	Leu	Val	Glu	Trp	Met	Pro	Thr	Arg	Gly
				310					315						320
Arg	Thr	Val	Ala	325	Gly	Ile	Leu	Leu	Gly	Tyr	Ser	Phe	Thr	Leu	Gly
				330					335						340
Gln	Leu	Ile	Leu	345	Ala	Gly	Val	Ala	Tyr	Leu	Ile	Arg	Pro	Trp	Arg
				350					355						360
Cys	Leu	Gln	Phe	365	Ala	Ile	Ser	Ala	Pro	Phe	Leu	Ile	Phe	Phe	Leu
				370					375						380
Tyr	Ser	Trp	Trp	385	Leu	Pro	Glu	Ser	Ser	Arg	Trp	Leu	Leu	Leu	His
				390					395						400
Gly	Lys	Ser	Gln	405	Leu	Ala	Val	Gln	Asn	Leu	Gln	Lys	Val	Ala	Ala
				410					415						420
Met	Asn	Gly	Arg	425	Lys	Gln	Glu	Gly	Glu	Arg	Leu	Thr	Lys	Glu	Val
				430					435						440
Met	Ser	Ser	Tyr	445	Ile	Gln	Ser	Glu	Phe	Ala	Ser	Val	Cys	Thr	Ser
				450					455						460
Asn	Ser	Ile	Leu	465	Asp	Leu	Phe	Arg	Thr	Pro	Ala	Ile	Arg	Lys	Val
				470					475						480
Thr	Cys	Cys	Pro	485	Ala	Leu	Arg	Phe	Ser	Asn	Ser	Val	Ala	Tyr	Tyr
				490					495						500
Gly	Leu	Ala	Met	505	Asp	Leu	Gln	Lys	Phe	Gly	Leu	Ser	Leu	Tyr	Leu
				510					515						520
Val	Gln	Ala	Leu	525	Phe	Gly	Ile	Ile	Asn	Ile	Pro	Ala	Met	Leu	Val
				530					535						540
Ala	Thr	Ala	Thr	545	Met	Ile	Tyr	Val	Gly	Arg	Arg	Ala	Thr	Val	Ala
				550					555						560
Ser	Phe	Leu	Ile	565	Leu	Ala	Gly	Leu	Met	Val	Ile	Ala	Asn	Met	Phe
				570					575						580
Val	Pro	Glu	Gly	585	Thr	Gln	Ile	Leu	Cys	Thr	Ala	Gln	Ala	Ala	Leu
				590					595						600
Gly	Lys	Gly	Cys	605	Leu	Ala	Ser	Ser	Phe	Ile	Cys	Val	Tyr	Leu	Phe
				610					615						620
Thr	Gly	Glu	Leu	625	Tyr	Pro	Thr	Gln	Ile	Arg	Gln	Met	Gly	Met	Gly
				630					635						640
Phe	Ala	Ser	Val	645	His	Ala	Arg	Leu	Gly	Gly	Leu	Thr	Ala	Pro	Leu
				650					655						660

WO 01/77174

PCT/US01/11206

```

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

```

<210> 8

<211> 549

<212> FRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> incyte ID No: 7473422C01

<400> 8

```

Met Val Ser Asp Arg Gly Leu Lys Pro Phe Glu Asp Leu Arg Pro
1 5 10 15
Pro Lys Ile Ser Pro Ala Asp Leu Gly Asn Ala Glu Glu Ala Ile
20 25 30
Glu Leu Glu Arg Glu Pro Ala Pro Val Arg Phe Val Pro Arg Arg
35 40 45
Lys Arg Pro Ser Ile Trp Val Val Leu Pro Val Leu Phe Leu Val
50 55 60
Ala Met Ser Leu Leu Pro Leu Leu Tyr Val Ala Ile Lys Ala Trp
65 70 75
Glu Ala Gly Trp Arg Glu Ala Leu His Leu Leu Trp Arg Pro Phe
80 85 90
Val Trp Gly Leu Met Arg Asn Thr Leu Met Leu Met Val Gly Val
95 100 105
Thr Leu Ala Cys Met Val Val Gly Leu Ala Leu Ala Trp Leu Leu
110 115 120
Glu Arg Ser Asn Leu Ala Gly Arg Arg Leu Trp Gly Val Val Leu
125 130 135
Cys Leu Pro Phe Ala Val Pro Ser Phe Val Ser Ser Phe Thr Trp
140 145 150
Val Ser Leu Ser Ser Asp Phe Glu Gly Leu Gly Gly Ala Ile Leu
155 160 165
Val Met Ala Leu Ser Lys Tyr Pro Leu Val Phe Leu Pro Val Ala
170 175 180
Ala Thr Leu Arg Asn Leu Asp Thr Ser Leu Glu Glu Ser Ala Arg
185 190 195
Thr Leu Gly Cys Ser Arg Trp Gly Val Phe Ile Lys Val Thr Leu
200 205 210
Pro Leu Leu Trp Pro Ser Met Leu Gly Gly Ala Leu Leu Ile Ala
215 220 225
Leu His Met Leu Val Glu Phe Gly Ala Leu Ser Ile Leu Gly Leu
230 235 240
Gln Thr Phe Thr Thr Ala Ile Tyr Gln Gln Phe Glu Leu Glu Phe
245 250 255
Ser Asn Ala Asn Ala Ala Met Leu Ser Ala Val Leu Leu Ala Met
260 265 270
Cys Leu Val Met Leu Trp Leu Glu Leu Arg Val Arg Gly Lys Ala
275 280 285
Arg His Val Arg Ile Gly Gln Gly Val Ala Arg Arg Ala Gln Pro
290 295 300

```

WO 01/77174

PCT/US01/11206

```

Val Arg Leu Arg Gly Trp Ala Val Pro Ala Gln Leu Leu Cys Val
305 310 315
Ala Leu Ala Val Leu Gly Ser Gly Ile Pro Leu Ala Met Leu Gly
320 325 330
Tyr Trp Leu Ser Val Gly Ser Ser Ala Ala Phe Pro Val Gly Ala
335 340 345
Ile Ser Lys Ala Leu Phe Thr Ser Leu Ser Val Ser Leu Gly Gly
350 355 360
Ala Gly Phe Cys Val Leu Leu Ala Leu Pro Ile Ser Phe Leu Val
365 370 375
Val Arg Tyr Lys Gly Arg Leu Ala Ile Trp Ala Glu Arg Leu Pro
380 385 390
Tyr Leu Leu His Ala Leu Pro Gly Leu Val Ile Ala Leu Thr Leu
395 400 405
Val Phe Phe Ala Leu His Tyr Val Pro Ala Leu Tyr Gln Thr Thr
410 415 420
Ala Leu Leu Leu Leu Ala Tyr Ala Leu Leu Phe Leu Pro Leu Ala
425 430 435
Gln Ser Pro Val Arg Thr Ala Leu Asn Lys Ala Ser Pro Thr Leu
440 445 450
Glu Glu Ala Ala Arg Thr Leu Gly Ala Ser Ser Phe Thr Ala Phe
455 460 465
Cys Arg Val Thr Leu Pro Ile Ile Phe Pro Ala Met Ala Ala Ala
470 475 480
Phe Ala Leu Val Phe Leu Arg Ala Met Lys Glu Leu Thr Ala Thr
485 490 495
Leu Leu Leu Ser Pro Thr Gly Met Thr Thr Leu Ala Thr Glu Val
500 505 510
Trp Ala His Thr Ala Asn Val Glu Phe Ala Ala Ala Ala Pro Tyr
515 520 525
Ala Ala Leu Leu Ile Val Val Ser Gly Leu Pro Val Tyr Leu Leu
530 535 540
Thr Thr Arg Met Tyr Leu Asn Lys Ala
545

```

<210> 9

<211> 634

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2864715CD1

<400> 9

```

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

```

WO 01/77174

PCT/US01/11206

Ile Met Trp Tyr Leu Phe Asn Ser Phe Gln Glu Pro Leu Pro Trp
 140 145 150
 Ser Asp Trp Pro Leu Asn Glu Asn Gln Thr Gly Tyr Val Asp Glu
 155 160 165
 Cys Ala Arg Ser Ser Pro Val Asp Tyr Phe Trp Tyr Arg Glu Thr
 170 175 180
 Leu Asn Ile Ser Thr Ser Ile Ser Asp Ser Gly Ser Ile Gln Trp
 185 190 195
 Trp Met Leu Leu Cys Leu Ala Cys Ala Trp Ser Val Leu Tyr Met
 200 205 210
 Cys Thr Ile Arg Gly Ile Glu Thr Thr Gly Lys Ala Val Tyr Ile
 215 220 225
 Thr Ser Thr Leu Pro Tyr Val Val Leu Thr Ile Phe Leu Ile Arg
 230 235 240
 Gly Leu Thr Leu Lys Gly Ala Thr Asn Gly Ile Val Phe Leu Phe
 245 250 255
 Thr Pro Asn Val Thr Glu Leu Ala Gln Pro Asp Thr Trp Leu Asp
 260 265 270
 Ala Gly Ala Gln Val Phe Phe Ser Phe Ser Leu Ala Phe Gly Gly
 275 280 285
 Leu Ile Ser Phe Ser Ser Tyr Asn Ser Val His Asn Asn Cys Glu
 290 295 300
 Lys Asp Ser Val Ile Val Ser Ile Ile Asn Gly Phe Thr Ser Val
 305 310 315
 Tyr Val Ala Ile Val Val Tyr Ser Val Ile Gly Phe Arg Ala Thr
 320 325 330
 Gln Arg Tyr Asp Asp Cys Phe Ser Thr Asn Ile Leu Thr Leu Ile
 335 340 345
 Asn Gly Phe Asp Leu Pro Glu Gly Asn Val Thr Gln Glu Asn Phe
 350 355 360
 Val Asp Met Gln Gln Arg Cys Asn Ala Ser Asp Pro Ala Ala Tyr
 365 370 375
 Ala Gln Leu Val Phe Gln Thr Cys Asp Ile Asn Ala Phe Leu Ser
 380 385 390
 Glu Ala Val Glu Gly Thr Gly Leu Ala Phe Ile Val Phe Thr Glu
 395 400 405
 Ala Ile Thr Lys Met Pro Leu Ser Pro Leu Trp Ser Val Leu Phe
 410 415 420
 Phe Ile Met Leu Phe Cys Leu Gly Leu Ser Ser Met Phe Gly Asn
 425 430 435
 Met Glu Gly Val Val Val Pro Leu Gln Asp Leu Arg Val Ile Pro
 440 445 450
 Pro Lys Trp Pro Lys Glu Val Leu Thr Gly Leu Ile Cys Leu Gly
 455 460 465
 Thr Phe Leu Ile Gly Phe Ile Phe Thr Leu Asn Ser Gly Gln Tyr
 470 475 480
 Trp Leu Ser Leu Leu Asp Ser Tyr Ala Gly Ser Ile Pro Leu Leu
 485 490 495
 Ile Ile Ala Phe Cys Glu Met Phe Ser Val Val Tyr Val Tyr Gly
 500 505 510
 Val Asp Arg Phe Asn Lys Asp Ile Glu Phe Met Ile Gly His Lys
 515 520 525
 Pro Asn Ile Phe Trp Gln Val Thr Trp Arg Val Val Ser Pro Leu
 530 535 540
 Leu Met Leu Ile Ile Phe Leu Phe Phe Phe Val Val Glu Val Ser
 545 550 555
 Gln Glu Leu Thr Tyr Ser Ile Trp Asp Pro Gly Tyr Glu Glu Phe
 560 565 570
 Pro Lys Ser Gln Lys Ile Ser Tyr Pro Asn Trp Val Tyr Val Val
 575 580 585
 Val Val Ile Val Ala Gly Val Pro Ser Leu Thr Ile Pro Gly Cys
 590 595 600
 Ala Ile Tyr Lys Leu Ile Arg Asn His Cys Gln Lys Pro Gly Asp

WO 01/77174

PCT/US01/11206

His Gln Gly Leu 605 610 615
Val Ser Thr Leu Ser Thr Ala Ser Met Asn Gly
Asp Leu Lys Tyr 620 625 630

<210> 10

<211> 491

<212> 2RT

<213> Homo sapiens

<320>

<321> misc_feature

<323> Incyte ID No: 1734724CDL

<400> 10

Met Asp Gly Asn Asp Asn Val Thr Leu Leu Phe Ala Pro Leu Leu
1 5 10 15
Arg Asp Asn Tyr Thr Leu Ala Pro Asn Ala Ser Ser Leu Gly Pro
20 25 30
Gly Thr Asn Leu Ala Leu Ala Pro Ala Ser Ser Ala Gly Pro Gly
35 40 45
Pro Gly Leu Ser Leu Gly Pro Val Pro Ser Phe Gly Phe Ser Pro
50 55 60
Gly Pro Thr Pro Thr Pro Glu Pro Thr Thr Ser Gly Leu Ala Gly
65 70 75
Gly Ala Ala Ser His Gly Pro Ser Pro Val Pro Ser Ala Leu Gly
80 85 90
Ala Pro Arg Ala Pro Val Leu Gly His Ala Ala Glu Pro Arg Ala
95 100 105
Glu Arg Val Arg Gly Arg Arg Leu Cys Ile Thr Met Leu Gly Leu
110 115 120
Gly Cys Thr Val Asp Val Asn His Phe Gly Ala His Val Arg Arg
125 130 135
Pro Val Ala Ala Leu Leu Ala Ala Leu Pro Val Arg Pro Pro Ala
140 145 150
Ala Ala Gly Leu Pro Ala Gly Pro Arg Leu Gln Ala Gly Arg Gly
155 160 165
Gly Arg Arg Gly Leu Leu Leu Cys Gly Cys Cys Pro Gly Gly Asn
170 175 180
Leu Ser Asn Leu Met Ser Leu Leu Val Asp Gly Asp Met Asn Leu
185 190 195
Arg Arg Ala Ala Leu Leu Ala Leu Ser Ser Asp Val Gly Ser Ala
200 205 210
Gln Thr Ser Thr Pro Gly Leu Ala Val Ser Pro Phe His Leu Tyr
215 220 225
Ser Thr Tyr Lys Lys Lys Val Ser Trp Leu Phe Asp Ser Lys Leu
230 235 240
Val Leu Ile Ser Ala His Ser Leu Phe Cys Ser Ile Ile Met Thr
245 250 255
Ile Ser Ser Thr Leu Leu Ala Leu Val Leu Met Pro Leu Cys Leu
260 265 270
Trp Ile Tyr Ser Trp Ala Trp Ile Asn Thr Pro Ile Val Gln Leu
275 280 285
Leu Pro Leu Gly Thr Val Thr Leu Thr Leu Cys Ser Thr Leu Ile
290 295 300
Pro Ile Gly Leu Gly Val Phe Ile Arg Tyr Lys Tyr Ser Arg Val
305 310 315
Ala Asp Tyr Ile Val Lys Val Ser Leu Trp Ser Leu Leu Val Thr
320 325 330
Leu Val Val Leu Phe Ile Met Thr Gly Thr Met Leu Gly Pro Glu
335 340 345
Leu Leu Ala Ser Ile Pro Ala Ala Val Tyr Val Ile Ala Ile Phe

WO 01/77174

PCT/US01/11206

```

350      355      360
Met Pro Leu Ala Gly Tyr Ala Ser Gly Tyr Gly Leu Ala Thr Leu
365      370      375
Phe His Leu Pro Pro Asn Cys Lys Arg Thr Val Cys Leu Glu Thr
380      385      390
Gly Ser Gln Asn Val Gln Leu Cys Thr Ala Ile Leu Lys Leu Ala
395      400      405
Phe Pro Pro Gln Phe Ile Gly Ser Met Tyr Met Phe Pro Leu Leu
410      415      420
Tyr Ala Leu Phe Gln Ser Ala Glu Ala Gly Ile Phe Val Leu Ile
425      430      435
Tyr Lys Met Tyr Gly Ser Glu Met Leu His Lys Arg Asp Pro Leu
440      445      450
Asp Glu Asp Glu Asp Thr Asp Ile Ser Tyr Lys Lys Leu Lys Glu
455      460      465
Glu Glu Met Ala Asp Thr Ser Tyr Gly Thr Val Lys Ala Glu Asp
470      475      480
Ile Ile Met Met Glu Thr Ala Gln Thr Ser Leu
485      490

```

<210> 11

<211> 525

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1563237CD1

<400> 11

```

Met Pro Ala Pro Arg Ala Arg Glu Gln Pro Arg Val Pro Gly Glu
1      5      10      15
Arg Gln Pro Leu Leu Pro Arg Gly Ala Arg Gly Pro Arg Arg Trp
20      25      30
Arg Arg Ala Ala Gly Ala Ala Val Leu Leu Val Glu Met Leu Glu
35      40      45
Arg Ala Ala Phe Phe Gly Val Thr Ala Asn Leu Val Leu Tyr Leu
50      55      60
Asn Ser Thr Asn Phe Asn Trp Thr Gly Glu Gln Ala Thr Arg Ala
65      70      75
Ala Leu Val Phe Leu Gly Ala Ser Tyr Leu Leu Ala Pro Val Gly
80      85      90
Gly Trp Leu Ala Asp Val Tyr Leu Gly Arg Tyr Arg Ala Val Ala
95      100      105
Leu Ser Leu Leu Leu Tyr Leu Ala Ala Ser Gly Leu Leu Pro Ala
110      115      120
Thr Ala Phe Pro Asp Gly Arg Ser Ser Phe Cys Gly Glu Met Pro
125      130      135
Ala Ser Pro Leu Gly Pro Ala Cys Pro Ser Ala Gly Cys Pro Arg
140      145      150
Ser Ser Pro Ser Pro Tyr Cys Ala Pro Val Leu Tyr Ala Gly Leu
155      160      165
Leu Leu Leu Gly Leu Ala Ala Ser Ser Val Arg Ser Asn Leu Thr
170      175      180
Ser Phe Gly Ala Asp Gln Val Met Asp Leu Gly Arg Asp Ala Thr
185      190      195
Arg Arg Phe Phe Asn Trp Phe Tyr Trp Ser Ile Asn Leu Gly Ala
200      205      210
Val Leu Ser Leu Leu Val Val Ala Phe Ile Gln Gln Asn Ile Ser
215      220      225
Phe Leu Leu Gly Tyr Ser Ile Pro Val Gly Cys Val Gly Leu Ala
230      235      240
Phe Phe Ile Phe Leu Phe Ala Thr Pro Val Phe Ile Thr Lys Pro

```

WO 01/77174

PCT/US01/11206

```

245      250      255
Pro Met Gly Ser Gln Val Ser Ser Met Leu Lys Leu Ala Leu Gln
260      265      270
Asn Cys Cys Pro Gln Leu Trp Gln Arg His Ser Ala Arg Asp Arg
275      280      285
Gln Cys Ala Arg Val Leu Ala Asp Glu Arg Ser Pro Gln Pro Gly
290      295      300
Ala Ser Pro Gln Glu Asp Ile Ala Asn Phe Gln Val Leu Val Lys
305      310      315
Ile Leu Pro Val Met Val Thr Leu Val Pro Tyr Trp Met Val Tyr
320      325      330
Phe Gln Met Gln Ser Thr Tyr Val Leu Gln Gly Leu His Leu His
335      340      345
Ile Pro Asn Ile Phe Pro Ala Asn Pro Ala Asn Ile Ser Val Ala
350      355      360
Leu Arg Ala Gln Gly Ser Ser Tyr Thr Glu Ser Trp Arg Trp Ser
365      370      375
Ala Leu His Tyr Ile His His Asn Glu Thr Val Ser Gln Gln Ile
380      385      390
Gly Glu Val Leu Tyr Asn Ala Ala Pro Leu Ser Ile Trp Trp Gln
395      400      405
Ile Pro Gln Tyr Leu Leu Ile Gly Ile Ser Glu Ile Phe Ala Ser
410      415      420
Ile Pro Gly Leu Glu Phe Ala Tyr Ser Glu Ala Pro Arg Ser Met
425      430      435
Gln Gly Ala Ile Met Gly Ile Phe Phe Cys Leu Ser Gly Val Gly
440      445      450
Ser Leu Leu Gly Ser Ser Leu Val Ala Leu Leu Ser Leu Pro Gly
455      460      465
Gly Trp Leu His Cys Pro Lys Asp Phe Gly Asn Ile Asn Asn Cys
470      475      480
Arg Met Asp Leu Tyr Phe Phe Leu Leu Ala Gly Ile Gln Ala Val
485      490      495
Thr Ala Leu Leu Phe Val Trp Ile Ala Gly Arg Tyr Glu Arg Ala
500      505      510
Ser Gln Gly Pro Ala Ser His Ser Arg Phe Ser Arg Asp Arg Gly
515      520      525

```

<210> 12

<211> 1310

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7473443CD1

<400> 12

```

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

```


WO 01/77174

PCT/US01/11206

Ser Ile Ser Ser Arg Ile Asp Gln Ala Glu Lys Arg Ile Ser Glu
 110 115 120
 Ile Glu Asp Gln Leu Asn Glu Ile Lys Arg Glu Asp Lys Ile Arg
 125 130 135
 Glu Lys Asn Glu Lys Asp Glu Gln Gly Leu Gln Glu Ile Trp Asp
 140 145 150
 Tyr Val Lys Arg Pro Asn Leu His Leu Ile Gly Val Pro Gly Leu
 155 160 165
 Leu Tyr Ser Asp Met Cys Arg Leu Leu Pro Ser Pro Arg Asn Gln
 170 175 180
 Pro Ala Leu Gln Ala Leu Glu Arg Gly Val Ile Leu Glu Val Lys
 185 190 195
 Cys Val Val Cys Ser Thr Gln Ala Gly Ala Ala Arg Arg Gly Val
 200 205 210
 Lys Ile Ser Ile Lys Gly Lys Gly Phe Ser Val Val Ser Val Val
 215 220 225
 Gly Thr Leu Gln Trp Leu Leu Trp Ala Arg Ala Ala His Ala Pro
 230 235 240
 His Trp Arg Phe Leu Arg Trp Met Ala Ala Leu Trp Asp Val Pro
 245 250 255
 Gly Lys Thr Gly Pro Ser Pro Ile Ser Leu Thr Gly Gln Arg Gly
 260 265 270
 Asn Arg Gly Pro Glu Ser Ser Ser Ile Leu Arg Gly Val Pro Lys
 275 280 285
 Asp Phe Ser Thr Gly Thr Ser Ala Gln Leu Arg Arg Ala Met Gly
 290 295 300
 Leu Ala Pro Glu Gly Gly Gly Phe Gln Ala Phe Phe Pro Arg Pro
 305 310 315
 Thr Met Pro Ala Thr Pro Asn Phe Leu Ala Asn Pro Ser Ser Ser
 320 325 330
 Ser Arg Trp Ile Pro Leu Gln Pro Met Pro Val Ala Trp Ala Phe
 335 340 345
 Val Gln Lys Thr Ser Ala Leu Leu Trp Leu Leu Leu Gly Thr
 350 355 360
 Ser Leu Ser Pro Ala Trp Gly Gln Ala Lys Ile Pro Leu Glu Thr
 365 370 375
 Val Lys Leu Trp Ala Asp Thr Phe Gly Gly Asp Leu Tyr Asn Thr
 380 385 390
 Val Thr Lys Tyr Ser Gly Ser Leu Leu Leu Gln Lys Lys Tyr Lys
 395 400 405
 Asp Val Glu Ser Ser Leu Lys Ile Glu Glu Val Asp Gly Leu Glu
 410 415 420
 Leu Val Arg Lys Phe Ser Glu Asp Met Glu Asn Met Leu Arg Arg
 425 430 435
 Lys Val Glu Ala Val Gln Asn Leu Val Glu Ala Ala Glu Glu Ala
 440 445 450
 Asp Leu Asn His Glu Phe Asn Glu Ser Leu Val Glu Pro Gly Val
 455 460 465
 Gly Val Gly Val Gly Met Ser Val Thr Gln Ser Gly Val Gly Val
 470 475 480
 Gly Val Gly Met Ser Val Thr Gln Ser Gly Val Gly Val Gly Val
 485 490 495
 Gly Met Ser Ile Thr Leu Ser Gly Val Gly Val Gly Val Gly Met
 500 505 510
 Ser Val Arg Gln Ser Gly Val Gly Val Gly Val Gly Met Ser Val
 515 520 525
 Thr Gln Ser Gly Val Gly Val Gly Val Gly Met Ser Val Thr Gln
 530 535 540
 Ser Gly Val Gly Val Gly Val Gly Met Ser Val Arg Gln Ser Gly
 545 550 555
 Val Gly Val Gly Val Gly Met Ser Val Thr Gln Ser Trp Gly Val
 560 565 570
 Phe Ser Ala Gln Arg Ala Ala Ala Gly Ala Cys Val Asp Ser Asp

WO 01/77174

PCT/US01/11206

Gly Arg Pro Ala	575	580	585
Pro Ala Leu Ser Ser Ser His Leu Arg Arg	590	595	600
Ser Ser Ser Leu Ser Ala Cys Pro Gly Ala Arg Ala Ala Ser	605	610	615
Gly Leu Thr Arg Pro Pro Gln Phe Asp Tyr Tyr Asn Ser Val Leu	620	625	630
Ile Asn Glu Arg Asp Glu Lys Gly Asn Phe Val Glu Leu Gly Ala	635	640	645
Glu Phe Leu Leu Glu Ser Asn Ala His Phe Ser Asn Leu Pro Val	650	655	660
Asn Thr Ser Ile Ser Ser Val Gln Leu Pro Thr Asn Val Tyr Asn	665	670	675
Lys Asp Pro Asp Ile Leu Asn Gly Val Tyr Met Ser Glu Ala Leu	680	685	690
Asn Ala Val Phe Val Glu Asn Phe Gln Arg Asp Pro Thr Leu Thr	695	700	705
Trp Gln Cys Phe Gly Ser Ala Thr Gly Phe Phe Arg Ile Tyr Pro	710	715	720
Gly Ile Lys Trp Thr Pro Asp Glu Asn Gly Val Ile Thr Phe Asp	725	730	735
Cys Arg Asn Arg Gly Trp Tyr Ile Gln Ala Ala Thr Ser Pro Lys	740	745	750
Asp Ile Val Ile Leu Val Asp Val Ser Gly Ser Met Lys Gly Leu	755	760	765
Arg Met Thr Ile Ala Lys His Thr Ile Thr Thr Ile Leu Asp Thr	770	775	780
Leu Gly Glu Asn Asp Phe Ile Asn Ile Ile Ala Tyr Asn Asp Tyr	785	790	795
Val His Tyr Ile Glu Pro Cys Phe Lys Gly Ile Leu Val Gln Ala	800	805	810
Asp Arg Asp Asn Arg Glu His Phe Lys Leu Leu Val Glu Glu Leu	815	820	825
Met Val Lys Gly Val Gly Val Val Asp Gln Ala Leu Arg Glu Ala	830	835	840
Phe Gln Ile Leu Lys Gln Phe Gln Glu Ala Lys Gln Gly Ser Leu	845	850	855
Cys Asn Gln Ala Ile Met Leu Ile Ser Asp Gly Ala Val Glu Asp	860	865	870
Tyr Glu Pro Val Phe Glu Lys Tyr Asn Trp Pro Asp Cys Lys Val	875	880	885
Arg Val Phe Thr Tyr Leu Ile Gly Arg Glu Val Ser Phe Ala Asp	890	895	900
Arg Met Lys Trp Ile Ala Cys Asn Asn Lys Gly Tyr Tyr Thr Gln	905	910	915
Ile Ser Thr Leu Ala Asp Thr Gln Glu Asn Val Met Glu Tyr Leu	920	925	930
His Val Leu Ser Arg Pro Met Val Ile Asn His Asp His Asp Ile	935	940	945
Ile Trp Thr Glu Ala Tyr Met Asp Ser Lys Leu Leu Ser Ser Gln	950	955	960
Ala Gln Ser Leu Thr Leu Leu Thr Thr Val Ala Met Pro Val Phe	965	970	975
Ser Lys Lys Asn Glu Thr Arg Ser His Gly Ile Leu Leu Gly Val	980	985	990
Val Gly Ser Asp Val Ala Leu Arg Glu Leu Met Lys Leu Ala Pro	995	1000	1005
Arg Tyr Lys Leu Gly Val His Gly Tyr Ala Phe Leu Asn Thr Asn	1010	1015	1020
Asn Gly Tyr Ile Leu Ser His Pro Asp Leu Arg Pro Leu Tyr Arg	1025	1030	1035
Glu Gly Lys Lys Leu Lys Pro Lys Pro Asn Tyr Asn Ser Val Asp	1040	1045	1050

WO 01/77174

PCT/US01/11266

Leu Ser Glu Val Glu Trp Glu Asp Gln Ala Glu Ser Leu Arg Thr
 1055 1060 1065
 Ala Met Ile Asn Arg Glu Thr Gly Thr Leu Ser Met Asp Val Lys
 1070 1075 1080
 Val Pro Met Asp Lys Gly Lys Arg Val Leu Phe Leu Thr Asn Asp
 1085 1090 1095
 Tyr Phe Phe Thr Asp Ile Ser Asp Thr Pro Phe Ser Leu Gly Val
 1100 1105 1110
 Val Leu Ser Arg Gly His Gly Glu Tyr Ile Leu Leu Gly Asn Thr
 1115 1120 1125
 Ser Val Glu Glu Gly Leu His Asp Leu Leu His Pro Asp Leu Ala
 1130 1135 1140
 Leu Ala Gly Asp Trp Ile Tyr Cys Ile Thr Asp Ile Asp Pro Asp
 1145 1150 1155
 His Arg Lys Leu Ser Gln Leu Glu Ala Met Ile Arg Phe Leu Thr
 1160 1165 1170
 Arg Lys Asp Pro Asp Leu Glu Cys Asp Glu Glu Leu Val Arg Glu
 1175 1180 1185
 Val Leu Phe Asp Ala Val Val Thr Ala Pro Met Glu Ala Tyr Trp
 1190 1195 1200
 Thr Ala Leu Ala Leu Asn Met Ser Glu Glu Ser Glu His Val Val
 1205 1210 1215
 Asp Met Ala Phe Leu Gly Thr Arg Ala Gly Leu Leu Arg Ser Ser
 1220 1225 1230
 Leu Phe Val Gly Ser Glu Lys Val Ser Asp Arg Lys Phe Leu Thr
 1235 1240 1245
 Pro Glu Asp Glu Ala Ser Val Phe Thr Leu Asp Arg Phe Pro Leu
 1250 1255 1260
 Trp Tyr Arg Gln Ala Ser Glu His Pro Ala Gly Ser Phe Val Phe
 1265 1270 1275
 Asn Leu Arg Trp Ala Glu Gly Pro Gly Arg Pro Ser Ala Lys Gly
 1280 1285 1290
 Leu Pro Pro Pro Leu Cys Gln Thr Ile Leu Lys Arg Arg Asp Gly
 1295 1300 1305
 Lys Met Ser Trp Ser
 1310

<210> 13

<211> 400

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7473438CD1

<400> 13

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

WO 01/77174

PCT/US01/11206

Lys Phe Gly Phe Ser Ala Glu Asp Tyr Arg Glu Leu Glu Arg Leu
 125 130 135
 Ala Leu Gln Ala Glu Pro His Arg Ala Gly Arg Gln Trp Lys Phe
 140 145 150
 Pro Gly Ser Phe Tyr Phe Ala Ile Thr Val Ile Thr Thr Ile Glu
 155 160 165
 Tyr Gly His Ala Ala Pro Gly Thr Asp Ser Gly Lys Val Phe Cys
 170 175 180
 Met Phe Tyr Ala Leu Leu Gly Ile Pro Leu Thr Leu Val Thr Phe
 185 190 195
 Gln Ser Leu Gly Glu Arg Leu Asn Ala Val Val Arg Arg Leu Leu
 200 205 210
 Leu Ala Ala Lys Cys Cys Leu Gly Leu Arg Trp Thr Cys Val Ser
 215 220 225
 Thr Glu Asn Leu Val Val Ala Gly Leu Leu Ala Cys Ala Ala Thr
 230 235 240
 Leu Ala Leu Gly Ala Val Ala Phe Ser His Phe Glu Gly Trp Thr
 245 250 255
 Phe Phe His Ala Tyr Tyr Tyr Cys Phe Ile Thr Leu Thr Thr
 260 265 270
 Gly Phe Gly Asp Phe Val Ala Leu Gln Ser Gly Glu Ala Leu Gln
 275 280 285
 Arg Lys Leu Pro Tyr Val Ala Phe Ser Phe Leu Tyr Ile Leu Leu
 290 295 300
 Gly Leu Thr Val Ile Gly Ala Phe Leu Asn Leu Val Val Leu Arg
 305 310 315
 Phe Leu Val Ala Ser Ala Asp Trp Pro Glu Arg Ala Ala Arg Thr
 320 325 330
 Pro Ser Pro Arg Pro Pro Gly Ala Pro Glu Ser Arg Gly Leu Trp
 335 340 345
 Leu Pro Arg Arg Pro Ala Arg Ser Val Gly Ser Ala Ser Val Phe
 350 355 355
 Cys His Val His Lys Leu Glu Arg Cys Ala Arg Asp Asn Leu Gly
 360 365 370
 Phe Ser Pro Pro Ser Ser Pro Gly Val Val Arg Gly Gly Gln Ala
 375 380 385
 Pro Arg Leu Gly Ala Arg Trp Lys Ser Ile
 390 395 400

<210> 14

<211> 269

<212> EMT

<213> Homo sapiens

<220>

<221> misc_feature

<222> Incyte ID No: 7474286CD1

<400> 14

Met Met Trp Ser Asn Phe Phe Leu Gln Glu Glu Asn Arg Arg Arg
 1 5 10 15
 Gly Ala Ala Gly Arg Arg Arg Ala His Gly Gln Gly Arg Ser Gly
 20 25 30
 Leu Thr Pro Glu Arg Glu Gly Lys Val Lys Leu Ala Leu Leu Leu
 35 40 45
 Ala Ala Val Gly Ala Thr Leu Ala Val Leu Ser Val Gly Thr Glu
 50 55 60
 Phe Trp Val Glu Leu Asn Thr Tyr Lys Ala Asn Gly Ser Ala Val
 65 70 75
 Cys Glu Ala Ala His Leu Gly Leu Trp Lys Ala Cys Thr Lys Arg
 80 85 90
 Leu Trp Gln Ala Asp Val Pro Val Asp Arg Asp Thr Cys Gly Pro
 95 100 105

WO 01/77174

PCT/US01/11206

```

Ala Glu Leu Pro Gly Glu Ala Asn Cys Thr Tyr Phe Lys Phe Phe
110 115 120
Thr Thr Gly Glu Asn Ala Arg Ile Phe Gln Arg Thr Thr Lys Lys
125 130 135
Glu Val Asn Leu Ala Ala Val Ile Ala Val Leu Gly Leu Ala
140 145 150
Val Met Ala Leu Gly Cys Leu Cys Ile Ile Met Val Leu Ser Lys
155 160 165
Gly Ala Glu Phe Leu Leu Arg Val Gly Ala Val Cys Phe Gly Leu
170 175 180
Ser Gly Leu Leu Leu Leu Val Ser Leu Glu Val Phe Arg His Ser
185 190 195
Val Arg Ala Leu Leu Gln Arg Val Ser Pro Glu Pro Pro Pro Ala
200 205 210
Pro Arg Leu Thr Tyr Glu Tyr Ser Trp Ser Leu Gly Cys Gly Val
215 220 225
Gly Ala Gly Leu Ile Leu Leu Leu Gly Ala Gly Cys Phe Leu Leu
230 235 240
Leu Thr Leu Pro Ser Trp Pro Trp Gly Ser Leu Cys Pro Lys Arg
245 250 255
Gly His Arg Ala Thr
260

```

```

<210> 15
<211> 489
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Inocyte ID No: 7472589CD1

```

```

<400> 15
Met Ser Ser Arg Ser Pro Arg Pro Pro Pro Arg Arg Ser Arg Arg
1 5 10 15
Arg Leu Pro Arg Pro Ser Cys Cys Cys Cys Cys Cys Arg Arg Ser
20 25 30
His Leu Asn Glu Asp Thr Gly Arg Phe Val Leu Leu Ala Ala Leu
35 40 45
Ile Gly Leu Tyr Leu Val Ala Gly Ala Thr Val Phe Ser Ala Leu
50 55 60
Glu Ser Pro Gly Glu Ala Glu Ala Arg Ala Arg Trp Gly Ala Thr
65 70 75
Leu Arg Asn Phe Ser Ala Ala His Gly Val Ala Glu Pro Glu Leu
80 85 90
Arg Ala Phe Leu Arg His Tyr Glu Ala Ala Leu Ala Ala Gly Val
95 100 105
Arg Ala Asp Ala Leu Arg Pro Arg Trp Asp Phe Pro Gly Ala Phe
110 115 120
Tyr Phe Val Gly Thr Val Val Ser Thr Ile Val Arg Glu Glu Ser
125 130 135
Pro Pro Leu Ala Leu Thr Pro Gly Arg Leu Cys Ser Asn Thr Gly
140 145 150
Arg Leu Cys Asp Leu Thr Phe Lys Ser Tyr Ile Asn Ile Ala Lys
155 160 165
Glu Gln Glu His Pro Ala Ile Gln Gln Ser Phe Pro Arg Val Ser
170 175 180
Thr Val Ser Ser Glu Asn Arg Lys Gly Gly Phe Gly Met Thr Thr
185 190 195
Pro Ala Thr Val Gly Gly Lys Ala Phe Leu Ile Ala Tyr Gly Leu
200 205 210
Phe Gly Cys Ala Gly Thr Ile Leu Phe Phe Asn Leu Phe Leu Glu
215 220 225

```

WO 01/77174

PCT/US01/11206

Arg Ile Ile Ser Leu Leu Ala Phe Ile Met Arg Ala Cys Arg Glu
 230 235
 Arg Gln Leu Arg Arg Ser Gly Leu Leu Pro Ala Thr Phe Arg Arg
 245 250
 Gly Ser Ala Leu Ser Glu Ala Asp Ser Leu Ala Gly Trp Lys Pro
 260 265
 Ser Val Tyr His Val Leu Leu Ile Leu Gly Leu Phe Ala Val Leu
 275 280
 Leu Ser Cys Cys Ala Ser Ala Met Tyr Thr Ser Val Glu Gly Trp
 290 295
 Asp Tyr Val Asp Ser Leu Tyr Phe Cys Phe Val Thr Phe Ser Thr
 305 310
 Ile Gly Phe Gly Asp Leu Val Ser Ser Gln His Ala Ala Tyr Arg
 320 325
 Asn Gln Gly Leu Tyr Arg Leu Gly Asn Phe Leu Phe Ile Leu Leu
 335 340
 Gly Val Cys Cys Ile Tyr Ser Leu Phe Asn Val Ile Ser Ile Leu
 350 355
 Ile Lys Gln Val Leu Asn Trp Met Leu Arg Lys Leu Ser Cys Arg
 365 370
 Cys Cys Ala Arg Cys Cys Pro Ala Pro Gly Ala Pro Leu Ala Arg
 380 385
 Arg Asn Ala Ile Thr Pro Gly Ser Arg Leu Arg Arg Arg Leu Ala
 395 400
 Ala Leu Gly Ala Asp Pro Ala Ala Arg Asp Ser Asp Ala Glu Gly
 410 415
 Arg Arg Leu Ser Gly Glu Leu Ile Ser Met Arg Asp Leu Thr Ala
 425 430
 Ser Asn Lys Val Ser Leu Ala Leu Leu Gln Lys Gln Leu Ser Glu
 440 445
 Thr Ala Asn Gly Tyr Pro Arg Ser Val Cys Val Asn Thr Arg Gln
 455 460
 Asn Gly Phe Ser Gly Gly Val Gly Ala Leu Gly Ile Met Asn Asn
 470 475
 Arg Leu Ala Glu Thr Ser Ala Ser Arg
 485

<210> 16

<211> 1735

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Inuyle ID No: 1784775CB1

<400> 16

atgtgctctc ttgtcttccc ccttgtcccc agtagtccag attacatcct accctgcagt 60
 cctggatggc gccctccgact tgcagcttcc ttctgtcttt cgtctctccc gctgctagac 120
 attcttccag ttgtcttggc accaggggca gcccacaggac ccatagggtt agaggtgttg 180
 gcagggttgc ttgcagctgt gccctggatc agcccacagcc tggcctgttg ggtgttgcca 240
 cattcccttc atggccactc ccgggggtccc ttggccttgg cctgggttag cttgctggca 300
 gctccagccc tagtgctgac cgtgttgttg catlgtccagc gaggccactc totgccccca 360
 cttctcccag ggcccatggc ccgcttatgc ttgtctatcc tgcagctggc tgcactcttg 420
 gcttatggac tgggatgggc agctctctgg ggaccacagag aacctgtggc tcaggagccc 480
 ctcttgcccg aggatcaaga acctgaggtg gctgaagatg gggagagttg gctgtcagc 540
 ttttctctat cctggctggc acccttgctg gcccttgggg cctyltgaga gctccggcag 600
 cctccaggaca ttggccgctt ccccccacaga ctgcagccaa cctacttggc tctgtgcttc 660
 caggcacacc ggcaggaggg ggcacaggctg tggagggctt tglatggggc ctttggccgg 720
 tggatcttgg cacttggact gctgaagctg gtcggagaca ttttgggatt ctcaggcccc 780
 ctgttgcttt cctcatgggt gggcttcttg gaagaggggc agagccactc aagccacggc 840
 ctgctctatg ctctggggct agccggctgg gctgtgctgg gtgctgtgct gcagatccag 900
 tatgggtatg aggtatataa ggtaacactt caggcacggg ggcctgtgct gaacatccctg 960

WO 01/77174

PCT/US01/11206

```

taacgcaagg ctttaccgct ggggcccagc cggccctcta ctggggaggg cctgaacctt 1020
ctaggcaactg actctgaacg gctgcttaac ttgtctggga gcttccatga agcctggggc 1080
ctggccctctg aactggccat caccctccac ctgctgtaac agcaggtagg cgtggccttc 1140
gtgggtgggc tcatcttggc actgctgctg gtacccgtca acaagtgat tggccaccgc 1200
atcatggcca gcaaccaggc aatgctacag cacaaggatg cgggggttaa gcttctgaca 1260
gagctgctga gtggcattcg ggtcatcaag ttctgggggt gggagcaggg actgggagct 1320
cgagtagagg cctgcccggc tggagagctg gggcgacttc ggttcacaa atacctggt 1380
gcggctgctg taccctcgta ggttgcctta ccggtcgta tctccatcg tatcttctt 1440
acctatgtcc tcatggggga cagctccact ggcaccaaag tggagaccgg yagggaaggy 1500
gaccagcctc aaggagactt cagcgaagtg aagacagagg cttgggcctt cagtcctggc 1560
tggagagggg ggggggagtc ctgactgctt catctctcaa ccaggggaaa actgaagaaa 1620
ccctttttgt ggggcccctg atataaacd ctccctctgt tgaaggagtt ccttctcttc 1680
ctctctctta ccttctactc cggcttctat tcaagttccag gcttgggggtt aggttc 1735

```

```

<210> 17
<211> 1041
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7473034CB1

```

```

<400> 17
atggttcaag catctgggca caggcggttc accgctggct ccaaaatggt ctctgggtcc 60
gtgatagcaa agatccaggc aatutggtyc gaggagatg agaggaaagt ggtggagag 120
ttcttggcgc agtcatgag cacaatgttc atgatggatc tgggcttggg ttctgggttc 180
catatggctc taatatataac atatgggagc tacttgggtg tcaacttggg ttctgggttc 240
ggggttcaaca tgggagttca cgtggcaggc cgcctctctg gaggccatct gactgcagct 300
gtgancttca cttaactgtg gctgggcccgc gtgcccctga ggaagtttcc agtccatgtg 360
ctggggcagc tcttgggctc ctctctggca gctgccaacc tctacagttc ctctcagag 420
gtcatctctc acttcttggg tggagagctg atggagagcg gtcccttctg tactagctggc 480
atttcttcca cctaccttcc tgaatcactg acatcttggc ggggcttctc gaagagagag 540
tgggtgacca ggtgtctcca gctgtgtctc ttaacatcca cgggacaggc gaacaaacct 600
gcaatgcccag gaacacacgc gctgggtgata agcatctctg tgggtcatct cagggtgttc 660
catgtcataa acacaggala tggcatcaat ccatctcggg acccggcccc cagcatcttc 720
accttctctg ctggctaggc caaacaggct ttcagcgatg gggaggaactg gtgttgggtg 780
cagctggttg caccacttct ggttgctctc ctagggtggc tcatctacct ggtcttcat 840
ggctccacca tcccaaggga gacctgaaa ttggaggact ctgttggcgt tgaagaccac 900
gggataaccc tattgcccaa gatgggatct catgaacca tgaatctctc cctccacctc 960
atctccgtga gcttgcaca cagatcttca gtccactctg ccccaacctt acatgaatcc 1020
atggccctag agcatttcta a 1041

```

```

<210> 18
<211> 2367
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1876581CB1

```

```

<400> 18
ggacgcgctc cggggacgag cagggtcgcc gtggcgggag acgcggttcc ggtggcggcg 60
gaggctgcac tgaagcggag ctgagagcag cggggcgpyc agcccggggg aagcgtccgg 120
gacctgtctt ggagaactac caccacacat taacatcaag gaacctcgat gggatcaag 180
caatltcatl ggacgagcca atcatttctt caatgttaact gacccacagg acatttctgt 240
aaccaacgaa caactcgaga gtgcpagaaa aatagtaaat gattacaggg cgggaattgt 300
tcttcttggc ctacagaaa atggaattgt gaggagcaag taactcatatg atctcagct 360
tcaatctgac actgggagga agatgatctt gatggagaa agtccagccc aggttccat 420
gaacatgacc aatcaggggt gtatgatgac gttttacag actacccagg ctgtctgttt 480
ctggcaggtg attaacagc cctcaatgc cgtcgtcaat taacccaaca gaagtggaga 540
cgaacccctc actgtcaatg agtgggagc agcttncgtt tatgcaaan ctggtccggt 600

```

WO 01/77174

PCT/US01/11206

```

agcaacagct ctaggactca atgcattgac caagcatgto tcaaacctga taggacgttt 660
tgttcccttt gctgcccgtag ctgctgctaa ttgcattaat attccattaa tggaggcaag 720
ggaactcana gttggcattc ccgtacacgga tggagaaagg aacccgttgg gggagtcggc 780
gaacgtctgc aaaaagccca tcaagcaagt tggctgtgctc aggtatttca tggcagcccc 840
tggcatggcc atccctccat tcaattatgaa cacttttgaa aagaaagcct tcttgaagag 900
gttcccatgg atgagtgcaac caattcaagt tgggttagtt ggtttctggt tgglytttgc 960
tacaacccctg tgttctgccc tgtttccctca gaaaagttcc atgtctgtga caagcttggc 1020
ggcggagttg caagctaaaga tccaaagagag ccaacctgaa ttggagcccg tgaacttcaa 1080
taaggatgtg taagcagagg aggaacctc tgcagctcat tctgccaatg caaagctgggt 1140
gtagccatgc tggtagaaga aatcctgttc aacctgggtt ctccagttc cggaaacctt 1200
ttaaacatcc acactagcct tttagaataa agctgctatt ttaacagagc acctggcgtg 1260
ggcccaagtc ctgatactcc cttaactctga atcatgttat gatttataya aatccctttc 1320
ctgtagcttt tatagtcatt gtttttcaaa gacgacatad cagcctccac ccaggtttta 1380
aaaaagcact ggtaggcata gaattagggtc tcaagtatat gtcagtaaat gttctataga 1440
ttatcaatca ctgaaacaaag aaattctgttt aaaaacttga atttccatct cactccctatt 1500
gcaaatcaag gagaactcag cagtgaactg ggaataatca aaagctctgg gctaatctat 1560
aaaaacttca cctgaatatat taagggaagt ttggttccag ttgggggatt ggcctagccc 1620
aatgaaggtg atgaagcttt tggattttga gggtaaaagc tctttccacac ccttccaaa 1680
agtcagtcac agaacactgc aacatgcctt cctgtctgga tcaattatata cactcagatt 1740
gtgggtggat tgccttggtt gacttttaat ttatgttttt ttgtctttat aaagatgata 1800
atcttacttt gcaatgatgt aatttatatt caattatata catcaataaa cgaatgaact 1860
gaaatgtaca atgtcaaat tttagagta tattcaatad caatgctgta tgaagtggtc 1920
gaatccagtt catgtttttt tttttggtaa gaagtgaagc tcaagttoca gctactaca 1980
tgccttttct tgcctactct atagatctct ttggctttca gaaagatata gtcataatgt 2040
gtgtatgaat cagtcccaat gaatttttact tgaatattgt atgttgcat ccaactcat 2100
tgaataatct gaacacatgt accactgttt acatcatctg tagtgatttc atagatcata 2160
tatttaatat cagagattat gtttccactc tgtatgtgt taaagctata gacagttggc 2220
cctctgtatc cgtgagctct atactgtga attcaacca gtttgatgg aaaaattttt 2280
tttttttttt ttttcttgag acggagtcct gctctgtcac ccaagctgga gtgcagttgc 2340
gtagttctgg ctcaactgaa gctccgcg 2367

```

<219> 19

<211> 3343

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2346292cbl

<220>

<221> unsure

<222> 57

<223> a, t, c, g, or other

<400> 19

```

cttagacggg cgcacgcggg atgggagggc gggggaagca gggggaagag cggacgncgc 60
accgcgtccg ggcgcgtctt caatgagcag cgcggaaact gcaacccaga cccagacctg 120
ctgcgcgcgc cctcccgagc ctcaactgtt gacaggtaac aggcctggcc tgcgccctgt 180
gatgatgatg gcccttgccc cgtgagctac aacctggcct tcagcaccgc ccaacttcca 240
accagcagga tgcggctgtg gaggcggtg gtggtgactt tggccttaat gagtgtgac 300
atctgcgtga ccacggccat ctatgtcttc agccactcgg accgcagcct cctggaggac 360
atccgcactc tcaacatctt tgcctcgtgt ctgagatctc gggcagcctg cctgtaccgc 420
agctgcctgc tgcctggagc caacattggt gtggccaaga acagtgcgct ggggccccgc 480
cggctgcggg cctcgtggct ggtcatcagc ctctgtgtgc ccttctggg catctatgcc 540
atgggtgaagc tgcctgctct ctcaagagtg cgcaggccca tccgggaccc ctggttttgg 600
ccctcgttct gttggacgla catctcaact ggcgcatcct tcttgcctgt gtgcgtgctg 660
tccacatgac ggcagggcac ccaggccctg gggccagggg cggccacgga gctcagggc 720
tctctgggga gggccgggct accgcgcgag caggcgtctg gggccacgct gcaagagctg 780
cttctctaca ccaagccgga cggggccttc cctgtggcgc ccttctctct cctcctgtg 840
gcagctctgg gaaagacctt cctgacctac tacaagggcc gcgccattga tggcactcgt 900
atccagaaaa gcatggatca gttcagcaag cgtgtcttca tctgtgcct gctgcacatt 960
ggcagctcat ttgcgcagg tattcggggc ggcattttta cctctcatat tgcacagactg 1020

```


WO 01/77174

PCT/US01/11206

```

aacattatgc  ttccaaactg  tctcttcggc  tcaclggltg  cccaggagac  aagctctttt  1050
gtcgagaaac  gcaacaggga  cctcatctcc  cgcctgacct  cggacaccac  catggbcagc  1140
gaacgggtct  cccagaaacg  caatgtcttc  ctggcggaac  cagtcagggt  caggggcgctg  1200
gtggctctca  tgttcagcct  ctcatggcag  ctctctctgg  tcaacctcat  gggcttcccc  1250
atcatcatga  tgggtgtcaa  catctacggc  aagtcactac  agaggctctc  caaaggagtc  1320
cagaatgcgc  tggccagagc  gagcaacacg  gcygagggag  ccatcagtcg  catgaagact  1380
gtccggagct  tggccaatga  ggaaggagag  gcaggggtgt  acctgcggaa  gctgcagcag  1440
gtctcaaacg  tgaacaggaa  ggaagcactc  gcttactgtg  actacgtatg  gggcaggcgg  1500
tcggcgagct  cgtctctacg  tggcctgatg  caggagatgg  gggctgctga  gaaggtgttc  1550
gagttcatcg  aacggcagcc  gacatgggtg  cccgatggca  gcttggccgc  cgcacacctg  1620
gagggcgagg  tggactttga  gaatgtgacc  tccacctacc  gcactcgacc  ccacacccag  1680
gtcctgcaga  atgtctctct  cagcctgtcc  cccggcaagg  tgaacggcct  gctggggccc  1740
tcgggcagtg  ggaagagctc  ctgtgtcaac  atcctggaga  atttctaccc  cctggagggg  1800
ggcggcgctc  tgcctgagcg  caagcccaac  agcgcctacg  accacagta  cttgcacdg  1850
gtgatctccc  tgggtgagca  ggaagccgtg  ctgttcgcc  gctccatcac  ggataacac  1920
tctcaagccc  tgcctcactg  gctctctcag  atgttgggtg  aggcgcacac  gaagggccaa  1980
gcccaagcct  tcatcatgga  actccaggac  ggctacagca  cagagacagg  ggaagagggc  2040
gcccaagcct  caggctggca  gaagcagcgg  gtggccatgg  cccgggctct  ggtgcggaac  2100
ccccagtcct  tcatctctga  tgaagccacu  agcgtcttgg  atgcggagag  agagtactct  2150
atctaggcag  ccatccatgg  caactctcag  aagcatcacg  tactcatcat  cggcaccagg  2220
ctgagcaacc  tgaagcagcg  gcaactcatt  gtggtcttgg  accagggcgg  cgtagtgcag  2280
cagggcacct  accgagagct  gctggccag  ggcggcctct  accgtacgct  ggtgcagcgt  2340
cagatgctgg  ggcctcagcc  ggcgcagac  ttccagctg  gccacaacga  gctctgagcc  2400
aacggcgagt  caaaggccct  atggggggcc  cctgtctctc  ccgattgggg  agaggacccc  2450
gtcctcctcc  ggcagatctg  cccacggagg  ccccccgtcg  cctctcagcg  caggccctcg  2520
agcactgaaa  gacgacctgc  catgtcccat  ggatcacccg  ttcttgcac  ttgcccactg  2580
tccctgcctcc  attccacagg  cactctctac  cctctctgcc  ctgagccaac  gctctccagg  2640
acctccctag  cctccctaac  aaaggtagag  ctgctctctt  aaaaactaggt  cttacacagg  2700
ttttactactg  ttgggttgag  gcaacccact  caactctcag  alttcaaaaa  ccttttctta  2750
attgggagta  atggcgggca  cttccaccaa  gatgtcttag  aaactctcga  gccagagtg  2820
aatggccctt  ccttagtagc  ctgggggatg  tccagagact  aggcctctcc  cctttacccc  2880
ccagagagag  gggctctctct  gtcccgagg  gagaacccgg  gaacgggatt  ttccctctct  2940
ccctcttccc  agctctgtga  gtctggccag  ggcgggttag  gaggctggag  ggcattcttc  3000
tgcctatccc  cgtctgcact  ctgaagcagt  ctccctgtga  acccaacgaa  cctctcaactg  3050
ggggagttag  gggctggcca  ggtctggagg  ggcctcaggg  gtgccccag  cccggcacc  3120
agcgtcttcc  cccctcgtcc  accacccctc  ggtctgagc  ctccctcccc  accacccccc  3180
ctgtctctct  ctgtctggag  gccacgtgga  tgttcatgag  atgcattctc  ttctgtctct  3240
ggttgatggg  atgtctggca  agccacggat  ctggctttgc  cagggtgtgc  aacatgttga  3300
gagaaacccg  tcaataaagt  gtaataacct  ctacccctaa  aaa  3343

```

<210> 20

<211> 3517

<213> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5151730CB1

<400> 20

```

ccggaacgag  ctgggctggg  ctgggctggg  ctggggcggg  cgcaggcgcc  agggggcggg  60
gcggggggga  agaagcagg  ggggctctgg  ctctcccggg  gagcgcccg  ggaatgcacc  120
gggaacagcg  cctcccgcct  tggcgtctgc  ctgggctcgc  ccccgggccc  ggggtggatga  180
gccgcgcgcc  cgggggacat  ggaagcgtg  acgtctgtgc  ttctcccctg  gatattccag  240
tgcgtgtcgg  tggggggcga  ctccatctac  cactctgggt  ccatctctga  ggaagacccg  300
gccagggtag  acaggggtgt  ccagtggcgg  gtatccgacc  tgaacctcaa  cgaatgacat  360
ctgcagagcg  agaagatcac  ctactccatc  aagttcatcg  aggcacacaa  cccattccag  420
gctgtgaggg  aagcctgtga  ctcatgacc  caaggatatt  tggccttgg  caagtcacac  480
ggctgtgcat  ctgccaatgc  ctgcagctcc  ctccagatg  ccatgcagat  cccacacccc  540
tttgtctcgg  gcaacccggg  aggtctgcca  ggcctcgact  gccactgaa  cccacggccc  600
gatgtgaggg  cttaacacat  ggcctcagaa  ccacccctcc  gcttcaatga  tttcatgctc  650
aggctgtgtg  cggagctagc  ctggcagag  ttctctatgt  tctacgacag  cagatgatat  720
atccgtgggg  ttcaaaagct  tctggaccag  gccctcggcg  tgggctctga  cgtctcttta  780

```

WO 01/77174

PCT/US01/11206

```

caaaagggtgg acaagaacat tagcaacgta ttaaccagcc tcttcaccac gatgaagaca 840
gaggagctga atcgctaccg ggaacagctt cgcgcgcgca tctgtctgct cagcccccag 900
ggagcccact ccttcaccaa cgaagccgtg gagaaccaac tggcttccaa ggaacccac 960
tgggtccttg tgaatgagga aatcagtgac ccggagatcc tggatctggt ccacagtgcc 1020
cttgggaagg tgaacgtggt cgggcaaac cttcgtctcg caaaggacaa tcaagaatgc 1080
acgaggaaac accacccgat ctcctccctg ctctgggacc cccaggaggg ctacctccag 1140
atgctgcaga tctccaaact caactctgat gacagctatc tgaatctgag caacgctctc 1200
caaggaagc tggagaccg gaaatggcat agcatggcga gcttcaactg caacgggaaa 1260
tccactaagc cctgggaatg tgggaggtcc atgctggata ccatcaaaa ggcaccatcc 1320
actgacctca ctgggtgat ggaatltcgg gaggacagtt ggaatcccta tctccagttt 1380
gaatctcttg gcactaccta tagtagact tttggcaag acatggcga gttggcgaca 1440
tgggactcag agaaaggctc gaatggcagc ttccaaagaa ggcacatggg cagccgcctc 1500
caaggataga ctcttaaaat ggtgactgtc ttgggaagac ctltctgat ggtggtgag 1560
aacatcttag gaaagcccaa ggcctacaaa ggggttctca tagatgtcct ggaatgactg 1620
gccaaagctc tgggtcttaa atatgagatt taccaaagcc ctgatggcag gtacgggtac 1680
cagctccata acacctcttg gaacgggagc atcggggagc tcatcagcaa gagaagcag 1740
ttggctactc ctgcctcac catcaccca gagagggaga gcttctgga ctccagcag 1800
cgtacatagg antattcagt ggggtattcta attaagaagc cggaggagaa aatccagcgc 1860
ttctccctct ttgctccatt tgatttcgct ggtgggtccc gcatlycagc agcatccct 1920
gtgtgttggt tctgtatatt tctgtgaac aggaatacag ctgtgaggg tcaagatgct 1980
gctcagcga ggcgttcagc tctgtcact ctgcacagcg ccatctggat tctctataga 2040
gcttctgac agcaagtgga cgaatcttcc gtgaactcca tggcctatcg cctctgtag 2100
ggcagctggt ggtcttctac gctcttctgt tctctctctc aaacagccaa ccttctgccc 2160
ttctccacag tctccagat ggaacaaccc ataaagactt tccaggccct gtcacaacaa 2220
gtggaaatgc ctctatggcc tctcgggat tctgtcttat atgagtactl ccagagcaag 2280
ggacaacacc cctggagca ggaacagcag ttgtctgaac tctggcggag catcagcaag 2340
aaggaagggg ctgacaactg cgtgtccagt ccttccagag gcatcaggaa ggcacaagag 2400
gggaactcag ccttctgtg ggaatgtggc gtgtgtggaat agcatccct gaaggatgac 2460
gactctctgg tgaatgtcat ggcacaacgc atcagcagca aggttcaagg gattgcccc 2520
cagcatggca gccctacag ggaactcttc tccagagga tctggagct gacagacaa 2580
ggggccttgg atgtctgaa gcaagaatgg tggcgcgaca tgggcccgtg tggcctccac 2640
agcctatgca ggcctcaggc gacggcaca tccctcaagc tgcacagctt cgcggggc 2700
tctccatccc tggcactag cctgctctg gctggcctga tggctgccc ggaattgtgg 2760
tggaaacaga acgggtgcca ccaggagacc cccaagagg acaaaagagt gaacttggag 2820
caggttccac ggcgcagca cagctctatg gatcaagaca ttgctccaa ccagatttcc 2880
ccaggttcca ttgagctctc ggcctcggag atggggggcc tggctccac ccagatttcc 2940
gagcgcagac gggagtacca gaaccccag cctcgggtca gcaacttctc gccagagcag 3000
agcagcagc gccacagcgc gacactctca tccggggcca gcaacaaact gccgctgag 3060
ctgagcagct cggcagccat gccctctatg cagtgcacac acaggttacc caacggggg 3120
ctgttccggc agagcccggt gaagaccctc atccccatg ccttccagcc cgtctctgga 3180
gggtctcttc cagaggtctc ggaacacctc caggggacct ccatctgact ggcgcgctg 3240
ccctctgccc caccctccca cccaccgcag cagcagagct ttttaataca aggaataaac 3300
aacacaaacc acacacatc gacacacac acatacacag agactcttc attttctctg 3360
tcatatagc taataataga agaatggag tggggtaaaa gtgtatttct astattccca 3420
atttctgag cagtaaaa aacacaaaa ctgtatgaat gactttgtaa attttctct 3480
atatgaataa aaagcaaat tacttctgaa aaaaaa 3540

```

<210> 21
 <211> 1248
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Inctye ID No: 7472584081

```

<400> 21
atgaagagc gactctgtgc ctactgtctt gacacttacc ctactgtctg 60
gtgggagcc cgtgtctcga ggcctctcag tgggacacag agatgtgcga ggaagagaa 120
ccaaagcgc aggaatctcg gatcaagggg aagtacaaca tcaagcgcga ggaactcccg 180
cagctggagc tgtctatcct ccagtcggaa ccgacccgag ccggcgtcca gtggaatttc 240
gctggtctct tctactttgc gatcagcgtc atcaccacca taggttatgg gcaagctgaa 300
cctggcagcg atggggcaa ggcctctctc atgttctcag ccgtgtcggg catcccgctg 360

```

WO 01/77174

PCT/US01/1206

```

aacctggtca  tttccagag  cctgggagag  cgcattgaca  ccttcgtgag  ctacctgctg  430
aagcgattca  agaatgctg  tggcatgagc  aacactgagc  tgttcattga  yaacatggcg  450
actgtgggct  tcttctctg  catgggagag  ctgtgcatcg  gggcgggcgc  ctctcccgag  540
tgttaggaat  gtagcttctt  ccagcctac  tactactgct  tcatcacttl  gactaccatt  600
gggttcgggg  actacgtggc  cctgcagacc  aagggtagcc  tgcagaagaa  gccgctctac  660
ttgggtctta  gcttatgtta  tatctgggtg  gggctgacgg  tcatcggggc  ctccctcaac  750
ctagctgctc  tcaggttctt  gacatgaac  agtgaggatg  agcgggcgga  tctctgaagag  780
agggcatccc  tcgcccgaaa  ccgcaacagg  atggctcatc  acatccctga  gggagccggg  840
cccacccggc  ccaggtacaa  ggcggacgtc  ccggacccgc  agttctatlg  ctctgacacc  900
tgctacgctc  ccagggacta  tggcgggcgc  lcggtggcac  ccaggaactc  ctccagcgcc  960
aagcttgccc  cccactactt  ccactccacc  tcttacaaga  tcgaggagat  ctccacaaagc  1020
acattaaaaa  acagctctct  cccctcgcct  attagctcca  tctctcctgg  gttacaacagc  1080
tttacggacc  accagaggtc  gatgaacgc  cggaaagtcc  tttaggggaa  ctaactgcac  1140
attcaagaga  cggctccgtg  gatgctgggt  ctcactgcga  aagccgaaca  cggcttcggg  1200
attctctgcc  ttctcaagt  gacctcttgc  tgtgctgggc  ggaatgcc  1248

```

<210> 22

<211> 1770

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7472535CH1

<400> 22

```

atgtctgagc  tctcaaccc  tggcctgttc  ctccccctcc  cagggcacct  ccgggctctc  60
ctacataaag  ccgggggac  tgggcctcag  gtcaggacct  tgggtatggc  ctccacagac  120
ctgtcgggag  ctctgggag  catgggcgc  ttccagctca  accacacagc  cctgctgtgt  180
ctgcccctgg  gctgctggc  ctgcacacac  tccctgcaga  acttccacgc  cgtgtctccc  240
cccacacct  gcccgggccc  tcccacccac  actgaagctt  ccauccacga  ctggggggcc  300
tggctgaggg  ccacataacc  cccggaacag  ctggagcttg  ccggcgcttc  360
accagacctc  agtgggacct  gctgagcccc  aactcctcca  tcccgggggc  ggcacaggag  420
ggctccaggg  accgctgggt  ctataaacgc  agtctctccc  cgtcccccac  ggtgctggag  480
cagctgggac  tgggtgtgga  ggcgccacct  ctccagagac  tggcggaagt  cgtctacatt  540
gccggcgctg  tgggtggggc  tgcctgtgtt  ggcagcttgg  cagacaagct  gggctgcaag  600
ggcccccctg  tcgggtccta  cctgcagctg  gcagcttctg  gggccggcac  agcgtalltc  660
agctcttcca  gtgcctattg  cgtctctcgg  ttccgtatgg  gcattgacct  ctctgggcat  720
atctccaact  ccgtctccct  ggtggagtag  atgcccacac  gggcgccgac  tglggcggtt  780
atttctgtgg  ggtattctct  caccctgggc  cagctcacc  tggctggggg  agcctacctg  840
attcggccct  ggcggtccct  gcagtttgcc  atctctgtct  ctttctgtat  cttttctctc  900
tattcttggc  ggcctccaga  gtcattccgc  tggctctccc  tgcattggca  gtcccagtta  960
gctgtccaga  atctgcagaa  ggtggtcgca  atgaacgggg  gggagctggg  aggggagagg  1020
ctgacacagg  aggtgatgag  ctcttaccac  caaagcgagc  tlgcagagat  ctgcacctcc  1080
aactcaatct  tggacctctt  ccgaaccccg  ggcattcgca  aggtctcagt  ctgtccggcg  1140
ctgaggtctc  ccaactctgt  ggttacctat  ggcctggcca  tggacctgca  gaagtctggg  1200
ctcagcttat  accctgtgca  ggcctgtgtt  ggaatctcca  acatcccggc  cagtctgggt  1260
gccacggcca  ccatgattta  cgtgggcggc  cgtgcacagg  tggcctcctt  cctcatctgt  1320
cccgggctca  tgggtgacgc  caacatgttt  gtgcacagag  gacgcagat  cctgtgcaca  1380
gcccgggcag  cgtggggcaa  aggttgcttg  gccagctcct  tcatctgtgt  gtactgtttt  1440
accggcgagc  tgtacccccc  ggaatcaggg  cagatgggga  tgggctttgc  ctctgtccac  1500
gcccgcctcg  ggggcctgac  ggcgcctctg  gttaccacac  ttgggggaata  cagcaccalc  1560
ctgccaccgc  tgagctttgg  ggcacacgca  atctggctgt  ggtggcgct  ctgcttctgt  1620
actgagaccc  gcaacatgcc  cctggtggag  accatcgca  ccatggagag  gagggtcaaa  1680
gaaggtctct  caaagaacca  tttagaagag  aagagtgag  aaattttctc  tcagcagctg  1740
agagcatctc  cctcaaga  gaccatctaa  1770

```

<210> 23

<211> 2544

<212> DNA

<213> Homo sapiens

<220>

WO 01/77174

PCT/US01/11206

<221> misc_feature

<223> Incyte ID No: 7473422CB1

<400> 23

atgacgagatcc gccccgaacc gctgatgcgt accttggcgg ccgccgtgct gggccctgct 60
 atcggcgctcc cggccgcgcat ggcagagcga ccgtccacc ccgacatgta caacggccag 120
 cacaagagaa ccggcggaagc cgtcgcgcag gctccgagg ccagagccgg catcccaacc 180
 gatatccgca agggcgagcga caaccagctg gccagccaga tcatcgagga gggcgacggt 240
 tcggcagcgg acctcactca caccgaagag tcccccgcgc tgcacaacct gggcggaactg 300
 ggcctgctgg cgagagatcga cgaagccacc gcgaacatgg tgcaccaggga gtatctaggc 360
 gccaacggca cctggatggg catcacccgc cgcacgcgga tcatgggtga caaccggaag 420
 aaggctgatg aaaaagacct gccgaccaca gtgatggact tgcgaaccc tgggtgggaa 480
 ggcgcgctgg gctacgtacc caccagcggt gcattccagg agcggcgctg ggccatctcg 540
 aagatgcagc gtctgtgaag caccgaagaa tggctgaccg gccctaaagc ctccggcaaa 600
 acatcaccca acaacatggt cgcctcaaaa gccgtggaaa aagggtgaag ggctgcgga 660
 ctggtgaaca actactactg gtacgcactt gaacgcgaac gccgcaagct cgcacccaa 720
 ctctactacc tggcagatgg cgatgcgggc aacctgggtg caatctctag cgcgcctgtg 780
 gtcaaggcca ggaagcacc cgaagaagcc cgggcactgc tcaactgga ggccagcgaa 840
 gaaggccaac gctgatcac caagactacc gccgagtaac cgtctgacaa gggcatggtt 900
 tccggcggg gccctcaagcc gttccgaagc ctgcgcggc cgaacatctc gccagcgaa 960
 ctgggcaatg ccggcggaag catcgagctt gbaagcgagg ccggcgcggt acgtctctga 1020
 cggcgccgca agcgcctctc catctgggtg gtctgctctg tctctctctt ggtggcgatg 1080
 agcttctcac cactctctga ttctgccttc aaagcttggg aagccggctg cgttgaaagc 1140
 ttgcacctgc tctggcgccc cttctctctg gggtgatgag gcaatadct gatgctgatg 1200
 gtccgggtga catctggcctg catcgctgct gggctgctgc tggcctggt gatggagagc 1260
 agaaacctgc ctggcgccg cctgtggggc gtggtgcttc gcttgccttc cgtctgctgc 1320
 tctgtctgta gcagttctad ctgggtgtcg ctgagctcgg acttcgaagc gctggcgagg 1380
 gccatctctg tcatggccgc gtcacagtag caattgggtt tcttgcgggt ggccggcacc 1440
 ctgcgcaacc tcgacacctc gctggagagc tggcgcgca ccttgggttg tagccgcttg 1500
 ggcgtgttca tcaaggtcac cttgcgctg ctgtgacctc gcatgctcgg cggagcctg 1560
 ctgacgcgcc tgcacatgct ggtggagttc ggtgcgttgt cgaacctcgg cctgcagacc 1620
 ttcaaccgg cgaatcacca gcagttcgaa ctggatctca gcaatgcga cggcgcatg 1680
 cgtgtctggc tctgtttgag gattgtcttg gtgatgctgt ggtctgaatt cgtgtatcgt 1740
 ggcgaagccc gccatgtcgg catcgccagg gacttgagca gccgcgcga accgttcga 1800
 ctgctgtgct gggccgtacc ggcgcagcta ctctgcgtgg ccctggcagt gctggcgagc 1860
 ggcattccac tggccatgct cggctactgg ctgagcylgg gttcttccgg gcccttcccg 1920
 gtgggagcca tctccagcc cctgttccac tctgtctcgg tctgtcttgg cgtgtccgtt 1980
 ttctgtgtgc tctgtgcctg accgataagg tctctgttgg tgcgctctaa aggcctctt 2040
 gcgaltctgg ccgagcgctt gccgtacctg ctgcacgccc tgcacggcct ggtgattgca 2100
 ctgaccttgg tgttcttccg cctgcactac gtcgcggcgc tgtaccagac cagcgcttg 2160
 ctgctcttgg cgtatgctt cctgttctct ccatgggggc agtcaacggt ggcgacccg 2220
 ctgaccaagg cctgcgcaac actggagaa gccgcgcga ccttgggtgc cagcagcttc 2280
 accgcatctt ccggggcgac cctgcgagc atctctccgg ccatggcggt agcatcttgc 2340
 ctggtgtttc tggatgcat gaaagaactg accgctaccc tctgtctcag ccgacccgg 2400
 atgcacccc tggctacgga ggtgtggggc caaacggcca acgtcgagtt ccggcgagg 2460
 gccctctatg cagccttctg gatctgtgtg taaggcccg cagcttatct gctgacccg 2520
 cggatgtacc tgaacnaggc ataa 2544

<210> 24

<211> 2871

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3864715CB1

<220>

<221> unsure

<222> 2847

<223> a, t, c, g, or other

<400> 24

WO 01/77174

PCT/US01/1206

```

gctctctgccc tgcctctctgt gtcctgagccc gtcctgagccc caccatgggtt agctctctgccc 60
tgcccaaaccc cggccttagac gcccggatccc cgtccctggc tgcctctgag accatcgagc 120
aggagggaggg cagctcccgcc cggaggtggg acaacaaggcc gcatgacatg cccacctgccc 180
tgggtctcttg cgtggggcttc ggcacgctgt gngctctccc ctacctgtgt cagagccacg 240
gaggaggaggg ctctcatgac cgttctctca tctctgtggt cctgagaggc atctctctgccc 300
tgtacctgga gtctgcccac ggcctgagccc tgcctgagccc gctctctgccc tctctgagcc 360
ccatccatccc ggcctgaggg ggcctgagccc tgcctgagccc gctctctgccc tctctgagcc 420
gactctctctt caacccctctt atctctctgga tgcctgagccc gctctctgccc tctctgagcc 480
agctctgccc ctggagggagc tggcctgccc gacaggggat gacaggggat atctctctgccc 540
gctcagggag ctccctctggt gactactctt ggtacggaga gacgctcacc atctctctgccc 600
ccatccgcca ctctgggtccc atctctggtt ggtatgctgt gctctgagccc tgcctgagccc 660
gctctctgta catgtgccc atctctggtt tgcagagccc cgggagggccc gctctgagccc 720
ctccacgctt gctctctggt gctctgagccc tctctctggt cggagggccc gctctgagccc 780
gctcagggag tgcctctggt tctctctggt cggagggccc gctctgagccc gctctgagccc 840
acacctggtt ggcctgaggg ggcctgaggg tctctctggt cctctctggt tctctgagccc 900
tctctctggt cctcagggat aactctggtt acaacaaggc cgaagaggac tgcctgagccc 960
tgcctctggt caacgggttc acatctggtt atgtgagccc cgtggtgagc tgcctgagccc 1020
ggttctgagc cagcggggc tgcctgaggtt gctctgagccc gacatctggt accctctgccc 1080
acgggttctga cctgctgaga ggcacaggtt cccagggaga cttctgagccc atctgagccc 1140
gggtgagagcc cctcagggcc ggcctgaggtt gctctgagccc gctctgagccc tgcctgagccc 1200
agctctctctt ctgagggccc ggcagggccc cagggggccc cttctctggt tgcagggccc 1260
ccatccacaa gatgctggtt ggcctgaggtt ggtctggtt cctctctggt atctctctgccc 1320
gctcaggggtt gctctctggt tttgagagca tggagggggt cgtctgagccc cctcagggccc 1380
tgcagggggtt cctcagggag tgcctgaggtt aggtgagccc aggtctgagccc tgcctgagccc 1440
cctctctggtt tgcctgaggtt tgcctgaggtt aggtgagccc aggtctgagccc tgcctgagccc 1500
acaggggagc cgtctgaggtt cctctgaggtt tgcctgaggtt cgtctgaggtt tgcctgaggtt 1560
tgcctgaggtt cgtctgaggtt aggtgaggtt aggtgaggtt gctctgaggtt ggcagggccc 1620
ccacacacac ctgagggagc aggtgagggc tgcctgaggtt cgtctgaggtt cgtctgaggtt 1680
tctctctggt cgtctgaggtt aggtgaggtt aggtgaggtt cgtctgaggtt tgcctgaggtt 1740
gctcagggag atctcagggc tgcctgaggtt cctcagggc cctcagggc cctcagggc tgcctgaggtt 1800
tgcctgaggtt gctcagggc cctcagggc cctcagggc cctcagggc cctcagggc tgcctgaggtt 1860
tgcagggccc ctgagggag cagggggccc atctgaggtt aggtgaggtt cgtctgaggtt 1920
cctcagggc cagggggccc aggtgaggtt aggtgaggtt cgtctgaggtt cgtctgaggtt 1980
gtctcagggc tgcctgaggtt cgtctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2040
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2100
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2160
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2220
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2280
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2340
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2400
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2460
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2520
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2580
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2640
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2700
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2760
tgcctgaggtt tgcctgaggtt gctctgaggtt aggtgaggtt aggtgaggtt aggtgaggtt 2820
cctctctggt tgcctgaggtt accatctgga ggcggggccc gctgaggtt a 2871

```

<210> 25
 <211> 2141
 <213> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1734724CBI

```

<200> 25
gctcagggc cgggagggc agagggcagg gacgggaggc gacgggagc cgtcaggtccg 60
gctcagggc agagggcagg gacgggagc cctcaggtctt gcccaggtctt cgtcagggcagg 120
gctcagggc ggcggggcagg ggcgggagc tgggagggc cagggaggtc ggcctcagggc 180
cgtcaggtccg ggcggggcagg ggcgggagc tgggagggc ggcgggagc aggtcaggtccg 240

```

WO 01/77174

PCT/US01/1206

```

tdagcccccgg ccccaactccg accccgggagc ccaacgaccag cggccctcgcg ggcgggagggg 360
cgagccacagg cccctcccccg gttccctcggg cccctggggcg cccacggcgt cccgtctctgg 360
gaacagccggc tgaacccagg gctgaacgtg ttcgtggggcg cccctctgtgc atcaccatgc 420
tgggcccgggg ctgcaacggcg gaagctgaacc acttcggggcg gcaactccgt cggcccgctgg 480
cggcagctgct ggcagctctg caggttcggcg cccctcggcg tgcctggcctt cccgctggcc 540
cttgccttca agctggacga ggtggccggc ggggggtgct cctgtgtggc tgcctgtccg 600
ggcccaatct ctcaactctt atgcccctcg tggctggcg cgacatgaac ctgagagctg 660
ctctctctct ggcaactctcc tgggatgtag gtctctccca gcttcccaac cggagacttg 720
cagctctccc gttccacccc tactcaacat acaagaaaaa ggttaactgg ctgtttagt 780
caagctcgt tctgatttct gcaacttccc ttttcccgag cctcactctg accatctctt 840
ccacgcttct ggccctcgtc ttgatgcccc ttgtccctgt gatctacagc tgggatttga 900
tcaaaccccc tatcgttcgg tactacccc tagggacgtt gaactgtact ctcctcagca 960
ctctcatacc tatcgggttg ggcgtcttca ttgcctacaa ctacagccgg gtgctgact 1020
acattgtgaa ggtttccctg tggctctctgc tagtgactct ggtgtctctt ttcataatga 1080
cgggacttat gttaggacct gaactgtctg caagtatccc tgcagctgtt tatgtgatag 1140
caatttttat gcttllggca ggtacgctt caggttatgg tttagctact ctcttccatc 1200
ttccacccaa ctgcaaggag actgtatgtc tggaaacagg tagtcagact gtgcagctct 1260
gtacagccat totaaaaact gcttttccac cgcacttcat aggaagcatt tactgtttcc 1320
ctttgtgaa tgcacttttc cagttctgcg aagcggggat tttgtttta ctctataaaa 1380
tctcttataa aaacttaaaa gaaagggaaa tggcagacac ttctatggc acagtgaag 1440
cagaaatat aataatgatg gaaacccgct agactctct ctaaaagtgg agatcacag 1500
gagcttctat ctgtctgaaa tattgtctca tatttatggc ctgtgtgagt gcaactggtt 1560
caactaaag ataacactgg ttcaactaat acatgttaca attctgtact ttttaagttt 1620
caactgtgt ttaacccaaa gtttccaaa actcaaatc aatgctgtaa tataaatctg 1680
acctggaatg gctaaactga agctgaatt aaatctggt tttagtttt accatccca 1740
attcttatga ctgttcgaaa tacagaatct atcgaagaa aggtgtcttg aautgagaa 1800
ttttggcgca ctatgggaaa aaaaagctc tttttgaaa gcalaattga gtttaagtga 1860
atgttlytaa aaaaaaaagt gtgcttgcct tacttaaat tcttccaat gttgaatttt 1920
gaactgtatt cagaaagaat ccaaaacagg tcagttaaat aaggaatat agtatttctc 1980
aaacagtat caggaagaa ttacattaat gtatttgatt acttgatctg gtatttactt 2040
attaatgat aatcaacatt tttctagtga aaaaaaaa a 2100
2141

```

<210> 26

<211> 1903

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Inr: ID No: 1563237CB1

<400> 26

```

ccggctagtg agcggggggg caggaaaggtc agggcaagaa agggagaggg gacaggaggg 60
aagggtgggg caaagcgggg agaaagaggg gcccagcagt tgcctggggg aaggttgccg 120
agacccgggg gcaagaggtg agggctctga ggcggggaga ggaagggaga gaagagcccg 180
ggggggagcc gcccggagcc aggatgcggc cgcgcggcgc cgggggagag ccccgctgca 240
ccggggagcg ccagccctcg ctuocctcg gtagccgggg ccttcgacgg tggcgggggg 300
cggcgggcgg ggcctgtctg ctggctggaga tgcctggagc cgcgcgcttc ttccggctca 360
cgcctcaact cgtctgttat ctaaacagca caaacttcaa ctggacgggc gaggcggcga 420
cgcgcgcgcg gctgttatc ctggcgccct cctaactgat ggcggccgtg ggcggctggc 480
tggcggaagt gtaactgggc cgtacccggc cggtagcgct cagctgtctg ctctacatgg 540
cggcctcggg cctgtctgcc gccacccgct tcccagagg ccgcagctcc ctctcgggag 600
agatgcctcg gtcgcggctg ggaactgctt gccctcaggc cggctgcggc cgtctctctc 660
ccagccctca ctgcgggccg gtcctctcag cgggcttctt gctactgggc cggcgccgca 720
gctccgtcgg gaggcaactc acctccttgc gtcgcgacca ggttatgat ctgggctggc 780
acgcaacccg ccgtcttctc aactggtttt actggagcat caactgggtt gctgtgtctc 840
cgtctgtggt ggtggcattt attcagaga acaatagctt cctgtctggg taagagatcc 900
ctgtagctgt tggggggcgg gaatttttca tcttctctt tggcaacccc gtcctcaaca 960
caaaagccccc gttggggcagc caag-gtctc ctatgcttca gctcctctc caaaactgct 1020
gcccccagct gtcgcaacga caactcggca gaggccgcca atgtgcccgc gtgctggcgg 1080
acgagaggtc tcccagacca ggggtcttcc cgcagagga catcgcaac ttcagagctc 1140
tggltgaagat cttgcctcgc atggcgaccc tggctgctta ctggatggtc taacttcaga 1200

```

WO 01/77174

PCT/US01/11206

```

tgcagtcac  ctatgtctctg  cagggtcttc  acctccacat  cccaacatc  ttcccagcca  1260
accggggcaa  catctctgtg  gccctgagag  cccaggccag  cagctacacg  gactccctga  1320
gagcgagcgc  ctacactcac  atccacacaa  acgagacogt  gtcccagcag  attggggagg  1380
tccctgtacaa  cggggcaaca  ctgtccatct  ggtggcagat  cccctcagta  ctactcattg  1440
ggatccgtga  gatctttgcc  agcatccctg  gccctggagt  tgcctactca  gaggccccgc  1500
gtcccatgca  gggcgccctc  acgggcctct  tctctcgctt  gtgggggggt  ggtccattgt  1560
tgggtccag  ctgagtgaga  ctgctatctt  tggccggggg  ctgggtgac  tgcctccagg  1620
attttgggaa  catcaacaa  tggcggatgg  accctcaact  ctctctgtg  gctggcattc  1680
aggccgtcac  ggcctctcta  tttgtctgga  tgcctggacg  ctatgagagg  gcctccagg  1740
gccnagcctc  ccacagccgt  ttacgacggg  accagggctg  aacagggccc  attccagccc  1800
ccttctctca  ctctacggga  cagacggcag  cagtccagc  tctgggttca  ttctcgggtt  1860
attctgttag  aatgaatatg  tccataaat  aaggccatgg  cc  1920

```

<210> 27

<211> 1125

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No. 7475443C51

<400> 27

```

atggggaaaa  aacagtgcac  aaaggctaaa  aattccaaaa  accagatgc  ctctctctct  60
ccaaagggtc  acactctctc  gccgcaggg  gaacaaact  ggaaggga  tgaattgaaa  120
gaacagggtc  tcagaaagt  ggtgttaata  aactcclgca  agctaaagg  gaattgttta  180
accacatgta  aggaagccaa  gaaccttgaa  aaagggttag  ggaattgct  aactagaata  240
accagttag  agaaagaact  aaatgacct  atggagctga  aaacacagc  acgagacatt  300
cgtgatgct  acataagtat  cagtaccca  attgatcaag  cagaacaaag  gatatacagg  360
attgaagatc  aacttaata  aataaaggt  gaagacaga  ttgagaaaa  aatgaaaag  420
gatgaacaa  gccctccaa  aatattggac  tatgtgaaa  gacccaaact  acatttgatt  480
gggtgactct  gccgtctca  ttacagcatg  tgcgcgctc  tccctctcc  aagaatcag  540
ccgtgactgc  aggtcttgg  acgttgagcc  ctctggagg  tcaaatggt  cgtttgcagg  600
accagagctg  ggcagccgc  gcgtggtgta  aagataagta  taaaggsga  aggtatttct  660
gtctgtctag  tcttaggaac  ctggcaatgg  ttgctttgg  ccaggccgc  gaattgtcaa  720
cactggcctt  ttctacgttg  gatggcagcc  ctgtgggatg  tgcagccaa  aactggcct  780
tcccacatca  gccctacggg  tcagagaggg  aacgggggtc  cggagtctc  ctctatctc  840
cgaggcgtgc  ccaaggatth  cagcagggg  acatcagccc  aactaaagg  agccatgggg  900
ctggcccttg  agggcggagg  ttccagggc  ttctcccca  ggcacacat  gccatgaact  960
cccaacttcc  tgcuaaacc  cagctccagc  agcgcctgga  ttccctcca  gccaatgcc  1020
gtggccctgg  ccttptgca  gaagacctgc  gccctctgt  ggtctgctg  tctaggcacc  1080
tccctgtccc  ctgctgggg  aaggcccaag  attctctgg  aaacagtga  gctatgggtc  1140
gacaccttg  gccgggacut  gtataacact  gtgccaact  actcaggtc  tctcttgcg  1200
cagacgaact  ccaaggatgt  ggaatccagt  cgaagatcc  aggaagtag  tggcttggag  1260
ctggtgagg  agttctaga  ggaatggg  aacagctgc  ggaagaaat  cgaagcgtc  1320
caaatctgg  tgaagctgc  cgaaggggc  gaactgaac  accaatcca  tgaatccctg  1380
gtggaactg  ccttgggaat  tggcgtggg  atgtccctga  cgcagtcgg  cgtggaggtc  1440
ggcctgggga  tctcctgac  gcaatccgc  gtggagttg  gcgtgggat  gccaataag  1500
ctctccggcg  tggaggttgg  cctgggagtc  tccgtggg  agtcggcgct  gggagttggc  1560
gtggggatgt  cggtagcga  gtcagcgtg  ggaattggcg  tgggaggtc  cgtgacgag  1620
tccggcggtg  gatttggtg  ggggatgtc  gtgagcag  cggcgttgg  agttggcgtg  1680
gggatgtcgg  tgcacagtc  ctggggggcg  tcaagtccc  agcgcgcgc  cgcgggtgcl  1740
tgttagact  ctgagggcc  cccggcccg  gccctctgt  cctctcaact  ggcgcttct  1800
tcttctctc  tctccggcg  tccgggtgt  cgggcgcct  cgttggcct  caccgttcca  1860
ccccgtctg  actattaca  ctgggtctg  atcaacgaga  gggacggaa  gggcaacttc  1920
gtggagctgg  gcccgagtt  cctctggag  tccaatgctc  acttcagaa  cctgcccgtg  1980
aacacatcca  ccagcaggt  gcaatgccc  accaagctg  accaacaaga  cccagatatt  2040
tcaattggg  tcaactgtc  tgaagcttg  aatgtgtct  tctgtggag  cctccagat  2100
gaaccaagt  tgaactggg  aattttggc  agtgcactg  gattttcag  gatctatca  2160
ggtaataaat  ggaacactga  tgaagatgga  gtaattactt  ttgactggc  aaacccggcg  2220
tggtaaatc  aagctcctac  ttctccaaag  gacttagtg  ttttggtag  cgtgagcggc  2280
agtatgaagg  gctgaggat  gactattgct  aagcaacca  tcaacacat  cttgacaccc  2340
ctgggggaga  atgaattcat  taatatcata  gctacaaag  actacgtcca  ttacatcgg  2400

```

WO 01/77174

PCT/US01/11206

```

cctcgtttta aagggtatcct cgtccaggcg gaccgagaca atcgagagca tttoaaactg 2460
ctgggtggaggg agttgaggggt caaaggtgtg ggggtcgtgg accaagccct gagagaagcc 2550
ttccagatcc tgaagcaggtt caaaggagcc aagcaaggaa gctctctgca ccaggccatc 2640
atgctcatca ggcagggggt cgtggaggac tacgagccgg tgtttgagaa gbatnacttg 2730
ccagactgta aggtcggagt ttctacttac ctcatcgaga gagaagtatc ttttactgac 2820
cgcatgaaat ggtatgcatg caacaacaaa ggcatactaca cgcagatctc aacgtctggg 2910
gacacccagg agaacgtggt ggaatcacct cactgtgctca gcagcccat ggtcatcaac 3000
cagcaaccag acatcatctg gacggaggcc taccatggaa gaaagctcct cagctctgag 3090
gctcaggagcc tgacactgct caacactgtg gccatgcccag tcttcagcaa gaagaaggaa 3180
accgcatccc atggcattct cntgggtgtg gtagggtcag atgtggccct gagagagctc 3270
atgagctggg ggcocogtta caagcttgga gtgcacggat acgctcttct gaaccccaac 3360
aatggtatca tctctcccca tcccgaccct cggccctctg acagagaggg gaagaaacta 3450
aaacccaaac ctactacaaa cagtggtgat ctctccgaag tggagtgagg agaccaggct 3540
gaatctctga gaacagccat gatcaatag gaacaggtta ctctctggt ggtatgtgag 3630
gttccgatgg ataaaggga gggagttctt tctctgacca atgactact ctctcaggac 3720
atcagcgaca ccccttctag ttgggggttg gtcgtgtccc ggggcccagg agaatlacat 3810
ctctgggga aacgtctgtt ggaagaaggc ctgcatgact tgcttcacc agacttggcc 3900
ctggcgggtg atgggtatca ctgcatcaca gatattgacc cagaccacc gaagctcagc 4000
cagtgatagg ccagatccg ctctctcacc aggaaggacc cagacttggg gtgtgagag 4090
gggttggtcc ggaaggtgct gtttgacggc gtggtgacag ccccatgga agctactg 4180
acagcgctgc cctcaacat gtcagaggag tctgaacac tgggtggaat ggtctctctg 4270
ggcccccggg ctggccctct gagaagcagc ctgtctctga gctccgaga ggtctccgac 4360
aggaagtctc tgacacctga ggaagagccc agcctgtcca cctctggacc ctctccgctg 4450
tggtatccgc aggtctcaga gcatcctgct ggcagcttgc tcttcaact ccgtctggga 4540
gaagagacag gacgcccctc tgccaaaggc ctctccacac cactttgcca aaccatctc 4630
aagcgtcggt atggaataat gtcttgagac tgatggctgg ggaagccag gatccgggg 4720
atcctctctg gaccaggaca ggaaggga acagacggga ggaagctcac ccggccctct 4810
caggaaatcc gcttctctca cctgtgtgta tgcctctact cccactgtg ctggtctctc 4900
ccacacgtg caaatacta ctatgtctca caaaccaaaa aaaaaa 5000

```

<210> 28

<211> 2460

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7473438CB1

<400> 28

```

gggagggaggg gtagtgaaag gaaggaaagg aggaagggag gggagggagg 60
aggagggcag gggacggggg aggggagggg gggtaaaagg aaggagggaa ggacggaggg 120
aaggagggaa gactggcagc aggtcaccaac atgggttagt ggcgggaaca ggaactctgc 180
agggccgggg ttctgggctt ttctctagat ctacatattc acagacacac gaacccgggg 240
cagggaggtg ggaaggggac aggggagggg ctctctcggc cccacacagt ggcgatctcc 300
gcttcaaggt tcccggaagt ggcocggctt cgcctctcag aacccctgac cgtggggcgc 360
gggttgaggg gggcgagggc agcgtctcgg gcatcacagg caggttgaga ccgctggagg 420
taocgggggc gggggcccat ggcggggcgg agcctgcggc cggcccgact agtctctgac 480
accctgtgtt accctgcggt gggcgctgct gctctcagac cgtctcgact cgaaggggaa 540
agcgggcgcc agcagctgct ggtccagagc cggggcgctc tccggaggaa gttcggcttc 600
tcggccggag actaacggga gctggagggc ctggcgctcc aggtctgagc ccaccggcc 660
ggcgccagct ggaagtctcc gggctccttc tacttgcaca tcaccgltac cactaccatc 720
gagtagggcc agcggcgccc gggtagggac tccggcaagg ctctctcact gctctacgag 780
ctctgggga tcccgctgac gctgtcact tccagagacc tggggcaacg gctgaacgag 840
gtggtgcggc gctcctggtt ggcggccaag tctgtgctgg gcttggggc gactgtgctg 900
tccacggaga acctgggtgt ggcggggctg ctggcgcttg ccgccacctt ggcctcggg 960
gcccctgctt tctcgcactt ggaagggctg acctctcttc agcctacta ctactgtctc 1020
atacactca ccaacacggc ctctggcgac tccgtggcac tgcagagcgg cgaagggctg 1080
cagagggagc tccctcaagt ggtctcagc tctctcaaa tccctcagg gctcagggc 1140
atcgagcact tctcacaact ggtgtgtctg cgtctctctg tgcagagcgc agactggctc 1200
gagctgctcg cccgcacccc cagcccgccc cctccggggg ggcacaaag cagtgacctc 1260
tggtctctcc gcccgcggc cctgtccgtg ggtccgctct ctgtctctg cccagtgacc 1320
aagctggaga ggtggggcgg cgaacaactg ggtcttctgc cccctcggag cccgggggtc 1380

```


WO 01/77174

PCT/US01/11206

```

gtgcgtggcg ggcaggtctcc caggcttggg gcccggtgga agtccatctg acaacccccc 1440
ccaggccaggg gtcgaatctg gaataggagg gcttggcttc agctatcagg gcaacccccc 1500
cagggtattg aaacggatga cgggctctca ggcggctctc tgcacagagc agtttctcat 1560
taattgtctg ggttaagtc cctccctctt ttccaaaat atattacagt caaaccaaaa 1620
gcacaaaaca ggcctccagg gtaacctgtg gtagcaaat ccttgtagtc caaatgttat 1680
gaggccatgg ccacatcagc atttagagga ggcctgagac aggtccacct cagagccagc 1740
cctccagagc aactttctg tttgagag gctggctttc tgagttaacg tggaaatg 1800
caactttcaa cactggaatc ttccgtgcaa tttagtgggt cagatgtcca caatacaga 1860
agcatggggg acaaaaggtg accaaatggg accgtgtaca gtccagtgtt gacagggggg 1920
tcaatctctt gcgttaaac atctcaatct ctgccagtg tatagagtgg aatcagcccc 1980
tgtgtgaatg atggcctctt tccccagac agctgagctt ctttctctt ggcagcccc 2040
acttcaagat gaggaggggc ttgctgtcac ccaaccacca tctcaatggg gccacttgga 2100
cacagcagaa tgcctgggga cgaagctctt tccgggtctt gccaatcac agcaatata 2160
ccttggagaa tgcctatccc tcccagggtc tcaatttcca catctgtaga atgacgggt 2220
tcatcaagaa cctgtctccc ttttaagcat ctctggaggg acttctctga acagggccct 2280
gcaccaaaag ggggagggga caaagcccca ggtctggcat cctgggcccc cgttgccact 2340
agcactttcc cctggcttcc cagatagggc tgtctcaaca caaagcctga acagccagg 2400
ttcgaaaagg caaaaacttt aatttcaaca tgagctacat ggtcaatgaa acaatgcaac 2460

```

<210> 29

<211> 895

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7474286CB1

<400> 29

```

gggaaactga gtccttcacc cctttcaaga cccagggcgg cctctgctc cggccctcg 60
aggcccttcg ccggtctgc cctctccccc ttcccgaacc caccggccat agagtgatgt 120
ggcccaactt ctctctgcaa gaggagaacc ggcggggggg ggccgggggc cggcgccggc 180
cgcaacggga ggcagatcg ggctgaacgc cggagcgaga ggggaaggtg aagctggcgc 240
tgcctgtgag cgccttgagg gccacgtctg cgttctgttc cgtggccacc gagttctagg 300
tggagctcaa caactacbaa gccaaagcca ggcctgtgtg cgaagcggcc caactggggc 360
tgttgaagga gtgcaccaa cggctgttgc agggagacgt gcccgctggc agggabacct 420
gcggcccccg ggaagctgccc ggagaagcaa actgcaccta ttttaaatc ttcaacacgg 480
ggagagatgc acgcatcttt cagagaacca caaagaagaa ggtgaacttg gcagctggcg 540
tgatagcagt gctgggcttg gtagtcatgg ccttgggttg cctctgtatc atcatggtgc 600
tcagtaaaag tgcagagttc ctgctccagc ttggagccgt ctgctttggc ctctcaggcc 660
tgtctctctt ggttagcttg gaggctcttc gcaatccctg gaggcccttg ctgcagagag 720
tcagacccga gctctcccg gccacagcc tcaactcaga gtaactcttg tccctgggct 780
gcggcgtagg ggcgggacta atcctgtgtt tggggggcgg ctgcttcttg ctgctcaaac 840
tgacttctct gcccctggggg tccctctgtc ccaagcgggg gcaacgggac acctag 895

```

<310> 30

<311> 2050

<312> DNA

<313> Homo sapiens

<320>

<321> misc_feature

<323> Incyte ID No: 7472589CB1

<400> 30

```

gaaggaagcg ctcccccga ctcggctcgg ctccagagct cgaagccga cgcgcctagc 60
tcagctcccg gggcgggagc aggaactgcc gcaacagccc cacttaggag ggcctgattc 120
cgaaagcctc atgagagccc cccgggggct ctctccagcc ctgtctcccg cgtcccggtc 180
ctagagcggc gggaacaaag gccaccccgc ccggcagccc gcggcctgtc tggagagagc 240
tatagagccc ggggccccgc ggaacccggc caccggggcg cagggtccct agctgagagc 300
cccggcggcc gaggagcgcg ggttaaggac cgcggggagc ggcgaaggct cggggcgctc 360
agagagcgcc cctgcccgc ggcaacccgt gggccacggg actcgcggtt ggcctgagcg 420
ccggggggga ggcggagggc cccctctgtc cgggctcttg gaaggcgagc aggggtcttg 480

```

WO 01/77174

PCT/US01/11206

```

cgaaggcggc gaggggctcc gcggcgccgc cggacccctg gccaccatcc tcacgctccr 540
gctcccgcgg gggggatgtc gggggccggg ccccgagcgc cggcccggcc cgggggctga 600
gctccggacc atgtccctcc gcaagccccc gcccccggcc cggcgtagcc ggcggcgccr 660
gcgcgcgcgc tccctctgct gctgctgctg cggccgttcc caccctcaag aggaacccgg 720
cggctctagg ctgctggcgg cgtctctagg cctctaccct gctgggggag ccacagctct 780
ctcgggctcc gaggcccccg ggcagggcga ggcggggggc cgtctggggc ccacgctgag 840
caacttcagg gctggcgcag gcttgcccga gccagagctg cggccttcc ccgggcacta 900
cgaggcgccg ctggcgccgc ggttccgcgc cgaagcgctg cggcccgctt gggacttccc 960
cgycgccttc tacttcgctg gacccgtggt gtcaaccata gtgagggaag aagcccacc 1020
tcggcgctcc accccgggcc gctgtgcttc caacactggc cggctctgtg atctgacctt 1080
taagagttac atcaatatg ccaagaaca ggggcaccca gcaatacagc agagcttccc 1140
acgggttctc acagtgcttt cagagaacgc caaggggggt ttggcactga ccaccccgc 1200
gacggtgggc ggggaaggctt tctctatcgc ctacgggctg ttggcttcg ctgggacctt 1260
cctgttcttc aacctcttcc tggagcgcat catctgctg ctggccttca tcatsgggca 1320
ctgcggggag cggcagctgc ggcggcggc cctgtcgccc gccaccttc gccggggctc 1380
cgcctctctg gaggccgaca gcttgccggc ctgggaagcc tcggcgtaac acgtctgctt 1440
catctctggc ctgttcggcg tgcctctgct ctgctgcgcc tcggccatgt accccagctt 1500
ggagggtctg gactacgttg actcggctta cctctgcttc gtacactta gccacatcgg 1560
cttcggggac ctggtagaca gccagcaccg cgcctacccg aaccaggggc tcacacgctt 1620
gggcaacttc cctctcacc tgcctcggcg gtgctgcatt taactgctct caacgtctt 1680
ctccatcttc atcaagcagg tgcctcaact gatgctgcgc aagctgagct gccctgctg 1740
cgcccgctgc tgcgcggctc ctggcgccgc cctggccggc cggcaatgca tcaacccagg 1800
ctcccgctgc cgcgcgcgc tggcggcgtc cgttgccgac ccggcgccgc ggcagagga 1860
cgccgagggc cgcgcgcctc cggcgagctt catctccatg cgcgacctca cggcttccaa 1920
caaggtgtgc ctggcgctgc tgcagagcga gctgtcggag accgccaacg gctacccgc 1980
cagcgtgtgc gtcacccagg gccaggaagg cttctcgggc ggcgtggggc cgtggggcat 2040
catgaacacc cggctggcgg agaccagcgc ctccaggtag 2080

```

【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
18 October 2001 (18.10.2001)

PCT

(10) International Publication Number
WO 01/077174 A3(51) International Patent Classification: C07K 14/47,
C12N 15/12, 5/10, A01K 67/027, C07K 16/18, C12Q
1/68, A61K 38/16, 39/00

(21) International Application Number: PCT/US01/1206

(22) International Filing Date: 6 April 2001 (06.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/195,595 6 April 2000 (06.04.2000) US
60/196,872 12 April 2000 (12.04.2000) US
60/199,020 20 April 2000 (20.04.2000) US
60/200,553 28 April 2000 (28.04.2000) US
60/202,548 5 May 2000 (05.05.2000) US
60/203,495 11 May 2000 (11.05.2000) US

(63) Related by continuation (CON) or continuation-in-part

(CIP) to earlier applications:

US: 60/195,595 (CIP)
Filed on: 6 April 2000 (06.04.2000)
US: 60/196,872 (CIP)
Filed on: 12 April 2000 (12.04.2000)
US: 60/199,020 (CIP)
Filed on: 20 April 2000 (20.04.2000)
US: 60/200,553 (CIP)
Filed on: 28 April 2000 (28.04.2000)
US: 60/202,548 (CIP)
Filed on: 5 May 2000 (05.05.2000)
US: 60/203,495 (CIP)
Filed on: 11 May 2000 (11.05.2000)(71) Applicant (for all designated States except US): INCVTE
GENOMICS, INC., [US/US]: 3160 Porter Drive, Palo
Alto, CA 94301 (US)

(72) Inventors; and

(75) Inventors/Applicants (for US only): REDDY, Roopa
[IN/US]: 1233 W. McKinley Avenue, #3, Sunnyvale, CA
94086 (US); THORNTON, Michael [US/US]: 9 Midway
Road, Woodside, CA 94062-2612 (US); BORKOWSKY,
Mark, L. [US/US]: 122 Orchard Avenue, Redwood City,
CA 94061 (US); TANG, Y., Tom [US/US]: 1770 Ramoth
Court, San Jose, CA 95118 (US); KHAN, Faruk, A.
[IN/US]: 233 Lacey Avenue, #221, Mountain View, CA
94040 (US); TRIBOULEY, Catherine, M. [FR/US]:
1121 Tennessee Street, #5, San Francisco, CA 94107 (US);GANDHI, Ameena, R. [US/US]: 837 Noble Avenue,
#1, Menlo Park, CA 94025 (US); YAO, Monique, G.
[US/US]: 111 Frederick Court, Mountain View, CA 94043
(US); SANKARANARAYAN, Muthu, S. [US/US]: 710 Sylvia
Court, Los Altos, CA 94024 (US); BAUGHN, Mariah, R.
[US/US]: 14244 Santiago Road, San Leandro, CA 94577
(US); NGUYEN, Daniel, B. [US/US]: 1403 Ridgewood
Drive, San Jose, CA 95118 (US); POLICKY, Jennifer,
L. [US/US]: 1511 Jarvis Court, San Jose, CA 95128 (US);
YU, Henry [US/US]: 826 Lois Avenue, Sunnyvale, CA
94087 (US); SEELHAMER, Jeffrey, J. [US/US]: 12555
La Cresta, Los Altos Hills, CA 94022 (US); WALIA,
Narinder, K. [US/US]: 890 Davis Street, #205, San
Leandro, CA 94577 (US); LAL, Preeti [IN/US]: PO. Box
5149, Santa Clara, CA 95050 (US); KEARNEY, Liam
[IE/US]: 50 Woodside Avenue, San Francisco, CA 94117
(US); WALSH, Roderick, T. [IE/GB]: 8 Boundary Court,
St. Lawrence Road, Canterbury, Kent CT1 3EZ (GB);
LU, Deyang, Aina, M. [US/US]: 734 Coy Drive, San
Jose, CA 95131 (US); LU, Yan [CN/US]: 3885 Corcoran
Way, Palo Alto, CA 94303 (US); GREENE, Barrie, D.
[US/US]: 1332 10th Avenue, #104, San Francisco, CA
94122 (US); RAI MANN, Brigitte, E. [US/US]: 5801
South Drexler Avenue, #5B, Chicago, IL 60637 (US);
PATTERSON, Chandra [US/US]: 490 Sherwood Way,
#1, Menlo Park, CA 94025 (US).(74) Agents: HAMLET-COX, Donna et al.: Lucyle Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GR, GU, GF, GM, HN,
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, 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, SI, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

with international search report

(88) Date of publication of the international search report:
10 October 2001

[Continued on next page]

WO 01/077174 A3

(54) Title: HUMAN 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 01/077174 A3

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

【国際調査報告】

INTERNATIONAL SEARCH REPORT		National Application No. PCT/US 01/11206
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C07K14/47 C12N15/12 C12N5/10 A01K57/027 C07K16/18 C12Q1/68 A61K38/16 A61K39/00		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C07K C12N A61K C12Q A61K		
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, EMBL, CHEM ABS Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DATABASE EMBL [Online] 26 May 1999 (1999-05-26) KA0: "Homo sapiens mRNA for multidrug resistance-associated protein 7, partial" retrieved from EMBL, accession no. HSA238889 Database accession no. AJ238889 XP002192367 * See the sequence; standard codes for conversion of alternatives (unfinished sequencing; this fragment may be the first half of the claimed sequence) * --- -/--	1-19,22, 25-45,60
☒ Further documents are listed in the continuation of box C. ☐ Patent family members are listed in annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document relating to an oral disclosure, use, exhibition or other means "P" document published only to the international filing date but later than the priority date claimed "I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principles 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 can still be considered to involve an inventive step when the document is combined with one or more other such documents, each 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	
4 April 2002	17. 07. 2002	
Name and mailing address of the ISA European Patent Office, P.O. Box 5010, Patentstra�e 2 NL - 2250 MV Rijswijk tel (+31-70) 340-2040, Telex 31 651 eps nl, Fax: (+31-70) 340-3216	Authorized officer Korsner, S-E	

Form PCT/ISA/210 (second version) (July 1999)

INTERNATIONAL SEARCH REPORT

National Application No.
PCT/US 01/11206

C ₂ (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DATABASE EMBL [Online] 3 March 2000 (2000-03-03) OHARA ET AL: "Homo sapiens mRNA for FLJ000002 protein, partial cds" retrieved from EMBL, accession no. AK000002 Database accession no. AK000002 XP002192368 * See the NA sequence (99.9% overlap 27-1482 [of SA]: 371-1826) * -----	1-19,22, 25-45,60

Form PCT/ISA/210 (continuation of second sheet) (July 1993)

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 01/11206
Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)		
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:		
1. ☒	Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: see FURTHER INFORMATION sheet PCT/ISA/210	
2. ☒	Claims Nos.: 20-21, 23-24 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: see FURTHER INFORMATION sheet PCT/ISA/210	
3. ☐	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 invoice 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.: 1-19, 22, 25-44, 45, 60 (PART., AS APPLICABLE)	
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 01/11206

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 218

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-19, 22, 25-44, 45, 60 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by SEQ ID NOS 1 and 16, respectively, and dependent claims as long as they relate to said sequences.

2. Claims: 1-19, 22, 25-44, 46, 61 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by SEQ ID NOS 2 and 17, respectively, and dependent claims as long as they relate to said sequences.

3. Claims: 1-19, 22, 25-44, 47, 62 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by SEQ ID NOS 3 and 18, respectively, and dependent claims as long as they relate to said sequences.

4. Claims: 1-19, 22, 25-44, 48, 63 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by SEQ ID NOS 4 and 19, respectively, and dependent claims as long as they relate to said sequences.

5. Claims: 1-19, 22, 25-44, 49, 64 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by SEQ ID NOS 5 and 20, respectively, and dependent claims as long as they relate to said sequences.

6. Claims: 1-19, 22, 25-44, 50, 65 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by SEQ ID NOS 6 and 21, respectively, and dependent claims as long as they relate to said sequences.

7. Claims: 1-19, 22, 25-44, 51, 66 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by SEQ ID NOS 7 and 22, respectively, and dependent claims as long as they relate to said sequences.

8. Claims: 1-19, 22, 25-44, 52, 67 (partially, as applicable)

International Application No. PCT/US 01/11296

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

TRICH polypeptide and nucleic acid as specified by
SEQ ID NOS 8 and 23, respectively, and dependent
claims as long as they relate to said sequences.

9. Claims: 1-19, 22, 25-44, 53, 68 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by
SEQ ID NOS 9 and 24, respectively, and dependent
claims as long as they relate to said sequences.

10. Claims: 1-19, 22, 25-44, 54, 69 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by
SEQ ID NOS 10 and 25, respectively, and dependent
claims as long as they relate to said sequences.

11. Claims: 1-19, 22, 25-44, 55, 70 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by
SEQ ID NOS 11 and 26, respectively, and dependent
claims as long as they relate to said sequences.

12. Claims: 1-19, 22, 25-44, 56, 71 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by
SEQ ID NOS 12 and 27, respectively, and dependent
claims as long as they relate to said sequences.

13. Claims: 1-19, 22, 25-44, 57, 72 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by
SEQ ID NOS 13 and 28, respectively, and dependent
claims as long as they relate to said sequences.

14. Claims: 1-19, 22, 25-44, 58, 73 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by
SEQ ID NOS 14 and 29, respectively, and dependent
claims as long as they relate to said sequences.

15. Claims: 1-19, 22, 25-44, 59, 74 (partially, as applicable)

TRICH polypeptide and nucleic acid as specified by
SEQ ID NOS 15 and 30, respectively, and dependent
claims as long as they relate to said sequences.

International Application No. PCT/US 01/11206

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although Claims 32, 34 are directed to a diagnostic method practised on the body, the search has been carried out and based on the alleged effects.

Although Claim 18 is directed to a method of treatment of the body, the search has been carried out and based on the alleged effects.

Continuation of Box I.2

Claims Nos.: 20-21, 23-24

Present Claims 20-21, 23-24 relate to undefined compounds (and their use) as identified by screening methods. Since the compounds are not defined, no meaningful search can be carried out - and no factual support can be found in the Description.

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 I1 procedure.

フロントページの続き

(51)Int.Cl.⁷

F I

テーマコード (参考)

A 6 1 P 1/16
 A 6 1 P 1/18
 A 6 1 P 3/00
 A 6 1 P 3/06
 A 6 1 P 3/08
 A 6 1 P 3/10
 A 6 1 P 3/12
 A 6 1 P 5/14
 A 6 1 P 5/38
 A 6 1 P 7/00
 A 6 1 P 7/06
 A 6 1 P 9/00
 A 6 1 P 9/04
 A 6 1 P 9/06
 A 6 1 P 9/10
 A 6 1 P 9/12
 A 6 1 P 9/14
 A 6 1 P 11/00
 A 6 1 P 11/06
 A 6 1 P 13/02
 A 6 1 P 13/12
 A 6 1 P 17/00
 A 6 1 P 17/06
 A 6 1 P 19/06
 A 6 1 P 19/10
 A 6 1 P 21/00
 A 6 1 P 21/04
 A 6 1 P 25/00
 A 6 1 P 25/02
 A 6 1 P 25/06
 A 6 1 P 25/16
 A 6 1 P 25/18
 A 6 1 P 25/20
 A 6 1 P 25/22
 A 6 1 P 25/24
 A 6 1 P 25/28
 A 6 1 P 27/02
 A 6 1 P 27/12
 A 6 1 P 27/16
 A 6 1 P 29/00
 A 6 1 P 31/00
 A 6 1 P 31/04
 A 6 1 P 31/10
 A 6 1 P 31/12
 A 6 1 P 33/00
 A 6 1 P 35/00
 A 6 1 P 37/00

A 6 1 P 1/18
 A 6 1 P 3/00
 A 6 1 P 3/06
 A 6 1 P 3/08
 A 6 1 P 3/10
 A 6 1 P 3/12
 A 6 1 P 5/14
 A 6 1 P 5/38
 A 6 1 P 7/00
 A 6 1 P 7/06
 A 6 1 P 9/00
 A 6 1 P 9/04
 A 6 1 P 9/06
 A 6 1 P 9/10
 A 6 1 P 9/12
 A 6 1 P 9/14
 A 6 1 P 11/00
 A 6 1 P 11/06
 A 6 1 P 13/02
 A 6 1 P 13/12
 A 6 1 P 17/00
 A 6 1 P 17/06
 A 6 1 P 19/06
 A 6 1 P 19/10
 A 6 1 P 21/00
 A 6 1 P 21/04
 A 6 1 P 25/00
 A 6 1 P 25/00
 A 6 1 P 25/02
 A 6 1 P 25/06
 A 6 1 P 25/16
 A 6 1 P 25/18
 A 6 1 P 25/20
 A 6 1 P 25/22
 A 6 1 P 25/24
 A 6 1 P 25/28
 A 6 1 P 27/02
 A 6 1 P 27/12
 A 6 1 P 27/16
 A 6 1 P 29/00
 A 6 1 P 31/00
 A 6 1 P 31/04
 A 6 1 P 31/10
 A 6 1 P 31/12
 A 6 1 P 33/00
 A 6 1 P 35/00
 A 6 1 P 37/00

1 0 1

1 0 1

4 C 0 8 4
 4 C 0 8 5
 4 H 0 4 5

A 6 1 P	37/02	A 6 1 P	37/02
A 6 1 P	37/06	A 6 1 P	37/06
A 6 1 P	37/08	A 6 1 P	37/08
A 6 1 P	43/00	A 6 1 P	43/00
C 0 7 K	14/46	A 6 1 P	43/00
C 0 7 K	14/47	C 0 7 K	14/46
C 0 7 K	16/18	C 0 7 K	14/47
C 1 2 N	1/15	C 0 7 K	16/18
C 1 2 N	1/19	C 1 2 N	1/15
C 1 2 N	1/21	C 1 2 N	1/19
C 1 2 N	5/10	C 1 2 N	1/21
C 1 2 N	15/02	C 1 2 P	21/02
C 1 2 P	21/02	C 1 2 P	21/08
C 1 2 P	21/08	C 1 2 Q	1/02
C 1 2 Q	1/02	C 1 2 Q	1/68
C 1 2 Q	1/68	G 0 1 N	33/15
G 0 1 N	33/15	G 0 1 N	33/50
G 0 1 N	33/50	G 0 1 N	33/53
G 0 1 N	33/53	G 0 1 N	33/53
G 0 1 N	33/566	G 0 1 N	33/566
		C 1 2 N	5/00
		C 1 2 N	15/00
		C 1 2 N	5/00
		A 6 1 K	37/02

- (31)優先権主張番号 60/200,552
(32)優先日 平成12年4月28日(2000.4.28)
(33)優先権主張国 米国(US)
(31)優先権主張番号 60/202,348
(32)優先日 平成12年5月5日(2000.5.5)
(33)優先権主張国 米国(US)
(31)優先権主張番号 60/203,495
(32)優先日 平成12年5月11日(2000.5.11)
(33)優先権主張国 米国(US)

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CR,CU,CZ,DE,DK,DM,DZ,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,SD,SE,S G,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

- (72)発明者 ソートン、マイケル
アメリカ合衆国カリフォルニア州94062-2612・ウッドサイド・メッドウェイロード 9
(72)発明者 ボロースキー、マーク・エル
アメリカ合衆国カリフォルニア州94061・レッドウッドシティー・オーチャードアベニュー 122
(72)発明者 タング、ワイ・トム
アメリカ合衆国カリフォルニア州95118・サンノゼ・ランウィックコート 4230
(72)発明者 カーン、ファラ・エイ
アメリカ合衆国カリフォルニア州94040・マウンテンビュー・#221・エスケラアベニュー

3 3 3

- (72)発明者 トリボレー、キャサリン・エム
アメリカ合衆国カリフォルニア州9 4 1 0 7・サンフランシスコ・# 5・テネシーストリート 1
1 2 1
- (72)発明者 ガンディー、アミーナ・アール
アメリカ合衆国カリフォルニア州9 4 0 2 5・メンロパーク・# 1・ローブルアベニュー 8 3 7
- (72)発明者 ヤオ、モニーク・ジー
アメリカ合衆国カリフォルニア州9 4 0 4 3・マウンテンビュー・フレデリックコート 1 1 1
- (72)発明者 サンジャンワラ、マデュー・エス
アメリカ合衆国カリフォルニア州9 4 0 2 4・ロスアルトス・シルビアコート 2 1 0
- (72)発明者 ボーグン、マライア・アール
アメリカ合衆国カリフォルニア州9 4 5 7 7・サンレアンドロ・サンティアゴロード 1 4 2 4 4
- (72)発明者 ニュエン、ダニエル・ビー
アメリカ合衆国カリフォルニア州9 5 1 1 8・サンノゼ・リッジウッドドライブ 1 4 0 3
- (72)発明者 ポリッキー、ジェニファー・エル
アメリカ合衆国カリフォルニア州9 5 1 1 8・サンノゼ・ジャービスコート 1 5 1 1
- (72)発明者 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 シールハマー、ジェフリー・ジェイ
アメリカ合衆国カリフォルニア州9 4 0 2 2・ロスアルトス・ラクレスタ 1 2 5 5 5
- (72)発明者 ワライヤ、ナリンダー・ケイ
アメリカ合衆国カリフォルニア州9 4 5 7 7・サンレアンドロ・# 2 0 5・デービスストリート
8 9 0
- (72)発明者 ラル、ブリーティ
アメリカ合衆国カリフォルニア州9 5 0 5 6・サンタクララ・ピーオーボックス 5 1 4 2
- (72)発明者 キーニー、ライアム
アメリカ合衆国カリフォルニア州9 4 1 2 7・サンフランシスコ・ウッドサイドアベニュー 5 0
- (72)発明者 ウォルシュ、ロドリック・ティ
イギリス国ケント州・シーティー1 3 イージー・カンタベリー・セントローレンスロード・バ
ウンダリーコート 8
- (72)発明者 リュ、デュング・アイナ・エム
アメリカ合衆国カリフォルニア州9 5 1 2 3・サンノゼ・コイドライブ 2 3 3
- (72)発明者 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 グリーン、バリー・ディー
アメリカ合衆国カリフォルニア州9 4 1 2 2・サンフランシスコ・# 1 0 4・テンスアベニュー
1 3 3 2
- (72)発明者 ローマン、ブリジット・イー
アメリカ合衆国イリノイ州6 0 6 3 7・シカゴ・# 3 ビー・サウスドーチェスターアベニュー 5
8 0 1
- (72)発明者 パターソン、チャンドラ
アメリカ合衆国カリフォルニア州9 4 0 2 5・メンロパーク・# 1・シャーウッドウェイ 4 9 0

Fターム(参考) 2G045 AA25 AA40 BB03 BB20 CB01 CB21 DA12 DA13 DA14 DA36

DA77 FB02 FB03 FB04 FB07

4B024 AA01 AA11 BA53 BA80 CA04 CA09 DA02 EA02 EA04 GA03

GA11 HA04 HA12 HA17

4B063 QA01 QA05 QA20 QQ08 QQ20 QQ43 QR32 QR55 QR62 QS25

QS34

4B064 AG01 AG27 CA10 CA19 CA20 CC24 DA01 DA13

4B065 AA26X AA77X AA80X AA88X AA90X AB01 AB04 AC14 BA02 BA08

	CA24	CA25	CA44	CA46						
4C084	AA01	AA17	BA01	BA02	BA08	BA22	CA18	DC50	NA14	ZA011
	ZA021	ZA051	ZA061	ZA081	ZA121	ZA151	ZA161	ZA181	ZA201	ZA211
	ZA221	ZA241	ZA331	ZA341	ZA361	ZA371	ZA381	ZA421	ZA551	ZA591
	ZA661	ZA751	ZA811	ZA891	ZA941	ZA961	ZA971	ZB011	ZB051	ZB071
	ZB081	ZB131	ZB151	ZB261	ZB311	ZB331	ZB351	ZB381	ZB391	ZC061
	ZC081	ZC211	ZC311	ZC331	ZC351	ZC412	ZC551	ZC802		
4C085	HH01	HH03	HH11	HH13	KA03	KA04	KA26	KB82	LL03	LL07
	LL09	LL13	LL17	LL20						
4H045	AA10	AA11	AA20	AA30	BA10	BA41	CA40	DA75	DA76	EA20
	EA50	FA72	FA73	FA74						

专利名称(译)	转运蛋白和离子通道		
公开(公告)号	JP2004507218A	公开(公告)日	2004-03-11
申请号	JP2001575644	申请日	2001-04-06
[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	レディルーパ ソーントンマイケル ボロースキーマークエル タングワイトム カーンファラエイ トリボレーキャサリーンエム ガンディーアミーナアール ヤオモニークジー サンジャンワラマデューエス ボーグンマライアアール ニュエンダニエルビー ポリッキージェニファーエル ユエヘンリー シールハマージェフリージェイ ワライヤナリンダーケイ ラルプリーティ キーニーライアム ウォルシュロドリックティー リュデユングアイナエム リュヤン グリーンバリーディー ローマンブリジットイー パターソンチャンドラ		
发明人	レディ、ルーパ ソーントン、マイケル ボロースキー、マーク・エル タング、ワイ・トム カーン、ファラ・エイ トリボレー、キャサリーン・エム ガンディー、アミーナ・アール ヤオ、モニーク・ジー サンジャンワラ、マデュー・エス ボーグン、マライア・アール ニュエン、ダニエル・ビー ポリッキー、ジェニファー・エル ユエ、ヘンリー シールハマー、ジェフリー・ジェイ ワライヤ、ナリンダー・ケイ ラル、プリーティ キーニー、ライアム ウォルシュ、ロドリック・ティー リュ、デユング・アイナ・エム リュ、ヤン グリーン、バリー・ディー		

ローマン、ブリジット・イー
パターソン、チャンドラ

IPC分类号	G01N33/50 A61K38/00 A61K45/00 A61K49/00 A61P1/04 A61P1/16 A61P1/18 A61P3/00 A61P3/06 A61P3/08 A61P3/10 A61P3/12 A61P5/14 A61P5/38 A61P7/00 A61P7/06 A61P9/00 A61P9/04 A61P9/06 A61P9/10 A61P9/12 A61P9/14 A61P11/00 A61P11/06 A61P13/02 A61P13/12 A61P17/00 A61P17/06 A61P19/06 A61P19/10 A61P21/00 A61P21/04 A61P25/00 A61P25/02 A61P25/06 A61P25/16 A61P25/18 A61P25/20 A61P25/22 A61P25/24 A61P25/28 A61P27/02 A61P27/12 A61P27/16 A61P29/00 A61P31/00 A61P31/04 A61P31/10 A61P31/12 A61P33/00 A61P35/00 A61P37/00 A61P37/02 A61P37/06 A61P37/08 A61P43/00 C07K14/46 C07K14/47 C07K14/705 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/02 C12N15/09 C12P21/02 C12P21/08 C12Q1/02 C12Q1/68 G01N33/15 G01N33/53 G01N33/566
CPC分类号	A61P1/04 A61P1/16 A61P1/18 A61P3/00 A61P3/12 A61P11/00 A61P11/06 A61P13/02 A61P13/12 A61P17/00 A61P17/06 A61P19/06 A61P19/10 A61P21/00 A61P21/04 A61P25/00 A61P25/02 A61P25/06 A61P25/16 A61P25/18 A61P25/20 A61P25/22 A61P25/24 A61P25/28 A61P27/02 A61P27/12 A61P27/16 A61P29/00 A61P31/00 A61P31/04 A61P31/10 A61P31/12 A61P33/00 A61P35/00 C07K14/705
FI分类号	C12N15/00.ZNA.A A61K45/00 A61K49/00.A A61P1/04 A61P1/16 A61P1/18 A61P3/00 A61P3/06 A61P3/08 A61P3/10 A61P3/12 A61P5/14 A61P5/38 A61P7/00 A61P7/06 A61P9/00 A61P9/04 A61P9/06 A61P9/10 A61P9/12 A61P9/14 A61P11/00 A61P11/06 A61P13/02 A61P13/12 A61P17/00 A61P17/06 A61P19/06 A61P19/10 A61P21/00 A61P21/04 A61P25/00 A61P25/00.101 A61P25/02 A61P25/06 A61P25/16 A61P25/18 A61P25/20 A61P25/22 A61P25/24 A61P25/28 A61P27/02 A61P27/12 A61P27/16 A61P29/00.101 A61P31/00 A61P31/04 A61P31/10 A61P31/12 A61P33/00 A61P35/00 A61P37/00 A61P37/02 A61P37/06 A61P37/08 A61P43/00.105 A61P43/00.111 C07K14/46 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/02 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A C12N15/00.C C12N5/00.B 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/DA77 2G045/FB02 2G045/FB03 2G045/FB04 2G045/FB07 4B024/AA01 4B024/AA11 4B024/BA53 4B024/BA80 4B024/CA04 4B024/CA09 4B024/DA02 4B024/EA02 4B024/EA04 4B024/GA03 4B024/GA11 4B024/HA04 4B024/HA12 4B024/HA17 4B063/QA01 4B063/QA05 4B063/QA20 4B063/QQ08 4B063/QQ20 4B063/QQ43 4B063/QR32 4B063/QR55 4B063/QR62 4B063/QS25 4B063/QS34 4B064/AG01 4B064/AG27 4B064/CA10 4B064/CA19 4B064/CA20 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA26X 4B065/AA77X 4B065/AA80X 4B065/AA88X 4B065/AA90X 4B065/AB01 4B065/AB04 4B065/AC14 4B065/BA02 4B065/BA08 4B065/CA24 4B065/CA25 4B065/CA44 4B065/CA46 4C084/AA01 4C084/AA17 4C084/BA01 4C084/BA02 4C084/BA08 4C084/BA22 4C084/CA18 4C084/DC50 4C084/NA14 4C084/ZA011 4C084/ZA021 4C084/ZA051 4C084/ZA061 4C084/ZA081 4C084/ZA121 4C084/ZA151 4C084/ZA161 4C084/ZA181 4C084/ZA201 4C084/ZA211 4C084/ZA221 4C084/ZA241 4C084/ZA331 4C084/ZA341 4C084/ZA361 4C084/ZA371 4C084/ZA381 4C084/ZA421 4C084/ZA551 4C084/ZA591 4C084/ZA661 4C084/ZA751 4C084/ZA811 4C084/ZA891 4C084/ZA941 4C084/ZA961 4C084/ZA971 4C084/ZB011 4C084/ZB051 4C084/ZB071 4C084/ZB081 4C084/ZB131 4C084/ZB151 4C084/ZB261 4C084/ZB311 4C084/ZB331 4C084/ZB351 4C084/ZB381 4C084/ZB391 4C084/ZC061 4C084/ZC081 4C084/ZC211 4C084/ZC311 4C084/ZC331 4C084/ZC351 4C084/ZC412 4C084/ZC551 4C084/ZC802 4C085/HH01 4C085/HH03 4C085/HH11 4C085/HH13 4C085/KA03 4C085/KA04 4C085/KA26 4C085/KB82 4C085/LL03 4C085/LL07 4C085/LL09 4C085/LL13 4C085/LL17 4C085/LL20 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/BA41 4H045/CA40 4H045/DA75 4H045/DA76 4H045/EA20 4H045/EA50 4H045/FA72 4H045/FA73 4H045/FA74
優先権	60/195595 2000-04-06 US 60/196872 2000-04-12 US 60/199020 2000-04-20 US 60/200552 2000-04-28 US 60/202348 2000-05-05 US 60/203495 2000-05-11 US

摘要(译)

本发明提供了鉴定和编码TRICH的人转运蛋白和离子通道（TRICH）以及多核苷酸。 本发明还提供表达载体，宿主细胞，抗体，激动剂和拮抗剂。 本发明还提供了用于诊断，治疗或预防与TRICH异常表达有关的疾病的方法。

化合物 编号	化合物 SEQ ID NO:	化合物 编号	化合物 SEQ ID NO:	化合物 编号	化合物 SEQ ID NO:
1704775	1	1704775C01	16	1704775C01	16
7473034	2	7473034C01	17	7473034C01	17
1878581	3	1878581C01	18	1878581C01	18
2246292	4	2246292C01	19	2246292C01	19
5151730	5	5151730C01	20	5151730C01	20
7472594	6	7472594C01	21	7472594C01	21
7472536	7	7472536C01	22	7472536C01	22
7473422	8	7473422C01	23	7473422C01	23
2864715	9	2864715C01	24	2864715C01	24
1734724	10	1734724C01	25	1734724C01	25
1563237	11	1563237C01	26	1563237C01	26
7473443	12	7473443C01	27	7473443C01	27
7473438	13	7473438C01	28	7473438C01	28
7474286	14	7474286C01	29	7474286C01	29
7472589	15	7472589C01	30	7472589C01	30