

(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2004-500814

(P2004-500814A)

(43) 公表日 平成16年1月15日(2004.1.15)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A
A O 1 K 67/027	A O 1 K 67/027	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 P 1/00	4 B O 6 3
A 6 1 P 1/00	A 6 1 P 3/00	4 B O 6 4
		4 B O 6 5
	審査請求 未請求 予備審査請求 有	(全 266 頁) 最終頁に続く

(21) 出願番号	特願2001-547167 (P2001-547167)	(71) 出願人	301005050
(86) (22) 出願日	平成12年12月20日 (2000.12.20)		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6月24日 (2002.6.24)		アメリカ合衆国カリフォルニア州94304・パロアルト・ポータードライブ 3160
(86) 国際出願番号	PCT/US2000/035095	(74) 代理人	100089266
(87) 国際公開番号	W02001/046258		弁理士 大島 陽一
(87) 国際公開日	平成13年6月28日 (2001.6.28)	(72) 発明者	ボーグン、マライア・アール
(31) 優先権主張番号	60/172,000		アメリカ合衆国カリフォルニア州94577・サンレアンドロ・サンティアゴロード 14244
(32) 優先日	平成11年12月23日 (1999.12.23)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/176,083		
(32) 優先日	平成12年1月14日 (2000.1.14)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/177,332		
(32) 優先日	平成12年1月21日 (2000.1.21)		
(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)、及び T R I C H を識別しコードするポリヌクレオチドに関する。本発明はまた、発現ベクター及び宿主細胞、抗体、アゴニスト、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: 27 (SEQ ID NO: 1 - 27) からなる群から選択されたアミノ酸配列と、

(b) SEQ ID NO: 1 - 27 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% の配列同一性を有する天然のアミノ酸配列と、

(c) SEQ ID NO: 1 - 27 からなる群から選択されたアミノ酸配列の生物学的に活性な断片と、

(d) SEQ ID NO: 1 - 27 からなる群から選択されたアミノ酸配列の免疫原性断片とで構成され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単離されたポリペプチド。 10

【請求項 2】

SEQ ID NO: 1 - 27 からなる群から選択された請求項 1 の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 4】

請求項 2 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 5】

SEQ ID NO: 28 - 54 からなる群から選択された請求項 4 の単離されたポリヌクレオチド。 20

【請求項 6】

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

【請求項 7】

請求項 6 の組換えポリヌクレオチドで形質転換された細胞。

【請求項 8】

請求項 6 の組換えポリヌクレオチドを含む遺伝子組換え生物。

【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、 30

(a) 前記ポリペプチドの発現に好適な条件下で、請求項 1 のポリペプチドをコードするポリヌクレオチドに機能的に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。

【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、 40

(a) SEQ ID NO: 28 - 54 からなる群から選択されたポリヌクレオチド配列と、

(b) SEQ ID NO: 28 - 54 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% の配列同一性を有する天然のポリヌクレオチド配列と、

(c) 前記 (a) に相補的なポリヌクレオチド配列と、

(d) 前記 (b) に相補的なポリヌクレオチド配列と、

(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ポリヌクレオチド配列を含む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 50

れたポリヌクレオチド。

【請求項 13】

サンプルにおいて、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と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、存在する場合には随意選択で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。 10

【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】

サンプルにおいて、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、存在する場合には随意選択で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。 20

【請求項 16】

有効量の請求項 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 - 27 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 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特徴とする治療方法。 30

【請求項 19】

請求項 1 のポリペプチドのアゴニストとして効果的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 のスクリーニング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。 40

【請求項 21】

機能的な TRICH の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 20 の組成物を投与する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効果的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ンタゴニスト活性を検出するステップとを含むことを特徴 50

とするスクリーニング方法。

【請求項 2 3】

請求項 2 2 のスクリーニング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 2 4】

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 2 3 の組成物を投与することを含むことを特徴とする治療方法。

【請求項 2 5】

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、 10

(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 2 6】

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 20

(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、前記試験化合物の不在下での請求項 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特徴とするスクリーニング方法。

【請求項 2 7】

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

(a) 前記標的ポリヌクレオチドの発現に好適な条件下で、前記標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、 30

(b) 前記標的ポリヌクレオチドの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、前記化合物の不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 2 8】

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と、請求項 1 1 のポリヌクレオチドの少なくとも 2 0 の連続するヌクレオチドを含むプローブをハイブリダイズさせるステップであって、このハイブリダイゼーションが、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われ、前記標的ポリヌクレオチドが、請求項 1 1 のポリヌクレオチドのポリヌクレオチド配列またはその断片を含むポリヌクレオチドである、前記ステップと、 40

(c) ハイブリダイゼーション複合体の収量を定量するステップと、

(d) 前記処理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。

【発明の詳細な説明】

【0001】

(技術分野)

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

【0002】

(発明の背景)

真核細胞は、殆どの極性分子に高度に非浸透性であるような疎水性脂質二重層膜に囲まれ、この膜によって機能的に異なるオルガネラに細区画される。細胞及びオルガネラは、必須栄養素及び、 K^+ 、 NH_4^+ 、 P_i 、 SO_4^{2-} を含む金属イオン、糖、ビタミン及び種々の廃棄物を移出入するための輸送タンパク質を必要とする。輸送タンパク質は、抗生物質抵抗性、毒物分泌、イオンバランス、シナプス神経伝達、腎機能、腸管吸収、腫瘍成長及びその他の多様な細胞機能においても役割を果たす(Griffith, J. 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。

【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)。

【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(Pao, 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

10

20

30

40

50

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連している (Mueckler, M. (1994) Eur. J. Biochem. 219 : 713 - 725, Longo, N. and L. J. Ellis (1998) Adv. Pediatr. 45 : 293 - 313)。

10

20

30

40

50

【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. Haggstrom (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、イオンチャンネル)、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)。

【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關与する。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、多数の疾患状態に關連してきた (Danks, D. M. (1986) J. Med. Genet. 23 : 99 - 106)。

【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)。

【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)。

【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漏洩せしめるよう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

ヤネルがある。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。

【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。

10

【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(Forger, M. (1999) J. Biol. Chem. 274: 12951-12954)。

20

30

【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+} の移動を細胞への Na^+ の輸送にリンクさせる(アンチポート)。

40

【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

50

ニン依存性、及びグルタミン依存性カチオンチャンネル、並びにGABA依存性及びグリシン依存性クロライドチャンネル)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。特定のイオンチャンネルの開口特性(即ち開閉に対するその閾値及び耐性)は、補助チャン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。

【0016】

機械依存性または機械感受性イオンチャンネルは、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。伸長不活性化チャンネル(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チャンネルは、細胞膜への圧力またはストレスによって活性化され、Ca²⁺及びNa⁺を共に伝導するような1群のチャンネルに属している(スズキ, M. ら(1999) J. Biol. Chem. 274: 6330-6335)。

10

【0017】

電位依存性カチオンチャンネルのポア形成サブユニットは、イオンチャンネルタンパク質のスーパーファミリーを形成する。チャン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Ca²⁺サブファミリーでは、このドメインは4回繰り返され、Fizzチャンネルサブファミリーでは、各チャンネルは、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、チャンネルに対するイオン選択性を特定する情報を含む。K⁺チャンネルの場合、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)。

20

【0018】

電位依存性Na⁺及びK⁺チャンネルは、神経及び筋肉細胞などの電氣的興奮性細胞の機能に必要である。活動電位は、神経伝達物質放出及び筋収縮の原因となるものであるが、Na⁺及びK⁺イオンへの膜の浸透性における大きな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⁺チャンネルを開く。ナトリウムイオンは、細胞へ流れ、更に膜を脱分極してより多く電位依存性Na⁺チャンネルを開き、こ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ンネルを開く。結果として、カリウムイオンは外向きに流れ、それによって膜の再分極が生じる。電位依存性チャンネルは、第4膜貫通セグメント(S4)において荷電した残基を利用して電圧変化を検出する。開状態は、約1ミリ秒しか持続せず、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ン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ンネルのN末端によって媒介される。不活性化状態から閉状態へ遷移するには、静止電位に戻すことが必要である。

30

【0019】

電位依存性Na⁺チャンネルは、2つの更に小さな補助サブユニット1及び2に関連する260 kDaのポア形成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。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2及び1サブユニットとの関連は、チャンネルの機能的発現の増加、そのゲーティング特性の変化、及び膜表面面積の増加に起因する全細胞キャパシタンスの増加と関連がある(Isom, L. L. ら(1995) Cell 83: 433-442)。

40

【0020】

非電位依存性Na⁺チャンネルには、アミロライド感受性Na⁺チャンネル/degenerin(NaC/DEG)ファミリーのメンバーが含まれる。このファミリーのチャンネルサブユニットは、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⁺チャンネル(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

50

を引き起こす。NaC / DEGファミリーには、最近特徴付けられたH⁺ 依存性カチオンチャンネルまたは酸感受性イオンチャンネル (ASIC) もある。ASICサブユニットは、脳で発現し、ヘテロ多量体のNa⁺ 浸透性チャンネルを形成する。これらのチャンネルは、活性化のための酸性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ンネルのファミリーであるdegenerinとの相同性を示す。Degenerinでの突然変異は、神経変性を引き起こす。組織アシドーシスが痛みを発生させるものであるので、ASICサブユニットは、神経機能または痛みの知覚においても役割を有し得る (Waldmann, R. and M. Lazdunski (1998) *Curr. Opin. Neurobiol.* 8 : 418 - 424 ; Eglén, R. M. ら (1999) *Trends Pharmacol. Sci.* 20 : 337 - 342)。

【0021】

K⁺ チャンネルはすべての細胞種に存在し、電圧、ATP濃度、またはCa²⁺ 及びcAMPなどのセカンドメッセンジャーによってK⁺ チャンネルを制御し得る。非興奮性組織では、K⁺ チャンネルが、タンパク質合成、内分泌性分泌の制御及び膜内外の浸透平衡の維持に関与している。ニューロン及びその他の興奮性細胞では、活動電位の調整及び膜の再分極に加えて、K⁺ チャンネルが静止膜電位の設定の責任を有する。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ンネルが含まれる。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 (Na⁺ : K⁺ の割合は3 : 2である)。原形質膜のイオンチャンネルは、K⁺ 及びCl⁻ が受動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ンネルサブユニットはすべて6膜貫通 / 1ポア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チャンネルを形成する。これらのポア形成サブユニットは、チャンネル不活性化動力学を変化させるような種々の細胞質サブユニットとも関連している。Shaker様チャンネルファミリーには、QT延長症、心律動異常症候群に関連するようなhether-a-go-go関連遺伝子 (HERG) などの遅延整流型チャンネルや、電圧依存性Kチャンネルがある (Curran, M. E. (1998) *Curr. Opin. Biotechnol.* 9 : 565 - 572、Kaczarowski, G. J. and M. L. Garcia (1999) *Curr. Opin. Chem. Biol.* 3 : 448 - 458)。

【0023】

K⁺ チャンネルの第2のスーパーファミリーは、内向き整流チャンネル (Kir) で構成される。Kirチャンネルは、優先的にK⁺ の流れを内向きに導く特性を有する。これらのタンパク質は、1つのカリウム選択性ポアドメイン及び2つの膜貫通ドメインから構成され、これらは電位依存性Kチャンネルの第5及び第6膜貫通ドメインに対応する。Kirサブユニットは、四量体としても関連している。Kirファミリーには、腎尿細管疾患のパ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+} チャネルは精製されており、両チャ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。チャネルは、3つのサブユニットを含む。1サブユニットは、膜ポア及び電圧センサを形成し、2及び3サブユニットは、チャネルの電圧依存、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1、1つの2、4つの遺伝子によってコードされる。第4サブユニットは、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】

クロライドチャネルは、内分泌性分泌及びサイトゾルオルガネラのpH調整に必要である。分泌性上皮細胞では、 Cl^{-} は Na^{+} 、 K^{+} / Cl^{-} 共輸送体によって側底膜を通過して細胞に入り、電気化学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)。

【0027】

電位依存性クロライドチャネル (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)。

【0028】

リガンド依存性チャ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ネルに結合するときである。神経伝達物質依存性チャネルは、神経伝達物質が細胞外ドメインに結合すると開くチャネルである。これらのチャネルは、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ネルには、2つの型がある。ナトリウムチャネルは、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。クロライドチャネルは、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ネルは、4

10

20

30

40

50

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）。

【0029】

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A28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。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）。

10

【0030】

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、嗅覚に關与するcAMP依存性 Na^+ チャンネルと、視覚に關与するcGMP依存性カチオンチャンネルである。両システムは、Gタンパク質結合受容体のリガンド媒介活性化に關与し、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チャンネルはまた、 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チャンネルは、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ンネルを形成し得るサブユニットとチャンネル特性を調節する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^+ チャンネルに類似の、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ン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）。

20

30

【0031】

別のタイプのイオンチャン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。多くのチャンネルは、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ンネル活性を調節する。Kirチャンネルは、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ンネルの局在化に關与している。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ン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）。

40

【0032】

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ン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

50

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

【0033】

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

10

20

【0034】

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

30

【0035】

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(Calahán, M. D. and K. G. Chandry (1997) *Curr. Opin. Biotechnol.* 8 : 749 - 756)。

40

【0036】

新たな輸送体及びイオンチャネル及びこれらをコードするポリヌクレオチドの発見は、輸送障害、神経障害、筋肉障害及び免疫疾患の診断、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

50

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。

【0037】

(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」、「TRICH-15」、「TRICH-16」、「TRICH-17」、「TRICH-18」、「TRICH-19」、「TRICH-20」、「TRICH-21」、「TRICH-22」、「TRICH-23」、「TRICH-24」、「TRICH-25」、「TRICH-26」、および「TRICH-27」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乃至27 (SEQ ID NO: 1-27) からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-27 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択された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。別法では、SEQ ID NO: 1-27 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。

10

20

【0038】

更に本発明は、(a) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含む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: 1-27 からなる一群から選択されたポリペプチドをコードする。別法では、このポリヌクレオチドは、SEQ ID NO: 28-54 からなる一群から選択される。

30

【0039】

更に、本発明は、(a) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含む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。

40

【0040】

また、本発明は、(a) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-27 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-27 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含む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

50

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。

【0041】

更に、本発明は、(a)SEQ ID NO: 1 - 27からなる一群から選択されたアミノ酸配列と、(b)SEQ ID NO: 1 - 27からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c)SEQ ID NO: 1 - 27からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d)SEQ ID NO: 1 - 27とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含む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。

10

【0042】

更に、本発明は、(a)SEQ ID NO: 28 - 54からなる一群から選択されたポリヌクレオチド配列と、(b)SEQ ID NO: 28 - 54からなる一群から選択されたポリヌクレオチド配列と90%以上の配列同一性を有する天然のポリヌクレオチド配列と、(c)前記(a)に相補的なポリヌクレオチド配列と、(d)前記(b)に相補的なポリヌクレオチド配列と、(e)前記(a)乃至(d)のRNA等化物とで構成される一群から選択されたポリヌクレオチド配列を含む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。

【0043】

更に本発明は、(a)SEQ ID NO: 28 - 54からなる一群から選択されたポリヌクレオチド配列と、(b)SEQ ID NO: 28 - 54からなる一群から選択されたポリヌクレオチド配列と90%以上の配列同一性を有する天然のポリヌクレオチド配列と、(c)前記(a)に相補的なポリヌクレオチド配列と、(d)前記(b)に相補的なポリヌクレオチド配列と、(e)前記(a)乃至(d)のRNA等化物とで構成される一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配列を有する標的ポリヌクレオチドをサンプルにおいて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構成する少なくとも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、存在する場合には随意選択で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。

20

30

【0044】

更に本発明は、(a)SEQ ID NO: 28 - 54からなる一群から選択されたポリヌクレオチド配列と、(b)SEQ ID NO: 28 - 54からなる一群から選択されたポリヌクレオチド配列と90%以上の配列同一性を有する天然のポリヌクレオチド配列と、(c)前記(a)に相補的なポリヌクレオチド配列と、(d)前記(b)に相補的なポリヌクレオチド配列と、(e)前記(a)乃至(d)のRNA等化物とで構成される一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配列を有する標的ポリヌクレオチドをサンプルにおいて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、存在する場合には随意選択でその収量を測定するステップとを含む。

40

【0045】

更に本発明は、(a)SEQ ID NO: 1 - 27からなる一群から選択されたアミノ酸配列と、(b)SEQ ID NO: 1 - 27からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c)SEQ ID NO: 1 - 27からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d)SEQ

50

I D N O : 1 - 2 7 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含む効果的な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、S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列を含む組成物を提供する。更に、本発明は、患者にこの組成物を投与することを含む、機能的 T R I C H 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。

【 0 0 4 6 】

更に本発明は、(a) S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列と、(b) S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列と 9 0 % 以上の配列同一性を有する天然のアミノ酸配列と、(c) S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) S E Q I D N O : 1 - 2 7 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドのアゴニストとして効果的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 T R I C H 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。

【 0 0 4 7 】

更に、本発明は、(a) S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列と、(b) S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列と 9 0 % 以上の配列同一性を有する天然のアミノ酸配列と、(c) S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) S E Q I D N O : 1 - 2 7 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドのアンタゴニストとして効果的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 T R I C H 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。

【 0 0 4 8 】

更に本発明は、(a) S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列と、(b) S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列と 9 0 % 以上の配列同一性を有する天然のアミノ酸配列と、(c) S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) S E Q I D N O : 1 - 2 7 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含む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。

【 0 0 4 9 】

更に本発明は、(a) S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列と、(b) S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列と 9 0 % 以上の配列同一性を有する天然のアミノ酸配列と、(c) S E Q I D N O : 1 - 2 7 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) S E Q I D N O : 1 - 2 7 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含む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、この試験化合物の不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という特徴を有する。

【0050】

更に本発明は、SEQ ID NO: 28 - 54 からなる一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効果的な化合物をスクリーニングする方法であって、(a)この標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、(b)この標的ポリヌクレオチドの発現の変化を検出するステップとを含む、該スクリーニング方法を提供する。

10

【0051】

本発明はさらに、試験化合物の毒性を評価する方法を提供する。この方法は、(a)核酸を含む生体サンプルを前記試験化合物で処理するステップと、(b)処理した前記生体サンプルの核酸をプローブとハイブリダイズするステップと、(c)ハイブリダイゼーション複合体の収量を測定するステップと、(d)前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処理の生体サンプルにおけるハイブリダイゼーション複合体の収量とを比較するステップとを含み、前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量の差異が試験化合物の毒性を示唆する。この方法における前記プローブは、(1)SEQ ID NO: 28 - 54 からなる一群から選択されたポリヌクレオチド配列と、(2)SEQ ID NO: 28 - 54 からなる一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%の配列同一性を有する天然のポリヌクレオチド配列と、(3)前記(1)に相補的なポリヌクレオチド配列と、(4)前記(2)に相補的なポリヌクレオチド配列と、(5)前記(1)乃至(4)のRNA等価物とで構成される一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の連続する少なくとも20個の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: 28 - 54 からなる一群から選択されたポリヌクレオチド配列と、(2)SEQ ID NO: 28 - 54 からなる一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%の配列同一性を有する天然のポリヌクレオチド配列と、(3)前記(1)に相補的なポリヌクレオチド配列と、(4)前記(2)に相補的なポリヌクレオチド配列と、(5)前記(1)乃至(5)のRNA等価物とを含む。代替的に前記標的ポリヌクレオチドは前記ポリヌクレオチド配列の断片である。

20

30

【0052】

(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。

40

【0053】

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。

【0054】

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

50

する技術分野の一般的な技術者が普通に解釈する意味と同じである。本明細書で記述したものと類似、或いは同等の全ての装置及び材料、方法は本発明の実施及びテストに使用できるが、好適な装置及び材料、方法をここに記す。本明細書に記載の全ての文献は、本発明に関連して使用する可能性のある文献に記載された細胞系、プロトコル、試薬、ベクターを記述し開示するために引用した。従来の特許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。

【0055】

(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。

10

【0056】

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

【0057】

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。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。

20

【0058】

TRICH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、TRICHと同じポリペプチド或いはTRICH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、TRICH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TRICH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TRICH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TRICH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。

30

40

【0059】

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。

【0060】

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。

50

【0061】

用語「アンタゴニスト」は、TRICH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、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。

【0062】

用語「抗体」は、抗原決定基と結合可能なFab及びF(ab')₂、及びそれらの断片、Fv断片などの無傷の分子を指す。TRICH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。

10

【0063】

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。

20

【0064】

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(benzylphosphate)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。

30

【0065】

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。

【0066】

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'AGT3'」が相補的な配列「3'TCA5'」と対をなす。

40

【0067】

「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

50

溶液に展開され得る。

【0068】

「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。

10

【0069】

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
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

20

30

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

【0070】

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。

40

【0071】

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。

【0072】

50

「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。

【0073】

用語「断片」は、TRICHまたはTRICHをコードするポリヌクレオチドの固有の部分であって、その親配列 (parent sequence) と同一であるがその配列より長さが短いものを指す。「断片」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。

10

【0074】

SEQ ID NO: 28 - 54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: 28 - 54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: 28 - 54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: 28 - 54を関連ポリヌクレオチド配列から区別する類似の方法に有用である。ある断片と一致するSEQ ID NO: 28 - 54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。

20

【0075】

SEQ ID NO: 1 - 14のある断片は、SEQ ID NO: 15 - 28のある断片によってコードされる。SEQ ID NO: 1 - 14のある断片は、SEQ ID NO: 1 - 14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 - 14のある断片は、SEQ ID NO: 1 - 14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 - 14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。

30

【0076】

「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。

【0077】

「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。

【0078】

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。

40

【0079】

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.

50

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。

【0080】

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

【0081】

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

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。

【0082】

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。

【0083】

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。

【0084】

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。

10

【0085】

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。

【0086】

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

20

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。

30

【0087】

「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。

【0088】

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。

【0089】

「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

40

50

、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。

【0090】

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。

10

【0091】

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

20

【0092】

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、フィルター、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。

30

【0093】

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。

【0094】

「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。

【0095】

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GVRED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。

40

【0096】

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。

【0097】

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。

50

【0098】

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。

【0099】

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。

【0100】

「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。

10

【0101】

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

20

【0102】

TRICH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、TRICHの酵素環境に依存し、細胞の種類によって異なり得る。

【0103】

「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、TRICH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。

30

【0104】

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。

40

【0105】

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 Molecula

50

r 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。

【0106】

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

【0107】

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。

【0108】

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。

10

20

30

40

50

【0109】

「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。

【0110】

「レポーター分子」は、核酸、アミノ酸または抗体の標識に用いられる化学的または生化学的部分である。レポーター分子には、放射性核種、酵素、蛍光剤、化学発光剤、発色剤、基質、補助因子、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。

10

【0111】

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。

【0112】

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

20

【0113】

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。

【0114】

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。

30

【0115】

「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。

【0116】

用語「基板」は、任意の好適な固体或いは半固体の支持物を指し、膜及びフィルター、チップ、スライド、ウエハ、ファイバー、磁気または非磁気ビード、ゲル、チューブ、プレート、ポリマー、微小粒子、毛細管が含まれる。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。

40

【0117】

「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。

【0118】

「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

50

。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。

【0119】

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。

10

【0120】

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

20

30

【0121】

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、ある核酸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。

40

【0122】

（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、及び免疫疾患の診断、治療、及び予防に関する。

【0123】

表1は、本発明の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

50

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) に関連する。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 (Polypeptide 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 Polypeptide ID) の両方によって示されている。

【0124】

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。

【0125】

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

【0126】

表4に示されているように、本発明の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 : 28 - 54を同定するため、或いはSEQ ID NO : 28 - 54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、列5の配列に対応するcDNA配列およびゲノム配列の開始ヌクレオチド (5') 位置および終了ヌクレオチド (3') 位置を示す。

【0127】

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、6813453H1はインサイトcDNA配列の識別番号であり、ADRETUR01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、70207988V1) に由来する

10

20

30

40

50

。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、g1947104)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、GNNG6554406_006は、Genscan推定コード配列の識別番号であって、g6554406がGenscan分析によって得られたGenBankの配列の識別番号である。このGenscan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、"exon-stitching"アルゴリズムによってcDNAおよびGenscan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、"exon-stretching"アルゴリズムによってcDNAおよびGenscan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cDNA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cDNAの識別番号は示されていない。

10

20

30

40

50

【0128】

表5は、インサイトcDNA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cDNAライブラリを示す。代表的なcDNA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cDNA配列を最も多く含むインサイトcDNAライブラリのことである。表5に示されているcDNAライブラリを作製するために用いた組織およびベクターが表6に示されている。

【0129】

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

【0130】

本発明はまた、TRICHをコードするポリヌクレオチドを提供する。特定の実施例において、本発明は、TRICHをコードするSEQ ID NO: 28-54からなる一群から選択された配列を含むポリヌクレオチド配列を提供する。配列表に示したSEQ ID NO: 28-54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RNA配列を含む。

【0131】

本発明はまた、TRICH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、TRICH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、SEQ ID NO: 28-54からなる一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SEQ ID NO: 28-54からなる一群から選択された配列を含むポリヌクレオチド配列の変異配列を提供する。上記したポリヌクレオチド変異配列は何れも、TRICH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。

【0132】

遺伝暗号の縮重により作り出され得るTRICH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TRICH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。

【0133】

TRICH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TRICH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TRICH或いはその誘導体をコードするヌクレオチド配列を作るとは有利となり得る。特定のコドンが宿主によって利用される頻度に基づいてコドンを選択して、ペプチドの発現が特定の真核細胞または原核宿主に発生する割合を高めることが可能である。コードされたアミノ酸配列を変えないで、TRICH及びその誘導体をコードするヌクレオチド配列を実質的に変更する別の理由は、天然の配列から作られる転写物より例えば長い半減期など好ましい特性を備えるRNA転写物を作ることにある。

10

【0134】

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

【0135】

更に本発明には、種々のストリンジェントな条件下で、請求項に記載されたポリヌクレオチド配列、特に、SEQ ID NO: 28 - 54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

【0136】

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

30

40

【0137】

当分野で周知のPCR法をベースにした種々の方法で、部分的なヌクレオチド配列を利用して、TRICHをコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば制限部位PCR法を利用する1つの方法では、一般的

50

なプライマー及び入れ子プライマー (nested primer) を用いてクローニングベクター内のゲノムDNAから未知の配列を増幅する (例えば、Sarkar, G. (1993) PCR Methods Applic 2: 318 - 322 を参照)。逆PCR法を用いる別法では、広範な方向に伸長して環状化した鋳型から未知の配列を増幅するプライマーを用いる。この鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限断片に由来する (例えば、Triglia, T.ら (1988) Nucleic Acids Res 16: 8186 を参照)。キャプチャPCR法を用いる第3の方法は、ヒト及び酵母菌人工染色体DNAの既知の配列に隣接するDNA断片のPCR増幅を含む (例えば、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。この方法ではライブラリをスクリーニングする必要がなく、イントロン/エキソン接合部を探すのに有用である。全てのPCR法をベースにした方法では、プライマーは、市販のOLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) 或いは別の好適なプログラムなどを用いて、長さが22 ~ 30ヌクレオチド、GC含量が50%以上、約68 ~ 72の温度で鋳型に対してアニーリングするよう設計される。

【0138】

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)ライブラリが完全な長さの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。

【0139】

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

【0140】

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。

【0141】

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、スプライスバリアントの作製等が可能である。

【0142】

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。

【0143】

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。

【0144】

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)。

【0145】

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)。

10

【0146】

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)。

20

【0147】

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) *EMBO J.* 6:307-311; *The McGraw Hill Yearbook of Science and Technology* (1992) McGraw Hill, New York NY, pp. 191-196、Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659、Harri-
ngton, 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

30

40

50

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。

【0148】

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、PBLUESCRIPT (Stratagene, La Jolla CA) または pSPORT1 プラスミド (GIBCO BRL) などの多機能の大腸菌ベクターを用いることができる。ベクターの多数のクローニング部位にTRICHをコードする配列をライゲーションすると lacZ 遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。更に、これらのベクターを用いて、クローニングされた配列の *in vitro* での転写、ジデオキシンスクリーニング、ヘルパーファージによる一本鎖の救出、入れ子状態の欠失を作り出すことが可能である (例えば、Van Heeke, G. 及び S.M. Schuster (1989) J. Biol. Chem. 264: 5503 - 5509. を参照)。例えば、抗体の産生のためなどに多量の TRICH が必要な場合は、TRICH の発現をハイレベルで誘導するベクターが使用できる。例えば、強力に発現を誘発する T5 または T7 バクテリオファージプロモーターを含むベクターを使用できる。

10

【0149】

TRICH の発現に酵母の発現系の使用が可能である。因子やアルコールオキシダーゼや PGH プロモーターなどの構成型或いは誘導型のプロモーターを含む多種のベクターが、酵母菌 サッカロミセス - セレビジエ または *Pichia pastoris* に使用可能である。更に、このようなベクターは、発現したタンパク質の分泌か細胞内への保持のどちらかを誘導し、安定した増殖のために宿主ゲノムの中に外来配列を組み込む。(例えば、Ausubel, 1995, 前出、Bitter, G.A. ら (1987) Methods Enzymol. 153: 516 - 544、及び Scorer, C.A. ら (1994) Bio/Technology 12: 181 - 184. を参照)。

20

【0150】

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

30

【0151】

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

40

【0152】

50

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

【0153】

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

【0154】

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^r または apr^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 (cTRICHsulfuron)、ホスフィンオトリ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)、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

40

【0155】

マーカー遺伝子の発現の存在/不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。

【0156】

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

50

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。

【0157】

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)。

【0158】

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 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。

【0159】

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。

【0160】

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。

【0161】

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

【0162】

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。

30

【0163】

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。

【0164】

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

【0165】

50

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。

【0166】

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不在下での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。

【0167】

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「ロックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの作製に有用である(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ロックアウトする(Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. 他(1997) Nucleic Acids Res. 25:4323-4330)。形質転換したES細胞を同定し、例えばC57BL/6マウス系等から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を決め、これを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、潜在的な治療薬や毒性薬剤で検査することができる。

【0168】

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)。

【0169】

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「ロックイン」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。ロ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

10

20

30

40

50

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)。

【0170】

(治療)

TRICHのある領域と輸送体及びイオンチャネルのある領域との間に、例えば配列及びモチーフの文脈における化学的及び構造的類似性が存在する。従って、TRICHは、輸送障害、神経の疾患、筋疾患、及び免疫疾患においてある役割を果たすと考えられる。TRICH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、TRICHの発現または活性を低下させることが望ましい。また、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の発現または活性を増大させることが望ましい。

10

【0171】

従って、一実施例において、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、患者にTRICH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、及び免疫疾患が含まれ、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 ツース病、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、トゥレット病、妄想性精神病、及び分裂病などの輸送に関連した神経障害と、神経線維腫症、帯状疱疹後神経痛、3叉神経ニューロパシー、サルコイドーシス、鎌状赤血球性貧血、ウィルソン病、白内障、不妊症、肺動脈狭窄症、常染色体性感音性難聴、高ノ低血糖症、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、癩癧、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、クーラー及びクロイツフェルト ヤコブ病、ゲルストマン症候群、Gerstmann-Strausler-Scheinker症候群を含むプリオン病(prion disease)と、致死性家族性不眠症、神経系性栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜血管芽腫(cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、ダウン症を含む中枢神経系性精神薄弱及び他の発生障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィー及び他の神経筋障害、末梢神経疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性、及び中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分性及び不安性精神障害、及び妄想性精神病と、季節性の感情の障害(SAD)、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、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、及びトゥレット病が含まれ、筋疾患の中には、心筋

20

30

40

50

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、ミオクローヌス疾患、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的リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、膵炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれる。

10

20

【0172】

別の実施例では、限定するものではないが上に列記した疾患を含む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。

【0173】

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。

【0174】

更に別の実施例では、限定するものではないが上に列記した疾患を含むTRICHの発現または活性の低下に関連した疾患の治療または予防のために、TRICHの活性を調節するアゴニストを患者に投与することも可能である。

30

【0175】

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、及び免疫疾患が含まれる。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。

【0176】

別の実施例では、限定するものではないが上に列記した疾患を含むTRICHの発現または活性の増大に関連した疾患の治療または予防のために、TRICHをコードするポリヌクレオチドの相補配列を発現するベクターを患者に投与することも可能である。

40

【0177】

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

50

【0178】

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。

【0179】

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 (*bacilli Calmette - Guerin*) 及び *Corynebacterium parvum* が特に好ましい。

【0180】

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。

【0181】

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、Kohler, G. ら. (1975) *Nature* 256: 495 - 497; Kozbor, D. ら. (1985) *J. Immunol. Methods* 81 - 8 - 42; Cote, R. J. ら. (1983) *Proc. Natl. Acad. Sci.* 80: 2026 - 2030; Cole, S. P. ら. (1984) *Mol. Cell Biol.* 62: 109 - 120を参照）。

【0182】

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、Morrisson, S. L. 他. (1984) *Proc. Natl. Acad. Sci.* 81 - 4851 - 4855; Neuberger, M. S. 他. (1984) *Nature* 312: 604 - 608; Takeda, S. ら. (1985) *Nature* 314: 452, 454を参照）。別法では、当分野で周知の方法を用いて、一本鎖抗体の産生のための記載された技術を適用して、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ることもできる（例えば、Burton D. R. (1991) *Proc. Natl. Acad. Sci.* 88: 11120 - 3を参照）。

【0183】

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

10

20

30

40

50

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）。

【0184】

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、このような断片には、抗体分子のペプシン消化によって生成されるF(ab')に断片と、F(ab')に断片のジスルフィド架橋を減じることによって生成されるFab断片が含まれるが、これらに限定されるものではない。別法では、Fab発現ライブラリを作製することによって、所望の特異性とモノクローナルFab断片の迅速且つ容易な同定が可能となる（例えば、Huse, W.D. ら. (1989) Science 254: 1275 - 1281を参照）。

10

【0185】

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

20

【0186】

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Kaで表すが、このKa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Ka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Kaは、親和性の真の測定値を表す。Ka値が $10^9 \sim 10^{12}$ 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。Ka値が $10^6 \sim 10^7$ 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）。

30

【0187】

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1~2 mg/mlの特異的な抗体、好ましくは5~10 mg/mlの特異的な抗体を含むポリクローナル抗体医薬は一般に、TRICH抗体複合体を沈殿させなければならない処理に用いられる。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）。

40

【0188】

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ことができる。ある実施態様では、TRICHをコードする遺伝子のコーディング領域や調節領域に相補的な配列やアンチセンス分子（D

50

NA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。

【0189】

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ープ、及び当分野で周知のその他の系が含まれる(Rossi, J. J. (1995) *Br. Med. Bull.* 51(1):217-225; 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)。

【0190】

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; Poeschl, E. 他(1996) *Proc. Natl. Acad. Sci. USA.* 93:11395-11399)や、B型若しくはC型肝炎ウイルス(HBV、HCV)、Candida albicans及びParacoccidioides brasiliensis等の真菌寄生虫、Plasmodium falciparum及びTrypanosoma cruzi等の原虫寄生体)に対する防御機能を有するタンパク質を発現させて行うことができる。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。

【0191】

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 vitro 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トランスポゾン(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

【0192】

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、PCDNA3.1、EPITAG、PRCCMV2、PREP、PVAXベクター(Invitrogen, Carlsbad CA)、PCMV-SCRIPT、PCMV-TAG、PEGSH/PERV(Stratagene, La Jolla CA)、PTE-T-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 - アクチン遺伝子等)、(ii) 誘導性プロモーター(例えば、市販されているT-REXプラスミド(Invitrogen)に含まれている、テトラサイクリン調節性プロモーター(Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89:5547-5551; Gossen, M. 他(1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456)、エクジソン誘導性プロモーター(市販されているプラスミドPVGRXR及びPINDに含まれている: Invitrogen)、FK506/ラパマイシン誘導性プロモーター、またはRU486/ミフェプリスト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。

20

30

【0193】

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(Neumann, 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。

40

【0194】

本発明の別の実施例では、TRICHの発現に関連する遺伝子欠損によって起こる疾患や異常症は、(i) レトロウイルス末端反復配列(LTR)プロモーター若しくは独立したプロモーターのコントロール下でTRICHをコードするポリヌクレオチドと、(ii) 好適なRNAパッケージングシグナルと、(iii) 追加のレトロウイルス・シス作用性RNA配列及び効率的なベクターの増殖に必要なコーディング配列を伴うRev応答性エレメント(RRE)とからなるレトロウイルスベクターを作製して治療することができる。レトロウイルスベクター(例えば、PFB及びPFBNEO)はStratagene社から入手可能であり、公表データ(Riviere, I. 他.(1995) *Proc. Natl. Acad. Sci. U.S.A.* 92:6733-6737)に基

50

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)。

【0195】

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、Antinozzi, 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。

【0196】

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 I

CP22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。

【0197】

別法では、ウイルス(正の一本鎖RNAウイルス)ベクターを用いて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。プロトタイプのウイルスであるセムリキ森林熱ウイルス(Semliki Forest Virus, SFV)の生物学的な研究が広範に行われ、遺伝子伝達ベクター(gene transfer vector)がSFVゲノムに基づいていることが分かった(Garoff, H. and K. - J. Li (1998) *Cun. Opin. Biotech.* 9: 464 - 469)。

ウイルスRNAの複製中に、通常はウイルスカプシドタンパク質をコードするサブゲノムRNAが作り出される。このサブゲノムRNAが完全長のゲノムRNAより高いレベルで複製されるため、酵素活性(例えばプロテアーゼ及びポリメラーゼ)を有するウイルスタンパク質に対してカプシドタンパク質が過剰に産生される。同様に、TRICHをコードする配列をウイルスゲノムのカプシドをコードする領域に導入することによって、ベクター導入細胞において多数の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、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を導入することができる。ある集団における細胞のサブセットの特定の形質導入には、形質導入する前に細胞のソーティングを必要とする場合がある。ウイルスの感染性cDNAクローンの操作、ウイルスcDNA及びRNAのトランスフェクション、並びにウイルスの感染方法は当分野で周知である。

【0198】

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。三重式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、Geer, J. E. ら. (1994) *In: Huber, B. E. 及び B. I. Carr, Molecular and Immunological 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。

【0199】

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マーヘッド型リボザイム分子が含まれる。

【0200】

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、標的分子をスキヤニングすることによって初めに同定される。一度同定されると、切断部位を含む標的遺伝子の領域に対応する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。

【0201】

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。

10

【0202】

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-、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*)、及びワイプトシン(*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。

20

【0203】

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。

30

【0204】

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生化学系が含まれ得る。TRICH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、TRICH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

40

50

ヨンにより、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)。

【0205】

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)。

【0206】

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

【0207】

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、TRICH、TRICHの抗体、擬態、アゴニスト、アンタゴニスト、またはTRICH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。

【0208】

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

【0209】

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。

【0210】

本発明に用いる好適な組成物には、目的を達成するため、効果的な量の活性処方成分を含

む組成物が含まれる。当業者は、十分に自身の能力で効果的な服用量を定めることができる。

【0211】

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

【0212】

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。

【0213】

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

【0214】

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。

30

【0215】

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

【0216】

(診断)

別の実施例では、TRICH に特異的に結合する抗体が、TRICH の発現によって特徴付けられる疾患の診断、または TRICH や TRICH のアゴニストまたはアンタゴニスト、インヒビターで治療を受けている患者をモニターするためのアッセイに用いられる。診断に有用な抗体は、治療のところで記載した方法と同じ方法で製剤される。TRICH の診断アッセイには、抗体及び標識を用いてヒトの体液或いは細胞や組織から採取された

50

ものからTRICH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。

【0217】

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

10

【0218】

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

20

【0219】

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

30

【0220】

プローブはまた、関連する配列の検出に利用され、TRICHをコードする任意の配列と少なくとも50%の配列同一性を有し得る。目的の本発明のハイブリダイゼーションプローブには、DNAあるいはRNAが可能であり、SEQ ID NO: 28 - 54の配列、或いはTRICH遺伝子のプロモーター、エンハンサー、イントロンを含むゲノム配列に由来し得る。

【0221】

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

40

【0222】

TRICHをコードするポリヌクレオチド配列を用いて、TRICHの発現に関連する疾患を診断することが可能である。限定するものではないが、このような疾患には輸送障害、神経の疾患、筋疾患、及び免疫疾患が含まれ、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

50

ー マリー ツース病、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 10
 、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、筋萎縮性側策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 20
 中枢神経系疾患と、クールー及びクローイツフェルト ヤコブ病、ゲルストマン症候群、Gerstmann-Straussler-Scheinker症候群を含むプリオン病 (prion disease) と、致死性家族性不眠症、神経系性栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜血管芽腫 (cerebelloretinal hemangioblastomatosis)、脳3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 30
 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、ミオクローヌス疾患、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 40
 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、膝炎、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 50

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、外傷が含まれる。T R I C Hをコードするポリヌクレオチド配列は、サザーン法やノーザン法、ドットプロット法、或いはその他の膜系の技術、P C R法、ディップスティック (d i p s t i c k)、ピン (p i n)、E L I S A式アッセイ、及び変異T R I C H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。

【0223】

ある実施態様では、T R I C H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。T R I C H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、制御サンプルと比べて著しく変わっている場合は、サンプル内のT R I C H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。

10

【0224】

T R I C H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、T R I C H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。

20

【0225】

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ペースで繰り返して、被験者における発現のレベルが正常な患者に示される値に近づき始めたかどうかを推定することが可能である。繰り返し行ったアッセイの結果を、数日から数ヶ月の期間の治療の効果を見るのに用いることができる。

30

【0226】

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。

【0227】

T R I C H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、P C R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 *i n v i t r o* で生成され得る。オリゴマーは、好ましくはT R I C Hをコードするポリヌクレオチドの断片、或いはT R I C H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D N A或いはR N A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。

40

【0228】

或る実施態様において、T R I C H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 (S N P) を検出し得る。S N P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、S N Pの検出方法には、一本鎖立体構造多型 (S S C P) 及び蛍光S S C P (f S S C P) 法が含まれる。S S C Pでは、T R I C Hをコードする

50

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。fSCCP 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 シークエ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 (Chromatogram) の自動分析及び統計モデルを用いたシークエンシングエラーや研究室での DNA の調整に起因する配列のばらつきを排除する。別法では、例えばハイスループットの MASSARRAY システム (Sequenom, Inc., San Diego CA) を用いた質量分析により SNP を検出し、特徴付ける。

10

20

30

40

50

【0229】

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。

【0230】

更に別の実施例では、本明細書に記載した任意のポリヌクレオチド配列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。

【0231】

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。

【0232】

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。

【0233】

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。

【0234】

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

【0235】

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。

【0236】

別の実施例は、本発明のポリペプチド配列を用いて組織または細胞型のプロテオームを分析することに関連する。「プロテオーム」という用語は、或る特定の組織または細胞型におけるタンパク質発現の包括的パターンを指す。プロテオームを構成する各タンパク質は、個々に更なる分析をすることができる。プロテオーム発現パターン即ちプロファイルは、所定の条件下で所定の時間に発現したタンパク質の数及びそれらの相対的な存在量を定量することにより分析する。従って、ある細胞のプロテオームのプロファイルは、特定の組織または細胞型のポリペプチドを分離及び分析することにより作成し得る。或る実施例では、このような分離は2次元ゲル電気泳動によって行う。この2次元ゲル電気泳動法では、まず、1次元の等電点電気泳動によりサンプルからタンパク質を分離し、次に、2次元のドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に従って分離する (前出の

10

20

30

40

50

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

【0237】

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、Mendozze, 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。 20

【0238】

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。 30

【0239】

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。

【0240】

別の実施例では、試験化合物の毒性は、タンパク質を含む生体サンプルをその試験化合物で処理することにより評価する。生体サンプルから得たタンパク質を、本発明のポリペプチドに特異的な抗体と共にインキュベートする。その抗体により認識されたタンパク質の量を定量する。処理された生体サンプル中のタンパク質の量を、未処理生体サンプル中のタンパク質の量と比較する。両サンプルのタンパク質量の差が、処理サンプル中の試験化合物に対する毒性反応を示唆する。 40

【0241】

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 50

- 10619; Baldeschweiler 他(1995) PCT出願番号WO95/251116; Shalon, D. 他(1995) PCT出願番号WO95/35505; Heller, R.A. 他(1997) Proc. Natl. Acad. Sci. 94:2150-2155; 及び Heller, M.J. 他(1997) 米国特許第5,605,662号を参照)。様々なタイプのマイクロアレイが周知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, Londonに記載されている。また、この文献を引用することを以って本明細書の一部とする。

【0242】

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)の遺伝とが相関するような遺伝子連鎖地図を作成可能である(Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照)。

【0243】

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付けることができる。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。

【0244】

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

【0245】

本発明の別の実施例では、TRICH、その触媒作用断片或いは免疫原断片またはそのオ

10

20

30

40

50

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。

【0246】

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 8 4 / 0 3 5 6 4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。

10

【0247】

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。

【0248】

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。

20

【0249】

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、以下に記載する特定の好適な実施例は、例示目的であって本発明を限定するものではない。

【0250】

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60/172,000号、同第60/176,083号、同第60/177,332号、同第60/178,572号、同第60/179,758号、および同第60/181,625号、に言及することをもって本明細書の一部とする。

30

【0251】

(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。まず、組織の一部をホモジナイズしてグアニジニウムイソチオシアネート溶液に溶解する一方、この組織の別の一部をホモジナイズしてフェノールに溶解するか、或いはT R I Z O L (L i f e T e c h n o l o g i e s)、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などの好適な変性剤の混合液に溶解した。この溶解物を塩化セシウムにおいて遠心分離によって、或いはクロロホルムで抽出した。イソプロパノール或いは酢酸ナトリウムのどちらかとエタノール、或いは別の方法でこの溶解物からRNAを沈殿させた。

40

【0252】

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(P r o m e g a)またはO L I G O T E X ラテックス粒子(Q I A G E N , V a l e n c i a C A)、O L I G O T E X mRNA精製キット(Q I A G E N)を用いてポリ(A+)RNAを単離した。別法では、P O L Y (A) P U

50

RE 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。

【0253】

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した。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つの制限酵素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 S 1000またはSEPHAROSE CL 2B、SEPHAROSE CL 4Bカラムクロマトグラフィー(Amersham Pharmacia Biotech)、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)またはpSPORT1プラスミド(Life Technologies)、pcDNA2.1プラスミド(Invitrogen Carlsbad CA)、pINCYプラスミド(Incyte Pharmaceuticals, Palo Alto CA)などの好適なプラスミドのポリリンカーの適合性制限酵素部位にcDNAを結合させた。この組換えプラスミドを、Stratagene社のXL1-Blue, XL1-BlueMRF、SOLR、またはLife Technologies社のDH5 またはDH10B、ELECTROMAX DH10Bを含むコンピテント大腸菌細胞に導入し組み込んだ。

【0254】

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を、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。MagicまたはWIZARD Minipreps DNA精製システム(Promega)、及びAGTC Miniprep精製キット(Edge Biosystems, Gaithersburg MD)、QIAGEN社のQIAWELL 8 Plus Plasmid、QIAWELL 8 Plus Plasmid、QIAWELL 8 Ultra Plasmid 精製システム、REAL Prep 96プラスミドキットの内の少なくとも1つを用いてプラスミドを精製した。沈殿させた後、0.1mlの蒸留水に再懸濁して、凍結乾燥して或いは凍結乾燥しないで4℃で保管した。

【0255】

別法では、ハイスループットの直接結合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過程を単一反応混合液で行った。サンプルを処理してから384-ウェルプレートに移して保管し、増幅したプラスミドDNAの濃度をPICOGREEN色素(Molecular Probes, Eugene OR)及びFluoroskan II蛍光スキャナ(Labsystems Oy, Helsinki, Finland)を用いて蛍光定量的に測定した。

【0256】

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のシーケンシング反応は、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シーケンシングキットに含まれる試薬を用いて行った。cDNAシーケ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シーケンシングシステム(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シーケンシングシステム(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。

【0257】

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、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cDNA群をスクリーニングし、オープンリーディングフレームを決定した。これらの完全長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。次に、これらのポリペプチド配列をGenBankタンパク質データベース(genpept)、SwissProt、BLOCKS、PRINTS、DOMO、PRODOM、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 Software Engineering, South San Francisco CA)およびLASERGENEソフトウェア(DNASTAR)を用いて分析した。ポリヌクレオチド配列およびポリペプチド配列のアラインメントを、アラインメントした配列間のパーセント同一性を計算する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。

【0258】

表7は、インサイトcDNAの組み立ておよび完全長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）。

【0259】

完全長ポリヌクレオチド配列およびポリペプチド配列の組み立て及び分析に用いる上記のプログラムは、SEQ ID NO: 28 - 54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。

【0260】

4 ゲノムDNA由来のコード配列の同定および編集

推定輸送体及びイオンチャネルは、公共のゲノム配列データベース（例えば、gbpriやgbhtg）においてGenscan遺伝子同定プログラムを実行して初めに同定された。Genscanは、様々な生物に由来するゲノム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gbpri公共データベースの配列と比較した。必要に応じて、Genscan推定cDNA配列を、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を見つけて出すことにより、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。代わりに、ポリヌクレオチド配列は、その全体が編集された或いは未編集のGenscan推定コード配列に由来する、完全長コード領域である。

【0261】

5 ゲノム配列データとcDNA配列データとの組み立て

ステッチ配列 (Stitched Sequence)

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

10

20

30

40

50

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 genepept 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、genepept 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。

【0262】

10

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 あるいは 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。

20

【0263】

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: 28 - 54 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: 28 - 54 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。

30

【0264】

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。

40

【0265】

7 ポリヌクレオチド発現の分析

50

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）。

【0266】

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、

10

【数1】

(BLASTスコア×配列一致率)

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

として定義される積スコアである。積スコアは、0～100の標準化された値であり、以下のように求める。BLASTスコアにヌクレオチド配列の一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除する。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%の一致で一端が70%重畳しているか、或いは88%一致で他端が100%重畳しているかの何れかの場合である。積スコア50は、100%の一致で一端が50%重畳しているか、或いは79%の一致で他端が100%重畳しているかの何れかの場合である。

20

【0267】

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、膵臓、呼吸器系、感覚器官、皮膚、顎口腔系、分類不能/混合、または尿管などの1つの生物/組織のカテゴリーに分類される。各カテゴリーにおけるライブラリの数をカウントし、その合計数を全カテゴリーのライブラリ数で除す。同様に、各ヒト組織は、癌、細胞系、発生、炎症、神経、外傷、心血管、プール(pool)などの1つの疾患/症状のカテゴリーに分類され、各カテゴリーにおけるライブラリの数をカウントし、その合計数を全カテゴリーのライブラリ数で除す。得られるパーセンテージは、TRICHをコードするcDNAの疾患特異的な発現を反映する。cDNA配列およびcDNAライブラリ/組織の情報は、LIFESEQ GOLD データベース (Incyte Genomics, Palo Alto CA) から得ることができる。

30

40

【0268】

8 TRICHをコードするポリヌクレオチドの伸長

完全長のポリヌクレオチド配列は、完全長分子の好適な断片から設計したオリゴヌクレオチドプライマーを用いてその完全長分子の好適な断片を伸長して作製した。一方のプライマーは既知の断片の5'の伸長を開始するために合成し、他方のプライマーは既知の断片の3'の伸長を開始するために合成した。開始プライマーは、OLIGO 4.06ソフトウェア (National Biosciences) 或いは他の適切なプログラムを

50

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。

【0269】

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

【0270】

当分野で既知の方法を利用したPCR法で高い忠実度で増幅した。PCRはPTC-200 thermal cycler (MJ Research, Inc.)用いて96ウェルブロックプレートで行った。反応混合液は、鋳型DNA及び200 nmolの各プライマー、 Mg^{2+} と $(NH_4)_2SO_4$ とβ-メルカプトエタノールを含むバッファー、Taq DNAポリメラーゼ (Amersham Pharmacia Biotech)、ELONGASE酵素 (Life Technologies)、Pfu DNAポリメラーゼ (Stratagene)を含む。プライマーの組、PCI AとPCI Bに対して以下のパラメーターで増幅を行った。

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ーで増幅を行った。

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。

【0271】

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PICO GREEN定量試薬(0.25% (v/v) PICO GREEN; Molecular Probes, Eugene OR)を不透明な蛍光光度計プレート(Coming Costar, Acton MA)の各ウェルに分配してDNAが試薬と結合できるようにして測定する。このプレートをFluoroskan II (Labsystems Oy, Helsinki, Finland)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 µl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。

【0272】

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ラーゼ(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)を用いて増幅した。

t a g e n e) で制限部位の伸び出しを処理してコンピテント大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りとってLB / 2 X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。

【0273】

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。

10

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

20

【0274】

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。

【0275】

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: 28 - 54 から導き出されたハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記すが、より大きなcDNAフラグメントの場合でも基本的に同じ手順を用いる。オリゴヌクレオチドを、OLIGO4.06ソフトウェア (National Bioscience) のような最新式のソフトウェアを用いてデザインし、50 pmolの各オリゴマーと、250 μ Ciの[³²P]アデノシン三リン酸 (Amersham, Chicago, IL) 及びT4ポリヌクレオチドキナーゼ (DuPont NEN, Boston MA) とを組み合わせ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10⁷ カウントの標識されたプローブを含むアリコットを、次のエンドヌクレアーゼ、Ase I、Bgl II、Eco RI、Pst I、Xba I 或いはPvu II (DuPont NEN) の1つを用いて切断したヒトゲノムDNAの典型的な膜ベースのハイブリダイゼーション解析において用いる。

30

40

【0276】

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

【0277】

50

10 マイクロアレイ

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

【0278】

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。

【0279】

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

【0280】

マイクロアレイの準備

10

20

30

40

50

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 ~ 2 ng の初期量から 5 µg を超える最終量までアレイエレメントを増幅する。増幅されたアレイエレメントは、SEPHACRYL - 400 (Amersham Pharmacia Biotech) を用いて精製する。

【0281】

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 °C の天火で硬化させる。

10

【0282】

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

20

【0283】

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

【0284】

ハイブリダイゼーション

ハイブリダイゼーション反応液は、5 × SSC、0.2% SDS ハイブリダイゼーション緩衝液に Cy3 及び Cy5 標識した cDNA 合成産物を各 0.2 µg 含む 9 µl のサンプル混合体を含めたものである。サンプル混合液を、65 °C で 5 分間加熱し、マイクロアレイ表面上に一定量分注してから 1.8 cm² のカバーガラスで覆う。このアレイを、顕微鏡スライドより僅かに大きいキャビティを有する防水チャンバーに移す。チャンバーの角に 140 µl の 5 × SSC を加えて、チャンバー内を湿度 100% に保持する。このアレイを含むチャンバーを、60 °C で約 6.5 時間インキュベートする。アレイは、第 1 洗浄緩衝液中 (1 × SSC, 0.1% SDS) において 45 °C で 10 分間、第 2 洗浄緩衝液中 (0.1 × SSC) において 45 °C で 10 分間それぞれ 3 回洗浄し、その後乾燥させる。

30

40

【0285】

検出

レポーター標識されたハイブリダイゼーション複合体は、Cy3 を励起するための 488 nm、及び Cy5 を励起するための 632 nm のスペクトル線を生成し得る Innova 70 混合ガス 10 W レーザー (Coherent, Inc., Santa Clara CA) を備えた顕微鏡で検出する。20 倍の顕微鏡対物レンズ (Nikon, Inc., Melville NY) を用いて、アレイ上に励起レーザー光を集中させる。このアレイを含むスライドを顕微鏡のコンピュータ制御 X-Y ステージに置き、対物レンズを通してラスタスキャンする。本実施例で用いた 1.8 cm × 1.8 cm のアレイは、20

50

μmの解像度でスキャンする。

【0286】

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ー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ーを用いて、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。

10

【0287】

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。

【0288】

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(AID)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、先ずデータは蛍光体間の光学的漏話(重複発光スペクトルに起因する)に対して補正される。

20

【0289】

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。

30

【0290】

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。Oligo4.06ソフトウェア(National Biosciences)及びTRICHのコーディング配列を用いて、適切なオリゴヌクレオチドを設計する。転写を阻害するためには、最も独特な5'配列から相補的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TRICHをコードする転写物に結合するのを阻害する。

40

【0291】

1.2 TRICHの発現

TRICHの発現及び精製は、細菌若しくはウイルスを基にした発現系を用いて行うことができる。細菌でTRICHが発現するために、抗生物質耐性及びcDNAの転写レベルを高める誘導性のプロモーターを含む好適なベクターにcDNAをサブクロニングする。このようなプロモーターには、lacオペレーター調節エレメントに関連するT5また

50

はT7バクテリオファージプロモーター及びtrp-lac(tac)ハイブリッドプロモーターが含まれるが、これらに限定されるものではない。組換えベクターを、BL21(DE3)などの好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピル-Dチオガラクトピラノシド(IPTG)で誘発されるとTRICHを発現する。真核細胞でのTRICHの発現は、昆虫細胞株または哺乳動物細胞株に一般にバキュロウイルスとして知られているAutographica californica核多面性ウイルス(ACMNPV)を感染させて行う。バキュロウイルスの非必須ポリヘドリン遺伝子を、相同組換え或いは転移プラスミドの媒介を伴う細菌の媒介による遺伝子転移のどちらかによって、TRICHをコードするcDNAと置換する。ウイルスの感染力は維持され、強いポリヘドリンプロモーターによって高いレベルのcDNAの転写が行われる。組換えバキュロウイルスは、多くの場合はSpodoptera frugiperda(Sf9)昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる。(例えば、Engelhard, E. K. 他(1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. 他(1996) Hum. Gene Ther. 7:1937-1945.を参照)。

【0292】

殆どの発現系では、TRICHが、例えばグルタチオンSトランスフェラーゼ(GST)、またはFLAGや6-Hisなどのペプチドエピトープ標識で合成された融合タンパク質となるため、未精製の細胞溶解物からの組換え融合タンパク質の親和性ベースの精製が素早く1回で行うことができる。Schistosoma japonicumからの26キロダルトンの酵素GSTによって、タンパク質の活性及び抗原性を維持した状態で固定されたグルタチオンで融合タンパク質の精製が可能となる(Amersham Pharmacia Biotech)。精製の後、GST部分を特定の操作部位でTRICHからタンパク質的に切断できる。アミノ酸8個のペプチドであるFLAGで、市販のモノクローナル及びポリクローナル抗FLAG抗体(Eastman Kodak)を用いた免疫親和性の精製が可能となる。6個の連続するヒスチジン残基のストレッチである6-Hisによって、金属キレート樹脂(QIAGEN)で精製が可能となる。タンパク質の発現及び精製の方法は、Ausubel(1995, 前出, ch 10, 16)に記載されている。これらの方法で精製したTRICHを直接用いて以下の実施例16、17、及び18のアッセイを行うことができる。

【0293】

1.3 機能のアッセイ

TRICHの機能は、哺乳動物細胞培養系において生理学的に高められたレベルでのTRICHをコードする配列の発現によって評価する。cDNAを、cDNAを高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにサブクローニングする。このようなベクターには、pCMV SPORTTM(Life Technologies)及びpCR3.1(Invitrogen, Carlsbad, CA)が含まれ、どちらもサイトメガロウイルスプロモーターを含んでいる。5~10µg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1~2µg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, N.Y. に記載されている。

【0294】

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がCD64に対する抗体のどちらかで被覆された磁気ビードを用いて分離することができる(DYNAL, Lake Success, N.Y.)。mRNAは、当分野で周知の方法で細胞から精製することができる。TRICH及び目的の他の遺伝子をコードするmRNAの発現は、ノーザン分析やマイクロアレイ技術で分析することができる。

10

【0295】

1.4 TRICHに特異的な抗体の作製

ポリアクリルアミドゲル電気泳動法(PAGE; 例えば、Harrington, M. G. (1990) Methods Enzymol. 1816-3088-495を参照)または他の精製技術で実質的に精製されたTRICHを用いて、標準的なプロトコルでウサギを免疫化して抗体を作り出す。

20

【0296】

別法では、TRICHアミノ酸配列をLASERGENEソフトウェア(DNA STAR)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)。

【0297】

通常、約15残基の長さのオリゴペプチドを、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を用いてブロッキング処理し、ウサギ抗血清と反応させて洗浄し、さらに放射性ヨウ素標識されたヤギ抗ウサギIgGと反応させる。

30

【0298】

1.5 特異的な抗体を用いる天然TRICHの精製

天然TRICH或いは組換えTRICHを、TRICHに特異的な抗体を用いるイムノアフィニティークロマトグラフィ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。

40

【0299】

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

50

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。

【0300】

16 TRICHと相互作用する分子の同定

TRICHと相互作用する分子には、輸送体基質、アゴニスト若しくはアンタゴニスト、G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。

10

【0301】

別法では、TRICHと相互作用する分子がFields, S. 及びO. Song (1989, Nature 340:245-246)に記載の酵母2-ハイブリッドシステム(yeast two-hybrid system)で単離された。TRICHもしくはその断片は、Gal4若しくはlexAのDNA結合ドメインを有する融合タンパク質として発現し、また潜在的な相互作用タンパク質が活性ドメインを有する融合タンパク質として発現する。TRICH融合タンパク質と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方法は、Niethammer, M. およびM. Shengらの著書(1998, Meth. Enzymol. 293:104-122)に記載されている。

20

【0302】

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

30

【0303】

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。

【0304】

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。 - 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時間に渡ってインキュベートされる。

40

【0305】

- 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して青く染色される。染色された細胞は、本技術分野で既知の電気生理学技術を用いて膜電気伝導度の違いを試験される。形質転換されていない細胞、及び

50

ノ又はベクター配列単独、若しくは - 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は、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、確認されうる。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クする。

【0306】

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(Ishihara(既出)、Jegla, T. 及びL. Salkoff (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ーニングされ、また0, 5-5 ngのmRNAが成熟段階4の卵母細胞へ注入される。注入された卵母細胞は摂氏18度で1-5日間インキュベートされる。インサイドアウトマクロパッチが、116 mMのK-グルコン酸、4 mMのKC1、および10 mM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2-5 MΩ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0 mVの保持電位で実行される。-100 mVから100 mVの電圧ランプが、サンプリング周波数500 Hzで得られる。電流測定はアッセイに於けるTRICHの活性に比例する。

【0307】

TRICHの変換活性は、アフリカツメガエル卵母細胞への標識された基質の取り込みによって検査される。ステージ5及び6における卵母細胞は、TRICH mRNA (卵母細胞あたり10 ng)で注入され、OR2培地(82.5 mMのNaCl、2.5 mMのKC1、1 mMのCaCl₂、1 mMのMgCl₂、1 mM Na₂HPO₄、5 mMのHepes、3.8 mMのNaOH、50 μg/ml ゲンタマイシン、pH 7.8)内で摂氏18度で3日間インキュベートされ、TRICHの発現を可能とする。卵母細胞は、次に標準取り込み培地(100 mMのNaCl、2 mMのKC1、1 mMのCaCl₂、1 mMのMgCl₂、10 mMのHepes/Tris、pH 7.5)へと移される。様々な基質(例えばアミノ酸、糖、薬剤、イオン、及び神経伝達物質)の取り込みは、標識された基質(例えば³H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、卵母細胞を3回洗浄することで終了する。TRICH活性は内部移行されたラベル基質のレベルに比例する。

【0308】

TRICHに関連するATPase活性は、放射性ラベルされたATP-[³²P]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³²Pの定量化によって測定される。反応混合物は、ATP-[³²P]と、好適な時間をかけて摂氏37度でインキュベートされた好適なバッファ中の様々な量のTRICHとを有する。反応は、トリクロロ酢酸の沈殿によって終了し、塩基で中和され、反応混合物のアリコットは、反応混合物を分離するべく、膜若しくはフィルター上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³²P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TRICHのATPase活性に比例する。

【0309】

18 TRICHアゴニスト及びアンタゴニストの同定

TRICHは、CHO(チャイニーズハムスターの卵巣)またはHEK(ヒト胎児腎臓)

10

20

30

40

50

） 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 (Velicelibi, G.ら (1999) Meth. Enzymol. 294 : 20-47 ; West, M. R.及びC. R. Mollroy (1996) Anal. Biochem. 241 : 51-58)。これらアッセイは、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、 Ca^{2+} 指示薬Fluo-4 AM、SBFI及びナトリウムグリーンのようなナトリウム感知染料、Cl指示薬MQAE (すべてMolecular Probesより入手可能) のような蛍光染料を用い、FLIPR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は、DiBAC₄ (分子プローブ) のようなoxonyl染料を用いて測定される。DiBAC₄ 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DiBAC₄ の取り込みが検出可能となる (Gonzalez, J. E.及びP. A. Negulescu (1998) Curr. Opin. Biotechnol. 9 : 624-631)。候補アゴニスト若しくはアンタゴニストは、既知のイオンチャネルアゴニスト、若しくはアンタゴニスト、ペプチドライブラリ、若しくはコンビケムライブラリより選択されてもよい。

10

20

【0310】

当業者は、本発明の範囲及び精神から逸脱することなく本発明の記載した方法及びシステムの種々の改変を行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。

【0311】

(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。

30

【0312】

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

【0313】

表3は、推定上のモチーフおよびドメインを含む各ポリペプチド配列の構造的な特徴、並びに各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。

【0314】

表4は、各ポリヌクレオチド配列の組み立てに用いたcDNA断片およびゲノム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。

40

【0315】

表5は、本発明の各ポリヌクレオチドの代表的なcDNAライブラリを示す。

【0316】

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。

【0317】

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】

50

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
1416107	1	1416107CD1	28	1416107CB1
1682513	2	1682513CD1	29	1682513CB1
2446438	3	2446438CD1	30	2446438CB1
2817822	4	2817822CD1	31	2817822CB1
4009329	5	4009329CD1	32	4009329CB1
6618083	6	6618083CD1	33	6618083CB1
7472002	7	7472002CD1	34	7472002CB1
1812692	8	1812692CD1	35	1812692CB1
3232992	9	3232992CD1	36	3232992CB1
3358383	10	3358383CD1	37	3358383CB1
4250091	11	4250091CD1	38	4250091CB1
70064803	12	70064803CD1	39	70064803CB1
70356768	13	70356768CD1	40	70356768CB1
5674114	14	5674114CD1	41	5674114CB1
1254635	15	1254635CD1	42	1254635CB1
1670595	16	1670595CD1	43	1670595CB1
1859560	17	1859560CD1	44	1859560CB1
5530164	18	5530164CD1	45	5530164CB1
139115	19	139115CD1	46	139115CB1
1702940	20	1702940CD1	47	1702940CB1
1703342	21	1703342CD1	48	1703342CB1
1727529	22	1727529CD1	49	1727529CB1
2289333	23	2289333CD1	50	2289333CB1
2720354	24	2720354CD1	51	2720354CB1
3038193	25	3038193CD1	52	3038193CB1
3460979	26	3460979CD1	53	3460979CB1
7472200	27	7472200CD1	54	7472200CB1

10

20

30

40

【表2】

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 相同体
1	1416107CD1	g7018605	1.9e-302	グルコース輸送体 [ドブネズミ] (Ibberson, M.他(2000) J. Biol. Chem. 275:4607-4612)
2	1682513CD1	g5263196	1.4e-153	伸張抑制可能非選択性チャネル(SIC) [ドブネズミ] (Cloning of a stretch-inhibitable nonselective cation channel. J Biol Chem. 1999 3月 5;274(10):6330-6335)
3	2446438CD1	g4589141	0	パノロイド受容体-様タンパク質1 [ヒト] (A capsaicin-receptor homologue with a high threshold for noxious heat. Nature 1999 398:436-441)
5	4093229CD1	g3873983	1.9e-64	Na ⁺ /Dat, K ⁺ 逆輸送体類似 [線虫]
6	6618083CD1	g9230651	4.7e-268	促進性グルコース輸送体ファミリ-メンバー-GLUT9 [ヒト] (Phay, J.E.他(2000) Genomics 66:217-220)
7	7472002CD1	g433960	0	大動脈CNGチャネル(ACNG) [カイウサザ] (Primary structure and functional expression of a cyclic nucleotide-gated channel from rabbit aorta. FEBS Lett. 1993 8月 23;329(1-2):134-138)
8	1812692CD1	g3928756	4.5e-48	一過性受容体潜在的チャネル7 [ヒト] (Nagamine, K.他(1998) Molecular cloning of a novel putative Ca ²⁺ channel protein (TRPC7) highly expressed in brain)
9	3232992CD1	g3874275	3.2e-70	酵母の低親和性グルコース輸送体 HXT4類似 [線虫]
10	3358383CD1	g3004482	1.4e-163	推定膜内在性輸送タンパク質 [ドブネズミ] (Schoming, E.他(1998) Molecular cloning and characterization of two novel transport proteins from rat kidney. FEBS Lett. 425:79-86)
11	4250091CD1	g3880445	5.7e-16	VM106R1(K ⁺ チャネル四量体化ドメイン類似) [線虫]
12	70064803CD1	g3874275	7.0e-84	酵母の低親和性グルコース輸送体 HXT4類似 [線虫]
13	70356768CD1	g183298	4.1e-54	GLUT5タンパク質 [ヒト] (Kayano, Y.他(1990) Human facilitative glucose transporters. Isolation, functional characterization, and gene localization of cDNAs encoding an isoform (GLUT5) expressed in small intestine, kidney, muscle, and adipose tissue and an unusual glucose transporter pseudogen- like sequence (GLUT6). J. Biol. Chem. 265:13276-13282)

10

20

30

40

【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:	種率スコア	GenBank 相同体
14	5674114CD1	g5771352	1.3e-238	内向き整流カリウムチャネルKir2.4 [ヒト] (Topert, C. 他(1998) Kir2.4: a novel K ⁺ inward rectifier channel associated with motoneurons of cranial nerve nuclei. J. Neurosci. 18:4096-4105)
15	1254635CD1	g3953533	1.7e-210	内向き整流カリウムチャネルKir3.1 [ハツカネズミ] (Mouri, T. 他(1998) Assignment of mouse inwardly rectifying potassium channel Kcnj16 to the distal region of mouse chromosome 11. Genomics 54:181-182)
16	1670595CD1	g9502260	2.3e-146	カチオン-塩化物共輸送体-相互作用タンパク質 [ヒト] (Caron, L. 他(2000) J. Biol. Chem. 275:32027-32036)
17	1859560CD1	g5834394	1.4e-101	硫酸塩輸送体 [ショウジョウバエ]
18	5530164CD1	g4903004	3.5e-20	UDP-N-アセチルグルコサミン輸送体 [ヒト] (Ishida, N. 他(1998) Molecular cloning and functional expression of the human solgi UDP-N-acetylglucosamine transporter. J. Biochem. 126:68-77)
19	139115CD1	g8131858	1.5e-49	推定胸腺関連性共輸送体TSCOT [ハツカネズミ] (Kim, M.G. 他(2000) J. Immunol. 164:3185-3192)
20	1702940CD1	g5725224	2.5e-143	bk212A22 (推定アポリポタンパク質) [ヒト]
21	1703342CD1	g6003536	8.1e-06	カルシウムチャネル α -1サブユニット [カンジダ]
22	1727529CD1	g4529890	0.0	NG22 [ヒト]
23	2289333CD1	g4539333	5.6e-35	推定アミノ酸輸送タンパク質 [シロイヌナズナ]
24	2720354CD1	g3875242	4.3e-38	ミトコンドリアキャリアタンパク質類似 [線虫]
26	3460979CD1	g1931644	4.7e-08	膜タンパク質PTM1プレカーサー同系 (推定主要促進性スパーファミリー輸送体) [シロイヌナズナ]
27	7472200CD1	g2811254	2.8e-21	アミロライド感受性Na ⁺ チャネル [ショウジョウバエ] (Adams, C.M. 他(1998) J. Cell Biol. 140:149-152)

10

20

30

40

【表4】

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	1416107CD1	477	S99 T2 T281 S430 T205	N349	糖輸送体ドメイン: A29-F474 糖輸送体タンパク質シグネチャ: V108-L174, L293-S350, 641-I51, L124-V143, Q267-F277, V375-L396, S398-I410, グルコース輸送体シグネチャ: F257-Y278, V375-S388, 6439-V459	MOTIFS HMER-PFAM BLIMFS- BLOCS ProfileScan BLIMFS- PRINTS HMER
2	1682513CD1	498	S47 T131 S286 T367 S463 T67 S315 S382	N278 N411 N429	膜貫通ドメイン: L298-A279, L293-M313, L320-Y339, Y438-F457 グルコース輸送体シグネチャ: P310-M339	MOTIFS BLIMFS- PRINTS HMER
3	2446438CD1	764	T64 T329 S23 S67 T106 S268 S339 S348 S353 S464 S468 S667 T692 S697 T720 T101 T115 S325 T414 Y110 Y227 Y333	N570	膜貫通ドメイン: A95-Y117, V142-F160, L178-Y201 A200-F219, F244-L263, P319-M339 アンキリン反復: R162-Q194, F208-S243, Q293-F328 膜貫通ドメイン: L386-F405, I463-V486, F538-S557, L623-I642	MOTIFS HMER-PFAM HMER
4	2817822CD1	255	T30 S167 T33 T71 S23 T50 S134 T162 T238		カリウムチャネルシグネチャ: R76-T95	MOTIFS BLIMFS- PRINTS MOTIFS HMER HMER-PFAM
5	4009329CD1	584	S258 S321 T70 S271 S273 S468 S514 S62 T132	N60 N125	シグナルペプチド: M1-G29 ナトリウム/カルシウム交換体タンパク質 ドメイン: I113-Q252, L431-F576 膜貫通ドメイン: T101-F121, T166-I189, L234-Y251, Y382-A402, F492-R513, L560-M584	MOTIFS HMER HMER-PFAM HMER

【表 5】

10

20

30

40

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	潜在的リソ酸化部位	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	6618083CD1	416	T111 S4 S164 S274 T374 S16 S226 S269	N45 N61 N410	シグナルペプチド: M1-G37 糖輸送体ドメイン: A30-E416 糖輸送タンパク質 シグネチャ: A123-L189, S39-V49, I139-M158, Y298-F308 グルコース輸送体シグネチャ: V288-Y309, I356-Q376 膜貫通ドメイン: Y63-V99, I356-L375	MOTIFS SPScan HMMER HMMER-PFAM BLIMPS- BLOCKS ProfiLescan BLIMPS- PRINTS
7	7472002CD1	664	S402 S40 S46 S93 T107 T313 T337 T381 T422 S476 S552 T591 T606 T634 T2 S35 T124 T208 T418 T448 Y648	N311 N379	膜貫通領域, 環状ヌクレオチド結合ドメイン: K469-D665, A460-S476, G478-V501, G516-L525 膜貫通ドメイン: Y350-I375	MOTIFS HMMER-PFAM BLIMPS- BLOCKS
8	1812692CD1	242	T95 S35 S37 T124 S204 S221 S2 S9 S20 T21 S86 Y36	N15 N84	タンパク質メラスタチン染色体 膜貫通 PD018035: Y117-I227	BLAST- PRODOM
9	3232992CD1	398	T307 S338 S100 T133 T241 T303 S377 S395	N161	膜貫通ドメイン: A217-Q242, L247-F264, L350-F368 糖輸送タンパク質シグネチャ: L45-694, Y30-I86	HMMER MOTIFS BLIMPS- BLOCKS ProfiLescan
10	3358383CD1	553	S337 S352 S409 T58 S60 S109 T133 S337 T433 T527 S167 S201 T226 S282 T323 T405	N39 N56 N62 N102 N107 N473	膜貫通ドメイン: F204-A222, M470-Y493, I500-I519 輸送体のgIpIファミリー: V151-D168 有機輸送タンパク質, 腎臓細胞イオン輸送体, カチオン腎臓糖質 PD151320: N102-L144	HMMER MOTIFS BLIMPS- BLOCKS BLAST- PRODOM

10

20

30

40

【表 6】

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
11	4250091CD1	213	S2 S93 S172 S184 T17 T22 T137 S210 Y89 T365 S11 S364 S453 S292 T361 S390 S466	N188	カリウムチャネルシグネチャ: Q48-T67 膜貫通ドメイン: V222-R239, G327-V350, M413-F432 糖輸送タンパク質シグネチャ: L153-G202, V138-I204	MOTIFS BLIMPS- PRINTS HMMER
12	70064803CD1	476				MOTIFS BLIMPS- BLOCKS ProfileScan
13	70356768CD1	246	S100 S238 S118 S215	N34 N50	シグナルペプチド: M1-627 糖輸送タンパク質シグネチャ: I127-G176, A112-V178, T28-I38, M128-M147, M133-R168	MOTIFS SPScan HMMER BLIMPS- BLOCKS ProfileScan BLIMPS- PRINTS HMMER BLAST-DOMO
14	5674114CD1	436	S11 T80 S154 S340 S362 T263 S376 S422 Y47	N195	膜貫通ドメイン: M163-L181 糖輸送タンパク質: DM00135 (P46408): A112-C229 内向き整流カリウムチャネル ドメイン: V53-L394, R72-L118, P126-Q169, G170-V199, A204-Q238, D296-Y346, T358-E368 膜貫通ドメイン: W88-L114 内向き整流カリウムチャネル: DM00148 (P52188): N34-A395 内向き整流カリウムチャネル: PD001103: V53-Q372 KIR2.4 タンパク質: PD124342: A373-P436 PD063376: M1-F52	MOTIFS HMMER-PFAM BLIMPS-PFAM HMMER BLAST-DOMO BLAST- PRODOM

10

20

30

40

【表 7】

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
15	1254635CD1	453	S408 T16 T99 S416 S29 T209 T216 T250 S375		シグナルペプチド: M1-A32 膜貫通ドメイン: F109-A131, V182-I200 内向き整流カリウムチャネルドメイン: L72-T389 内向き整流カリウムイオンチャネルサブファミリー PD001103: K74-K403 電圧ゲート内向き整流カリウムチャネルBIR9, K1R5, I, 膜貫通 PD063375: M36-H73 内向き整流カリウムイオンチャネル: DM00448 (P52185) 27-380: K64-L379 内向き整流カリウムチャネルシグネチャ: S91-L137, P145-Q188, S189-R218, A223-R257, N310-G360, A371-W381	SPScan EMMER EMMER-PFAM BLAST-PRODOM BLAST-PRODOM BLAST-DOMO BLIMPS-PFAM
16	1670595CD1	299	S38 T211 S263 S33 T187	N193 N208	膜貫通ドメイン: L37-L58, A74-193, L134-Y154, F159-V179 感受性共輸送体, 塩化物イオンチャネル: DM01178 (S06903) 1-128: A32-L153	EMMER BLAST-DOMO

10

20

30

40

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
17	1859560CD1	606	T96 T116 S298 S571 T572 S595 T245	N294	膜貫通ドメイン: L45-163, I396-I421 硫酸塩輸送体ファミリードメイン: M137-A468 硫酸塩輸送体タンパク質 シグネチャ: A54-1107, L125-L176 硫酸塩輸送体タンパク質 膜貫通ドメイン: P0001255: M137-L465 硫酸塩輸送体タンパク質 膜貫通ドメイン: P0001121: L30-6143 硫酸塩輸送体: DM01229 (S64926) 69-531: L30-W428 硫酸塩輸送体モチーフ: P77-R98	HMMER HMMER-PFAM BLIMPS-BLOCKS BLAST-PRODOM BLAST-PRODOM BLAST-DCMO MOTIFS
18	5530164CD1	324	S2 S139 Y115	N99 N100 N101 N232	膜貫通ドメイン: A145-V163, L302-L319	HMMER
19	139115CD1	445	S54 S32 S77 S217 S424 S438 T150 S237 S443 Y23	N22 N30 N37 N127 N213 N235	膜貫通ドメイン: W182-F200, F242-1261, Y283-F302 糖輸送体モチーフ: L75-S91 グルコース輸送体シグネチャ: W182-L202	HMMER MOTIFS BLIMPS-PRINTS
20	1702940CD1	337	T30 S167 T208 S306		アポリポタンパク質Lプレカーサー, 脂質輸送, 糖タンパク質, シグナル, DJ6802.1, P0042084: M1-D836	BLAST-PRODOM
21	1703342CD1	273	T3 S63 T222 S248 S250 T10 S98 S219 S224		膜貫通ドメイン: F101-A118, M142-F161, F170-1186, I202-I218 イオン輸送タンパク質ドメイン: L97-L269 (スコア: -132.1, E-値: 0.72)	HMMER HMMER-PFAM

【表 9】

10

20

30

40

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
22	1727529CD1	710	S31, S102, S119, T135, S304, S22, S218, S430, S431, T494, S573, S619, Y13	N29, N69, N155, N197, N298, N393, N405, N416, N678	膜貫通ドメイン: C38-Y58, V241-L266, W309-V326, F356-T375, F440-L458, T499-I522, L598-F618, I645-V663 ABC 3 輸送ファミリー: S228-Q427 (スコア: -182.9, E-値: 2.1) 膜イオン交換体シグネチャ: A311-L330	HMMER
23	2289333CD1	476	T97, T7, S8, S125, T443, S272, S322, T351, T451, Y184	N166, N169, N212, N425, N467	膜貫通ドメイン: L54-I81, V127-F145, Y184-S208, L279-G297, I331-K357, I426-I451 膜貫通アミノ酸輸送体タンパク質ドメイン: A53-F436	BLIMPS- PRODOM HMMER
24	2720354CD1	237	T17, T64, S172		アミノ酸輸送体タンパク質: バニリアーゼ, 膜貫通, 推定 プロリン PD001875: D27-I337 シグナルペプチド: M1-G15 ミトコンドリアキャリアータンパク質ドメイン: S25-L109, L122-T202 ミトコンドリアエネルギ-伝達タンパク質シグネチャ: L128-Q152 ミトコンドリアエネルギ-伝達タンパク質シグネチャ: L27-I75, V123-Q171 ミトコンドリアキャリアータンパク質シグネチャ: G87-D107, V136-D154 アテニンヌクレオチド輸送体1シグネチャ: R63-V84, E176-R191 ミトコンドリアキャリアータンパク質モチーフ: P46-L54, P143-L151 輸送タンパク質, 膜貫通, 細胞内ミトコンドリアの, ADP/ATP: PD000117: L31-F200 ミトコンドリアエネルギ-伝達タンパク質: DM00026 (P38087) 243-325: L128-Y209	SPScan HMMER-PFAM BLIMPS- BLOCKS ProfileScan BLIMPS- PRINTS BLIMPS- PRINTS MOTIFS BLAST- PRODOM BLAST-DOMO

10

20

30

40

【表 10】

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
25	3038193CD1	345	T204 T251 S57 S243 T263 T308 S340	N246	膜貫通ドメイン: L87-L95, I134-I156, I224-F242 ナトリウム胆汁酸同歩輸送ファミリー: Y44-P212 (スコア: -7.0, E-値: 9.0e-4) リン酸塩輸送体シグネチャ: F153-6171	HMMER HMMER-PFAM
26	3460979CD1	521	S115 T184 S75 T93 S100 S126 S128 S134 S148 S183 S213 S256 S363 S389 S430 S510 T171 S180 S235 T247 S422 Y506	N70 N169 N211	膜貫通ドメイン: L265-L284, I335-I361 タンパク質ブレカ-サーPTM1: 膜貫通シグナル PD014374: G219-E517 (P-値: 7.1e-07)	BLIMPS- PRODOM HMMER BLAST- PRODOM
27	7472200CD1	555	T43 S56 T92 T148 T298 S423 S468 S20 S52 S82 S96 T184 S208 S252 S393	N132 N175 N311 N361 N421	アミロライド-感受性ナトリウムチャネル α-サブユニットシグネチャ PR01078: Y102-N118, Y342-Q353, Q353-P370, Q388-N404, G455-E471 膜貫通ドメイン: Y452-F475 アミロライド-感受性ナトリウムチャネル ASC: F38-L476	BLIMPS- PRINTS HMMER HMMER-PFAM
					アミロライド-感受性ナトリウムチャネル タンパク質 BL01206: R37-L47, Y342-F368, L427-L472	BLIMPS- BLOCKS

10

20

30

40

【表 1 1】

表4-1

ポリスクレオチド SEQ ID NO:	インサイト ポリスクレオチド ID	配列長さ	選択された断片	配列断片	5'位置	3'位置
28	1416107CB1	2080	1-109, 1901-2080, 1363-1446	α1941704	116	609
				6813453H1 (ADRETUR01)	319	870
				6605280H1 (UTREDIT07)	820	1447
				881845R1 (THYRN0T02)	889	1479
				1416107F6 (BRAINOT12)	1348	1896
				1416107T6 (BRAINOT12)	1579	2080
				71826604V1	1563	1974
				7448905T1 (BRAYDIN03)	1578	2050
				6300413H1 (UTREDIT07)	423	752
				71805807V1	750	1580
				71827149V1	1584	2080
				71807187V1	701	1241
29	1682513CB1	2128	1-1535, 1560-1581	6813453R6 (ADRETUR01)	1	312
				70207988V1	1	469
				70213506V1	394	872
				70211216V1	489	985
				70210573V2	852	1468
				70207907V1	988	1512
				70210540V2	1357	1948
				2866122T6 (KIDNNOT20)	1548	2108
				70211461V1	1597	2128
				5073532H2 (COLCTUT03)	1	311
				6309494H1 (NERDPTN03)	260	812
				30	2446438CB1	2825
70382927D1	996	1525				
70386205D1	1469	2075				
1798255F6 (COLNNOT27)	1632	2194				
1562088F6 (SPLNNOT04)	2178	2727				
2514370F6 (LIVRUT04)	2303	2825				
1502510F6 (BRAITUT07)	1	439				
70271734V1	183	768				
70273052V1	431	930				
70271651V1	891	1453				
2817822F6 (BRSTN0T14)	981	1538				
70272460V1	1094	1718				
31	2817822CB1	1718	1-71, 609-914			

10

20

30

40

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
32	4009329CB1	2000	1-962	6466193H1 (PLACFE01)	1	640
				6780428J1 (OVARDIR01)	582	1260
				6307863H1 (NERDFN03)	725	1364
				6781250H1 (OVARDIR01)	972	1639
				7253109J1 (PROSTME05)	1514	1842
33	6618083CB1	2216	1-96, 1201-2216	6759035J1 (HEAONOR01)	1515	2000
				5722362H1 (SEMNOT05)	1	581
				70789558V1	504	1127
				70787652V1	588	1203
				70791819V1	1050	1650
34	7472002CB1	1995	1-862, 1766-1995	70787819V1	1361	1984
				70791126V1	1695	2216
				g2121300.v113.gs_2.nt.edit	1	1995
				1812692F6 (PROSTUT12)	564	984
				5425924F6 (PROSTMF07)	1	488
35	1812692CB1	988	1-147, 244-570	g2525933	823	988
				5008333F6 (PROSTUT21)	283	804
				224000R6 (PANCNOT01)	2435	3087
				6825934J1 (SINTNOR01)	1	515
				7062063H1 (PENITMN02)	2683	3179
36	3232992CB1	3179	2106-2665, 1-1646	4491105H1 (BRADIT02)	2167	2861
				1698347F6 (BLADTUT05)	1762	2333
				70053653D1	1392	1870
				1807402F6 (SINTNOR13)	476	1019
				70055908D1	1317	1837
				7170824H2 (BRSTMC01)	719	1373
				6555265H1 (BRAFNON02)	1859	2372
				g1444660	992	1464
				g1009986	768	1254
				027195T6 (SPLNFF01)	1674	1986
				g1505781	552	1252
				3358383T6 (PROSTUT16)	1341	1691
				6221856U1	585	1252
6221857U1	1	727				
37	3358383CB1	1986	1465-1986, 1340-1371	g1444660	992	1464
				g1009986	768	1254

10

20

30

40

【表 1 3】

表4-3

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
38	4250091CB1	3294	1-920, 1991-2034, 2488-2760, 1365-1630	9715570 70759966V1 4250091F6 (BRADDI01) 5715843H1 (PANCNOT16) 707899723V1 966456R6 (BRSTNOT05) 70759467V1 70788682V1 7056848H1 (BRALNON02) 858645R1 (BRAITUT03) 70761829V1	2893 1864 1 2539 444 2862 1199 604 2373 1940 1333 1540 576 1361 1181 501 1 509 1072 1 368 1384 1060 1078 690 939 184 1 655 75 1 1126 548 960 368 1	3294 2477 532 3235 1032 3293 1842 1302 2989 2507 1947 2043 1280 2019 1797 1183 534 1081 1730 524 1078 1915 1720 1809 1040 1290 944 1311 1177 573 515 1730 724 1627 1147 641
39	70064803CB1	2043	1-22, 544-1285	2758549R6 (THPLAZS08) 6810024J1 (SKIRNOR01) 1676182T6 (BLADNOT05) 2109762R6 (BRAITUT03) 70503885V1 7177480H2 (BRAXDIC01)	1540 576 1361 1181 501 1	2043 1280 2019 1797 1183 534
40	70356768CB1	1915	1241-1263, 853-897, 1-143, 1450-1532, 668-820	70450108V1 70451575V1 1468307F6 (PANCUT02) 70451567V1 70449058V1 70449392V1	509 1072 1 368 1384 1060	1081 1730 524 1078 1915 1720
41	5674114CB1	1809	1-402	6776218J1 (OVARDI01) 3024042H1 (PROSDIN01) 6292787H1 (BMARUNA01) 6776218H1 (OVARDI01) 95686663.v113.gs_16.nt	1078 690 939 184 1	1809 1040 1290 944 1311
42	1254635CB1	1730	1-106, 1711-1730, 567-635, 696-900	2613664F6 (ESOGTUT02) SXBC01035V1 2863343F6 (KIDNNOT20) 2614317T6 (GBLANOT01) SCSA01493V1 3323244T6 (PTHYNOT03)	655 75 1 1126 548 960	1177 573 515 1730 724 1627
43	1670595CB1	1147	746-980, 1-696	SCIA02891V1 SCIA04658V1	368 1	1147 641

10

20

30

40

【表 1 4】

表4-4

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
44	1859560CB1	2745	1-820, 855-2059	6195927H1 (PIUNON01) 824186R1 (PROSNOT06) 1399644F6 (BRAITUT08) 6812696J1 (ADREURO1) 7255024H1 (FIBRTXC01) 2127239R7 (KIDNNOT05) 5990868H1 (PTVUT02) 6826978H1 (SINPNOR01) 6850265H1 (BRAIFEN08) 4677711H1 (NOSEDIT02) 1859560T6 (PROSNOT18) 4320381H1 (BRADDIT02) 7175713H1 (BRSTPMC01) 3217236H1 (TESFNOT07) 2552002H1 (LUNGTUT06) 70039789V1 70090155V1 6830248J1 (SINPNOR01) 6729730H1 (COLLITUT02) 2850920F6 (BRSTTUT13) 6125934H1 (BRAHNON05) 6059181H1 (BRAENOT04) 1956752F6 (CONNNOT01) 6455739H1 (COLNDIC01) 70077060U1 7126228H1 (COLNDIY01) 2468105F6 (THYRNOT08) 70079483U1 70122236V1 70122363V1 351595R1 (LVENNOT01) 70480521V1 4335403F6 (KIDCTMT01) 70466195V1 70466476V1	2367 1159 503 30 814 1732 1793 1 2056 1107 2247 1925 178 530 1 2452 1781 542 1214 929 2505 1923 1452 528 1900 1 827 2145 1424 1451 2056 1027 1 559 494	2745 1718 973 732 1364 2176 1795 479 2740 1381 2745 2202 811 812 242 3175 2492 1211 1889 1416 3204 2496 1893 1219 2535 580 1463 2763 1971 1976 2566 1639 516 1157 1102
45	5530164CB1	3204	1241-2064, 1-548, 572-1097			
46	1391115CB1	2763	1-770, 1345-1389, 1455-1683			
47	1702940CB1	1639	1-246, 1536-1639			

10

20

30

40

表4-5

ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチド ID	配列長さ	選択された断片	配列断片	5'位置	3'位置
48	1703342CB1	1600	1-812	3348562HL (BRAJUTU24)	1	282
				285125RL (FOSIHF02)	1041	1598
				7071066HL (BRAUTDR02)	250	854
				6494627HL (BONRN0T01)	1220	1600
				6879086JL (LNODNOR03)	360	1094
				957891HL (KIDNNOF05)	2085	2380
				60211961V1	691	1237
				6800135JL (COLENOR03)	1250	1983
				6798918JL (COLENOR03)	1693	2284
				60211964U1	262	807
49	1727529CB1	2380	1-569, 1228-1654	6798894HL (COLENOR03)	1089	1774
				3249035F6 (SEMVNOF03)	1	626
				3566495HL (BRONNOF02)	1984	2298
				2552315T6 (LUNGTUT06)	1322	1862
				9872898	848	1328
				1435329F1 (PANCNOF08)	2197	2725
				3553901HL (SYNONO1)	2570	2865
				2508452HL (CONVUT01)	1	114
				2771704HL (COLANOT02)	1815	2078
				6998443HL (HEALDIR01)	2	553
50	2289333CB1	3038	1-611, 2497-2524	91665184	2594	3038
				2289333R6 (BRAJN0N01)	1254	1708
				91156003	2524	3032
				5597992HL (UTREN0N03)	1029	1286
				95545742	612	1066
				5836345HL (BRAIDIT05)	2624	2880
				4220788F6 (PANCNOF07)	613	958
				1994713T6 (BRSTUT03)	1949	2451
				2040880R6 (HIPON0N02)	463	821
				2720354F6 (LUNGTUT10)	490	1046
51	2720354CB1	2608	1-2058	6942433HL (FTUEFUR01)	885	1437
				6121303HL (BRAHON05)	1630	2340
				6558224HL (BRAFN0N02)	1713	2406
				6940932HL (FTUEFUR01)	1	465
				91927466	325	872
				6826181JL (SINFNOR01)	1062	1666
				6197805HL (PIUNON01)	2154	2608

10

20

30

40

【表16】

表4-6

ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチド ID	配列長さ	選択された断片	配列断片	5'位置	3'位置
52	3038193CB1	3804	3392-3457, 1169-1264, 1-829, 2271-2483, 1319-1363	044564H1 (TBLYNOT01) 901446R6 (BRSTTUT03) 94088232 1428831H1 (SINTBST01) 2741328T6 (BRSTTUT14) 4970206H1 (KIDENC10) 2768967H1 (COLANOT02) 5688762F6 (BRAJUNT01) 3038193F6 (BRSTNOT16) 6477440H1 (PROSTMC01) 70809191V1 3154867H1 (TLYMXT02) 2257401R6 (OVARTUT01) 94268882 2257401T6 (OVARTUT01) 2237852F6 (PANCUT02) 3460979F6 (293TFLT01) 7161336H1 (PLACNOR01) 6800921J1 (COLENCOR03) 7057496H1 (BRALNOR02) GNN_96554406_006	2311 1733 3459 473 3235 1029 772 36 1284 1843 2411 1 2660 1421 2896 519 1000 582 1 1206 1	2562 2282 3804 661 3804 1303 1020 621 1710 2486 2832 272 3167 1815 3520 945 1500 1193 557 1894 1668
53	3460979CB1	1894	1-36 1746-1894			
54	7472200CB1	1668	1-1668			

10

20

30

40

【表17】

表5

ポリヌクレオチド SEQ ID NO:	インサイトプロジェクト ID	代表的ライブラリ
28	1416107CB1	UTREDIT07
29	1682513CB1	SPLNNOT11
30	2446438CB1	MCLDPXN03
31	2817822CB1	BRAITUT07
32	4009329CB1	OVARDIN02
33	6618083CB1	HELAUNT01
35	1812692CB1	PROSTUT12
36	3232992CB1	PANCN0T01
37	3358383CB1	PROSTUT16
38	4250091CB1	BRAITUT03
39	70064803CB1	THEPLAZS08
40	70356768CB1	HNT2AGT01
41	5674114CB1	OVARDIR01
42	1254635CB1	LUNGFET03
43	1670595CB1	BRAITUT24
44	1859560CB1	NGANN0T01
45	5530164CB1	BRAYDIN03
46	139115CB1	SININ0T18
47	1702940CB1	BRAVXT04
48	1703342CB1	EOSLHET02
49	1727529CB1	PROSN0T18
50	2289333CB1	LUNGFUT06
51	2720354CB1	PROSTUS23
52	3038193CB1	LIVRN0N08
53	3460979CB1	COLENOR03

10

20

30

40

ライブラリ	ベクター	ライブラリの説明
UTRED1T07	pINCY	ライブラリは、女性の子宮内膜検査の際に採取された病変子宮内膜組織から単離した RNA を用いて作製した。病変は、子宮内膜のβ3の欠損とII型異常を同時に示した。
SPLNN0T11	pINCY	ライブラリは、14歳のアジア人男性から全脾切除の際に採取した病変脾臓組織から単離した RNA を用いて作製した。病変は、突発性血小板減少性紫斑病と一致する変化を示した。患者にはあざがあった。患者の投薬は、ビンクリスチン(Vincristine)を含む。
MCLDTXN03	pINCY	ライブラリは、二つの樹状細胞ライブラリのプールから構築された。開始ライブラリは、男性の脐帯血液 CD34+前駆細胞からの未処置及び処置誘導樹状細胞から単離した RNA を用いて作製した。この細胞は、顆粒球/マクロファージコロニー刺激因子(GM-CSF)、腫瘍壊死因子α(TNFα)、及び幹細胞因子(SCF)によって誘導された。このライブラリは、著しく長い(48時間/回)アニーリングハイブリダイゼーションが用いられたの以外は Soares 他(1994)Proc. Natl. Acad. Sci. 米国 91:9228 及び Bonaldo 他(1996)Genome Res. 6:791の条件に従って標準化した。
BRA1TUT07	pINCY	ライブラリは、32歳の白人男性の大脳髄膜病変の切除中に採取した脳の左前頭葉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、低グレードの癒着性の神経腫瘍 (desmoplastic neuronal neoplasm) である。患者には、吐き気、嘔吐、及び頭痛の症状があった。患者の病歴は、飲酒、喫煙、及び6年間の週2回のマリファナ使用を含む。家族歴には、祖父母にアテローム硬化性冠動脈疾患患者がある。
OVARD1ND2	pINCY	ライブラリは、卵巣組織ライブラリから作製した。開始 RNA は39歳の白人女性の腹式子宮全切除、両卵管卵巣摘出術、拡張術及び掻爬術、部分的な結腸切除、随伴性虫垂切除、及び一時的な人工肛門形成術の際に採取した病変卵巣組織から単離した RNA を用いて作製した。病変は、右及び左の付属器、腸間膜及びS状結腸の筋結合組織が広範囲にわたり、子宮内膜の影響を受けていた。子宮内膜症の影響は、子宮の前及び後漿膜表面と盲管にも及んでいた。子宮内膜は増殖性であった。関連腫瘍組織の病変は、複数の平滑筋腫(壁内に3、漿膜下に1)を示した。患者は、腹痛と不妊症を患っていた。患者の病歴には、側彎症が含まれる。過去の手術には、腹腔鏡による胆嚢摘除術及び診査開腹術が含まれる。患者の投薬には Megace、Danazol、及び Lupron を含む。家族歴には、母親に高脂血症、良性高血圧症、高脂血症、アテローム硬化性冠動脈疾患、大動脈冠動脈バイパス術、抑うつ障害、脳癌、及びII型糖尿病を含む。このライブラリは、著しく長い(48時間/回)アニーリングハイブリダイゼーションが用いられたの以外は Soares 他(1994)Proc. Natl. Acad. Sci. 米国 91:9228 及び Bonaldo 他(1996)Genome Res. 6:791の条件に従って標準化した。
HELAUNT01	pINCY	ライブラリは、31歳の黒人女性から採取した頸部腺癌から得た未処置 HeLa 細胞株から単離した RNA を用いて作製した。

ライブラリ	ベクター	ライブラリの説明
BRA1TUT03	PSPORT1	ライブラリは、17歳の白人女性から脳腫瘍病巣切除の際に左前葉から採取した脳腫瘍組織から単離したRNAを用いて作製した。病理学的には、グレード4の纖維状の巨細胞及び小細胞星状細胞腫を示していた。家族歴には、良性高血圧症及び脳血管障害が含まれていた。
HNT2AGT01	PBLUESCRIPT	ライブラリは、抗原感受神経細胞前駆体の特性を示すヒト奇形癌を由来とするhNT2細胞株より単離したRNAを用いて、Stratagene社で作製された(STR937233)。細胞は5週間レチノイン酸で処理され、2週間有糸分裂阻害剤で処理され、調整培地中で更に4週間成熟することが許された。
OVARD1R01	pcDNA2.1	ライブラリは、45歳の白人女性の腹式子宮全切除、両卵巣摘出術、腫の拡張術及び固定、及び随伴性虫垂切除術の際に採取した右卵巣組織から単離したRNAを用いて作製した。病変は、右及び左卵巣の間質性卵巣膜細胞増殖症を示した。関連腫瘍組織の病変は、左卵巣に皮様嚢腫(良性嚢胞性奇形腫)を示した。複数(3)の壁内平滑筋腫が同定された。子宮頸部は扁平上皮化生を示した。患者の病歴は、異常子宮出血、女性ストレス失禁、脱毛、抑うつ障害、肺炎、通常分娩、及び欠乏性貧血を含む。家族歴は、良性高血圧症、アテローム硬化性冠動脈疾患、高脂質血症、及び初期変化群を含む。
PANCN0T01	PBLUESCRIPT	ライブラリは、頭部外傷で死亡した29歳の白人男性の膵臓組織から単離したRNAで作製した。
PROSTUT12	pINCY	ライブラリは、65歳の白人男性から根治前立腺切除の際に採取した前立腺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、腺癌(Gleasonグレード2+2)を示していた。腺癌腫瘍組織は過形成も示していた。患者には、前立腺特異的抗原(PSA)の上昇も見られた。
PROSTUT16	pINCY	ライブラリは、55歳の白人男性から採取した前立腺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、Gleasonグレード5+4の腺癌を示していた。また、腺癌腫瘍組織は過形成も示していた。患者の病歴には、前立腺特異抗原(PSA)の上昇があった。患者の病歴には、腎臓結石が含まれる。家族歴には、肺癌及び乳癌が含まれる。
THPIAZS08	PSPORT1	ライブラリは、5'-aza-2'-デオキシジチジン(AZ)処理THP-1前単核細胞株ライブラリよりの5.76×10 ⁶ 個のクローンをを用いて作製された。開始RNAは、0.8μMのAZで3日間処理したTHP-1前単核細胞からならなかった。AZ処理THP-1細胞ライブラリからの5.76×10 ⁶ 個のクローンは、次に未処理のTHP-1細胞ライブラリよりの5.0×10 ⁶ 個のクローンで、2ラウンドの差し引きハイブリダイゼーションにかけられた。差し引きハイブリダイゼーションの条件はSwaroopら(1991)のNucleic Acids Res. 19:1954、及びBonaldoら(1996)のGenome Research 6:791の方法論を基にした。THP-1(ATCC TIB 202)は、急性単球白血病の1歳男児の末梢血液に由来するヒト前単核細胞株である(Int. J. Cancer (1980) 26:171を参照)。

ライブラリ	ベクター	ライブラリの説明
BRAITUT24	pINCY	ライブラリは、50歳の白人女性から大脳髄膜病変部切除の際に採取した右前頭葉脳腫瘍組織より単離したRNAを用いて作製した。病変は髄膜腫を示した。家族歴には、大腸癌及び脳血管障害が含まれる。
BRAYDINO3	pINCY	この標準化された脳組織ライブラリは、脳組織ライブラリからの独立したクローン6.7×10 ⁶ 個から作製された。開始RNAは、脳血管発作で死亡した57歳の白人男性から採取した病変視床下部組織から単離したRNAを用いて作製した。患者の病歴には、ハンチントン病及び肺気腫を含む。このライブラリは、著しく長い(48時間/回)アニーリングハイブリダイゼーションが用いられたの以外は、Soares 他PNAS(1994) 91:9228及びBonaldi 他 Genome Research 6(1996):791の条件に従って2回標準化した。このライブラリは、クローンを含む遺伝子挿入の選択用に線形化及び再流布された。
LUNGFET03	pINCY	ライブラリは、妊娠20週目に死亡した白人の女胎児から採取された肺組織から単離したRNAを用いて作製した。
NGANNOT01	PSPORT1	ライブラリは、9歳白人男児の胸壁軟組織の切除の際に採取した神経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、節細胞腫を示していた。家族歴には喘息が含まれる。
BRAVXT04	PSPORT1	ライブラリは、IL-1を含むサイトカインとの組み合わせで4時間から6時間の間刺激されたヒトの星状細胞の分離した集団から単離したRNAを用いて作製した。このRNAは、ポリA RNAの単離及びライブラリの作製のためにプールされていたものである。
EOSIHET02	PBLUESCRIPT	ライブラリは、48歳の白人男性から末梢血液細胞を採取し、その末梢血液細胞から単離したRNAを用いて作製した。患者の病歴には、過好酸球増加症がある。ライト染色により、細胞集団の77%以上が好酸球であることが分かった。
LIVRNON08	pINCY	この標準化された脳組織ライブラリは、プールされている肝臓組織ライブラリからの独立したクローン5.7×10 ⁶ 個から作製された。開始RNAは、無酸素症で死亡した4歳のヒスパニック系男児及び妊娠16週目に無脳症で死亡した白人の女胎児から採取されプールされた肝臓組織から単離したRNAを用いて作製した。血清学的には、4歳児はサイトロメガウィルスに対して陽性であった。患者の病歴には、4歳児の喘息が含まれる。家族歴には、胎児の母親が産前に毎日ビタミンを摂取していたのと、僧帽弁脱出が含まれる。このライブラリは、著しく長い(48時間/回)アニーリングハイブリダイゼーションが用いられたの以外は、Soares 他 Proc. Natl. Acad. Sci. 米国(1994) 91:9228及びBonaldi 他(1996) Genome Research 6:791の条件に従って2回標準化した。

表 6-4

ライブラリ	ベクター	ライブラリの説明
LUNGTUT06	pINCY	ライブラリは、部分的な肺の切除術の際に80歳の白人女性から切除された尖端の肺腫瘍組織から単離されたRNAを用いて作製された。病理学報告は、転移性の顆粒膜細胞腫を示していた。患者の病歴には、骨盤の軟組織腫瘍、及び一年間の化学療法が含まれる。家族歴には、結核、肺癌、及びアテローム硬化型の冠状動脈疾患が含まれる。
PROSN0T18	pINCY	ライブラリは、58歳白人男性の根治膀胱切除及び根治前立腺切除、胃フィステル形成の際に採取した病変前立腺組織から単離したRNAを用いて作製した。病理学的には、腺線維筋腫性過形成であり、この組織はグレード3の移行性細胞癌に関連する。患者の病歴には、口咽炎及び気腫があった。家族歴には、急性心筋梗塞及びアテローム硬化性冠状動脈疾患、II型糖尿病があった。
PROSTUS23	pINCY	このサブトラクトされた前立腺組織は、プールされていた前立腺ライブラリよりの 1×10^6 個のクローンで、2ラウンドの差し引きハイブリダゼーションにかけられた、プールされた前立腺組織ライブラリからの独立したクローン 1×10^6 個から作製された。サブトラクションのための開始ライブラリは、リンパ節切除を伴う前立腺切除の際に58歳(A)、61歳(B)、66歳(C)、及び68歳(D)の白人男性から採取した前立腺腫瘍から単離したmRNAを用いた4つの前立腺腫瘍ライブラリから得られた等しい数のクローンをプールして作製した。病理学的には、すべてのドナーが腺癌を示した。病歴には、ドナーAに高値のPSA、硬化、ニコチン中毒；ドナーBに高値のPSA、硬化、前立腺肥大、腎不全、変形性関節症、腎動脈狭窄、良性HTN、血小坂減少症、高脂血症、喫煙/飲酒及びC型肝炎(保有者)；ドナーCに高値のPSA、硬化、及び喫煙；またドナーDに高値のPSA、硬化、高コレステロール血症、及び腎臓結石が含まれる。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、3人の異なるドナーからの前立腺組織、前立腺上皮細胞、及び前立腺の間質の線維芽細胞から得られた3つの前立腺組織ライブラリからの等しい数のcDNAクローンをプールして作製した。サブトラクティブハイブリダイゼーション条件は、Swaroop 他(1991) Nucleic Acids Res. 19:1954 及びBonaldo 他 Genome Research (1996) 6:791の方法論に基づいた。
SINTNOT18	pINCY	ライブラリは、59歳の男性から得られた小腸組織から単離したRNAを用いて作製した
COLENOR03	PCDNA2.1	ライブラリは、自動車事故で死亡した13歳の白人女性から採取した腸上皮組織から単離したRNAで作製した。

10

20

30

【表 2 2】

プログラム名	説明	引用文献	パラメーター閾値
ABI FACTURA	核酸配列においてベクター配列を除去して不定の塩基をマスクするプログラム。	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finder は、アミノ酸または核酸配列の比較及び注釈付け (annotation) に有用である。	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler	核酸配列を構築するプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	Basic Local Alignment Search Tool 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。	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. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E 値=1.0E-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 シグナルペプチドヒット: スコア=0 以上

表7-2

プログラム名	説明	引用文献	パラメータ-閾値
ProfileScan	Prosite で定義された配列パターンと一致するタンパク質配列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 ≧ 特定の Prosite モチーフに対する GCG 指定 "HIGH" 値 通常、スコア = 1.4-2.1
Phred	高い感度及び確率で自動配列決定機のトレースを調べる塩基読み出しアルゴリズムである。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的なインプリメンテーションに基づく SWAT や CrossMatch を含む Phils Revised Assembly プログラムであって、配列相同性の検索及び DNA 配列の構築に有用である。	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; Green, P., University of Washington, Seattle, WA.	スコア = 120 以上 一致長さ = 56 以上
Consed	Phrap で構築したものの表示及び編集をするためのグラフィックツールである。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキャンして分泌シグナルペプチドの存在を調べる重み付けマトリクス解析プログラムである。	Nielson, H. 他 (1997) Protein Engineering 10: 1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	スコア = 3.5 以上
TMAP	蛋白配列での膜貫通セグメントの描写及び配向の決定のための、加重マトリクスを用いたプログラムである。	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	蛋白配列での膜貫通セグメントの描写及び配向の決定のための、隠された Markov モデル (HMM) を用いたプログラムである。	Sonnhammer, E.L. 他 (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et 他, es., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	Prosites 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 Acid Res. 25: 217-221; Wisconsin Package Program Manual, versinn 9. naze M51-59. Genetics	

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
28 June 2004 (28.06.2004)

PCT

(10) International Publication Number
WO 01/46258 A2

- (51) International Patent Classification: C07K 14/705
- (21) International Application Number: PCT/US00/33095
- (22) International Filing Date: 22 December 2000 (22.12.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:

60/172,000	23 December 1999 (23.12.1999)	US
60/176,983	14 January 2000 (14.01.2000)	US
60/177,332	21 January 2000 (21.01.2000)	US
60/178,572	28 January 2000 (28.01.2000)	US
60/179,758	2 February 2000 (02.02.2000)	US
60/181,625	10 February 2000 (10.02.2000)	US

(US) LU, **Dyung, Anna M.** [US/US], 233 Coy Drive, San Jose, CA 95123 (US); **YANG, Junming** [CN/CN], 7125 Bark Lane, San Jose, CA 95129 (US); **REDDY, Roopa** [IN/US], 1253 W. McKinley Avenue, #5, Sunnyvale, CA 94086 (US); **LAU, Fred** [US/US], 2382 Liss Drive, Santa Clara, CA 95054 (US); **HILLMAN, Jennifer L.** [US/US], 230 Monroe Drive, #17, Mountain View, CA 94040 (US); **AZIZAL, Yalda** [US/US], 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US); **YUE, Henry** [US/US], 826 Lois Avenue, Sunnyvale, CA 94087 (US); **NGUYEN, Daniel B.** [US/US], 1403 Ridgewood Drive, San Jose, CA 95118 (US); **YAO, Monique, G.** [US/US], 111 Frederick Court, Mountain View, CA 94043 (US); **GANDHI, Ameena, R.** [US/US], 837 Roble Avenue, #1, Menlo Park, CA 94025 (US); **TANG, Y., Tom** [CN/US], 4230 Ranwick Court, San Jose, CA 95118 (US); **KHAN, Farrah, A.** [IN/CN], 333 Escuela Avenue, #221, Mountain View, CA 94040 (US).



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

US	60/172,000 (CIP)
Filed on	23 December 1999 (23.12.1999)
US	60/176,983 (CIP)
Filed on	14 January 2000 (14.01.2000)
US	60/177,332 (CIP)
Filed on	21 January 2000 (21.01.2000)
US	60/178,572 (CIP)
Filed on	28 January 2000 (28.01.2000)
US	60/179,758 (CIP)
Filed on	2 February 2000 (02.02.2000)
US	60/181,625 (CIP)
Filed on	10 February 2000 (10.02.2000)

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

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GR, GT, HK, HU, IL, IN, JP, KR, KZ, KP, KR, KZ, LC, LI, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NZ, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TH, TL, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW); Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, UZ, TM); European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, HU, IT, LI, 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.

WO 01/46258 A2

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

(72) Inventors; and Inventors/Applicants (for US only): BAUGHN, Mariah, R. [US/US], 14241 Sontago Road, San Leandro, CA 94577 (US); BURFORD, Neil [GB/US], 105 Wildwood Circle, Durham, CT 06422 (US); AI-YOUNG, Janice [US/US], 333 Golden Toyle Lane, Brisbane, CA 94005

(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 01/46258

PCT/US00/35095

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, and immunological disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of transporters and ion channels.

BACKGROUND OF THE INVENTION

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

Carrier proteins which transport a single solute from one side of the membrane to the other are called uniporters. In contrast, coupled transporters link the transfer of one solute with simultaneous or sequential transfer of a second solute, either in the same direction (symport) or in the opposite direction (antiport). For example, intestinal and kidney epithelium contains a variety of symporter systems driven by the sodium gradient that exists across the plasma membrane. Sodium moves into the cell down its electrochemical gradient and brings the solute into the cell with it. The sodium gradient that provides the driving force for solute uptake is maintained by the ubiquitous Na^+/K^+ ATPase system. Sodium-coupled transporters include the mammalian glucose transporter (SGLT1), iodide transporter (NIS), and multivitamin transporter (SMVT). All three transporters have twelve putative transmembrane segments, extracellular glycosylation sites, and cytoplasmically-oriented N- and C-termini. NIS plays a crucial role in the evaluation, diagnosis, and treatment of various thyroid pathologies because it is the molecular basis for radioiodide thyroid-imaging techniques and for specific targeting of radioisotopes to the thyroid gland (Levy, O. et al. (1997) Proc.

WO 01/46258

PCT/US00/35095

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).

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 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 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 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 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 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 specific and selective transporters for organic cations and organic anions in organs including the kidney, intestine and liver. Organic anion transporters are selective for

WO 01/46258

PCT/US00/35095

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.* 264:C761-C782; Price, N.T. et al. (1998) *Biochem. J.* 329:321-328; and Martindale, 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 (sulfonamide 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. Fatty acid transport protein (FATP), an integral membrane protein with four transmembrane segments,

WO 01/46258

PCT/US00/35095

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).

5 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
10 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 Mitochondrial energy transfer proteins signature; Online Mendelian Inheritance in Man (OMIM) *275000 Graves Disease).

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

Ion Channels

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

Ion Transporters

Ion transporters generate and maintain the resting electrical potential of a cell. Utilizing the energy derived from ATP hydrolysis, they transport ions against the ion's concentration gradient. These transmembrane ATPases are divided into three families. The phosphorylated (P) class ion

WO 01/46258

PCT/US00/35095

transporters, including Na⁺-K⁺ ATPase, Ca²⁺-ATPase, and H⁺-ATPase, are activated by a phosphorylation event. P-class ion transporters are responsible for maintaining resting potential distributions such that cytosolic concentrations of Na⁺ and Ca²⁺ are low and cytosolic concentration of K⁺ is high. The vacuolar (V) class of ion transporters includes H⁺ pumps on intracellular organelles, such as lysosomes and Golgi. V-class ion transporters are responsible for generating the low pH within the lumen of these organelles that is required for function. The coupling factor (F) class consists of H⁺ pumps in the mitochondria. F-class ion transporters utilize a proton gradient to generate ATP from ADP and inorganic phosphate (P_i).

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

The resting potential of the cell is utilized in many processes involving carrier proteins and 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., acetylcholine-, serotonin-, and glutamate-gated cation channels, and GABA- and glycine-gated

WO 01/46258

PCT/US00/35095

chloride channels) open their pores in the presence of a specific ion, nucleotide, or neurotransmitter. The gating properties of a particular ion channel (i.e., its threshold for and duration of opening and closing) are sometimes modulated by association with auxiliary channel proteins and/or post-translational modifications, such as phosphorylation.

5 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. 10 274:6330-6335).

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

Voltage-gated Na^+ and K^+ channels are necessary for the function of electrically excitable cells, 20 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 25 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 30 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 an integral membrane glycoprotein that contains an extracellular Ig domain, and its association with α and $\beta 1$

WO 01/46258

PCT/US00/35095

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 Ca²⁺ and cAMP. In non-excitabile tissue, K⁺ channels are involved in protein synthesis, control of endocrine secretions, and the maintenance of osmotic equilibrium across membranes. In neurons and other excitable cells, in addition to regulating action potentials and repolarizing membranes, K⁺ channels are responsible for setting 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, J. and T. Jogle (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 01/46258

PCT/US00/35095

related gene (HERG) associated with long QT, a cardiac dysrhythmia syndrome (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).

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 pacemaker
10 activity, seizures and epilepsy, and insulina regulation (Doupoik, 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 channel.
25 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).

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
30 Na⁺, K⁺/Cl⁻ cotransporter, accumulating in the cell above its electrochemical equilibrium concentration. Secretion of Cl⁻ from the apical surface, in response to hormonal stimulation, leads to flow of Na⁺ and water into the secretory lumen. The cystic fibrosis transmembrane conductance regulator (CFTR) is a chloride channel encoded by the gene for cystic fibrosis, a common fatal genetic disorder in humans. CFTR is a member of the ABC transporter family, and is composed of two domains each consisting of

WO 01/46258

PCT/US00/35095

six transmembrane domains followed by a nucleotide-binding site. Loss of CFTR function decreases transepithelial water secretion and, as a result, the layers of mucus that coat the respiratory tree, pancreatic ducts, and intestine are dehydrated and difficult to clear. The resulting blockage of these sites leads to pancreatic insufficiency, "meconium ileus", and devastating "chronic obstructive pulmonary disease" (Al-Awqati, Q. et al. (1992) *J. Exp. Biol.* 172:245-266).

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

Ligand-gated channels open their pores when an extracellular or intracellular mediator binds to the channel. Neurotransmitter-gated channels are channels that open when a neurotransmitter binds to their extracellular domain. These channels exist in the postsynaptic membrane of nerve or muscle cells. There are two types of neurotransmitter-gated channels. Sodium channels open in response to excitatory neurotransmitters, such as acetylcholine, glutamate, and serotonin. This opening causes an influx of Na^+ and produces the initial localized depolarization that activates the voltage-gated channels and starts the action potential. Chloride channels open in response to inhibitory neurotransmitters, such as γ -aminobutyric acid (GABA) and glycine, leading to hyperpolarization of the membrane and the subsequent generation of an action potential. Neurotransmitter-gated ion channels have four transmembrane domains and probably function as pentamers (Jentsch, *supra*). Amino acids in the second transmembrane domain appear to be important in determining channel permeation and selectivity (Sather, W.A. et al. (1994) *Curr. Opin. Neurobiol.* 4:313-323).

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

WO 01/46258

PCT/US00/35095

Curr. Opin. Neurobiol. 8:321-329).

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

The activity of other types of ion channel proteins may also be modulated by a variety of intracellular signalling proteins. Many channels have sites for phosphorylation by one or more protein kinases including protein kinase A, protein kinase C, tyrosine kinase, and casein kinase II, all of which regulate ion channel activity in cells. Kir channels are activated by the binding of the $\text{G}\beta\gamma$ subunits of heterotrimeric G-proteins (Reimann, F. 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 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.J. and D.S. Bradt (1998) Cell 93:495-498).

Disease Correlation

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

Human diseases caused by mutations in ion channel genes include disorders of skeletal muscle, cardiac muscle, and the central nervous system. Mutations in the pore-forming subunits of sodium and chloride channels cause myotonia, a muscle disorder in which relaxation after voluntary contraction is delayed. Sodium channel myotonias have been treated with channel blockers. Mutations in muscle

WO 01/46258

PCT/US00/35095

sodium and calcium channels cause forms of periodic paralysis, while mutations in the sarcoplasmic calcium release channel, T-tubule calcium channel, and muscle sodium channel cause malignant hyperthermia. Cardiac arrhythmia disorders such as the long QT syndromes and idiopathic ventricular fibrillation are caused by mutations in potassium and sodium channels (Cooper, E.C. and L.Y. Jan (1998) Proc. Natl. Acad. Sci. USA 96:4759-4766). All four known human idiopathic epilepsy genes code for ion channel proteins (Berkovic, S.F. and I.E. Scheffer (1999) Curr. Opin. Neurology 12:177-182). Other neurological disorders such as ataxias, hemiplegic migraine and hereditary deafness can also result from mutations in ion channel genes (Jen, J. (1999) Curr. Opin. Neurobiol. 9:274-280; Cooper, *supra*).

10 Ion channels have been the target for many drug therapies. Neurotransmitter-gated channels have been targeted in therapies for treatment of insomnia, anxiety, depression, and schizophrenia. Voltage-gated channels have been targeted in therapies for arrhythmia, ischemic stroke, head trauma, and neurodegenerative disease (Taylor, C.P. and L.S. Narasimhan (1997) Adv. Pharmacol. 39:47-98). Various classes of ion channels also play an important role in the perception of pain, and thus are

15 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

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

25 immunosuppressants (Calahan, 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, and immunological disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of

30 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,"

WO 01/46258

PCT/US00/35095

"TRICH-5," "TRICH-6," "TRICH-7," "TRICH-8," "TRICH-9," "TRICH-10," "TRICH-11,"
"TRICH-12," "TRICH-13," "TRICH-14," "TRICH-15," "TRICH-16," "TRICH-17," "TRICH-18,"
"TRICH-19," "TRICH-20," "TRICH-21," "TRICH-22," "TRICH-23," "TRICH-24," "TRICH-25,"
"TRICH-26," and "TRICH-27." In one aspect, the invention provides an isolated polypeptide

5 comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence
selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence
having at least 90% sequence identity to an amino acid sequence selected from the group consisting of
SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group
10 consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected
from the group consisting of SEQ ID NO:1-27. In one alternative, the invention provides an isolated
polypeptide comprising the amino acid sequence of SEQ ID NO:1-27.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising an
amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the
group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least
15 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-
27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of
SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group
consisting of SEQ ID NO:1-27. In one alternative, the polynucleotide encodes a polypeptide selected
from the group consisting of SEQ ID NO:1-27. In another alternative, the polynucleotide is selected
20 from the group consisting of SEQ ID NO:28-54.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter
sequence operably linked to a polynucleotide encoding a polypeptide comprising an amino acid
sequence selected from the group consisting of a) an amino acid sequence selected from the group
consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90%
25 sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c)
a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID
NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group
consisting of SEQ ID NO:1-27. In one alternative, the invention provides a cell transformed with the
recombinant polynucleotide. In another alternative, the invention provides a transgenic organism
30 comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising an amino acid
sequence selected from the group consisting of a) an amino acid sequence selected from the group
consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90%
sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c)

WO 01/46258

PCT/US00/35095

a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide

5 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 comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid

10 sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27.

The invention further provides an isolated polynucleotide comprising a polynucleotide sequence

15 selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous

20 nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence

selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, b) a naturally occurring polynucleotide sequence having at least 90% sequence

25 identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions

30 whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said

WO 01/46258

PCT/US00/35095

target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to 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.

10 The invention further provides a composition comprising an effective amount of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment the composition.

20 The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. 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.

30 Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising an amino acid sequence selected from the group consisting

WO 01/46258

PCT/US00/35095

of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. 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 comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. 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 comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. 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.

WO 01/46258

PCT/US00/35095

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:28-54, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered
5 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 comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID
10 NO:28-54, ii) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific
15 hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, ii) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, iii) a
20 polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological
25 sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

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

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

Table 3 shows structural features of each polypeptide sequence, including predicted motifs and
35 domains, along with the methods, algorithms, and searchable databases used for analysis of each

WO 01/46258

PCT/US00/35095

polypeptide.

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

Table 5 shows the representative cDNA library for each polynucleotide of the invention.

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

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

10 DESCRIPTION OF THE INVENTION

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

15 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.

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

25 All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

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

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of TRICH. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other

WO 01/46258

PCT/US00/35095

compound or composition which modulates the activity of TRICH either by directly interacting with TRICH or by acting on components of the biological pathway in which TRICH participates.

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

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

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

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of

WO 01/46258

PCT/US00/35095

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.

5 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
10 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
15 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
20 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
25 introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

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

WO 01/46258

PCT/US00/35095

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

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

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

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

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

WO 01/46258

PCT/US00/35095

Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

5

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.

10 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.

15 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.

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

A fragment of SEQ ID NO:28-54 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:28-54, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:28-54 is useful, for 35 example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:28-54 from related polynucleotide sequences. The precise length of a fragment of SEQ

WO 01/46258

PCT/US00/35095

ID NO:28-54 and the region of SEQ ID NO:28-54 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:1-27 is encoded by a fragment of SEQ ID NO:28-54. A fragment of SEQ ID NO:1-27 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-27. For example, a fragment of SEQ ID NO:1-27 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-27. The precise length of a fragment of SEQ ID NO:1-27 and the region of SEQ ID NO:1-27 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

10 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.

15 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.

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

25 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.

30 Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other

WO 01/46258

PCT/US00/35095

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

Matrix: BLOSUM62

Reward for match: 1

10 *Penalty for mismatch: -2*

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

15 *Filter: on*

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

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

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

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment

WO 01/46258

PCT/US00/35095

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 5 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

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

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

WO 01/46258

PCT/US00/35095

annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

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

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

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

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

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of

WO 01/46258

PCT/US00/35095

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

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of TRICH which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of TRICH which is useful in any of the antibody production methods disclosed herein or known in the art.

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

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

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

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

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

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

"Post-translational modification" of an TRICH may involve lipilation, 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

WO 01/46258

PCT/US00/35095

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 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that

WO 01/46258

PCT/US00/35095

hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing
5 primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

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

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

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

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

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose
30 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.

WO 01/46258

PCT/US00/35095

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.

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

WO 01/46258

PCT/US00/35095

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 95% or at least 98% 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 least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human 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, and immunological disorders.

WO 01/46258

PCT/US00/35095

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

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

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

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:28-54 or that distinguish between SEQ ID NO:28-54 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages

WO 01/46258

PCT/US00/35095

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.

5 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, 6813453111 is the identification number of an Incyte cDNA sequence, and ADRETUR01 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., 70207988V1). Alternatively, the identification numbers in column 5 may refer to
10 GenBank cDNAs or ESTs (e.g., g1947104) 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, GNN.g6554406_006 is the identification number of a Genscan-predicted coding sequence, with g6554406 being the GenBank identification number of the sequence to which Genscan was applied. The Genscan-predicted coding
15 sequences may have been edited prior to assembly. (See Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon-stretching" algorithm. (See Example V.) In some cases, Incyte cDNA coverage redundant
20 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
25 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
30 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:28-54, which encodes TRICH. The polynucleotide sequences of SEQ ID NO:28-54, as presented in the Sequence Listing, embrace the equivalent RNA sequences,

WO 01/46258

PCT/US00/35095

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:28-54 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:28-54.

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

15

Although nucleotide sequences which encode TRICH and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring TRICH under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding TRICH or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding TRICH and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

20

25

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.

30

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

WO 01/46258

PCT/US00/35095

hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:28-54 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in

5 "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 (PE Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or

10 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 (PE Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system

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

20 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., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.)

25 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.

30 (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo

WO 01/46258

PCT/US00/35095

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

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

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

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

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

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number

WO 01/46258

PCT/US00/35095

5,837,458; Chang, C.-C. et al. (1999) *Nat. Biotechnol.* 17:793-797; Christians, F.C. et al. (1999) *Nat. Biotechnol.* 17:259-264; and Cramer, A. et al. (1996) *Nat. Biotechnol.* 14:315-319) to alter or improve the biological properties of TRICH, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

15 In another embodiment, sequences encoding TRICH may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) *Nucleic Acids Symp. Ser.* 7:215-223; and Horn, T. et al. (1980) *Nucleic Acids Symp. Ser.* 7:225-232.) 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 (PE 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.

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

30 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

WO 01/46258

PCT/US00/35095

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.)

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); 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., T1 or pBR322 plasmids), or animal cell systems. (See, e.g., Sambrook, *supra*; Ausubel, *supra*; Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509; Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945; Takamatsu, N. (1987) *EMBO J.* 6:307-311; *The McGraw Hill Yearbook of Science and Technology* (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; and Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al. (1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344; Buller, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al. (1994) *Mol. Immunol.* 31(3):219-226; and Verma, I.M. and N. Somia (1997) *Nature* 389:239-242.)

WO 01/46258

PCT/US00/35095

The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding TRICH. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding TRICH can be achieved using a multifunctional *E. coli* vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding TRICH into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for *in vitro* transcription, diideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509.) When large quantities of TRICH are needed, e.g. for the production of antibodies, vectors which direct high level expression of TRICH may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

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 PGII promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *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.*

WO 01/46258

PCT/US00/35095

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 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; Colbero-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *oppB* 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

WO 01/46258

PCT/US00/35095

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.

5 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.

10 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.)

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

20 Host cells transformed with nucleotide sequences encoding TRICH may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing

WO 01/46258

PCT/US00/35095

polynucleotides which encode TRICH may be designed to contain signal sequences which direct secretion of TRICH through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and W138) 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 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), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunofluorescence 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

WO 01/46258

PCT/US00/35095

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 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)

WO 01/46258

PCT/US00/35095

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 5,175,583 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
5 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).
10 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
15 human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding TRICH can also be used to create "knockin" humanized animals
20 (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
25 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. Therefore, TRICH appears to play a
30 role in transport, neurological, muscle, and immunological 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

WO 01/46258

PCT/US00/35095

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, 5 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, bradycardia, tachyarrhythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline 10 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 15 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 disease, and Fanconi disease; a 20 neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural 25 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 30 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), 35 akathisia, amnesia, cataplexy, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses,

WO 01/46258

PCT/US00/35095

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,

5 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); and

10 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 (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema,

15 episodic lymphopenia with lymphocytotoxicus, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypercosinophilia, 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

20 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.

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

25 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.

30 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

WO 01/46258

PCT/US00/35095

disorders include, but are not limited to, those transport, neurological, muscle, and immunological 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.

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

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

15 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 25 such as aluminum hydroxide, and surface active substances such as lysolactin, 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 30 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.

WO 01/46258

PCT/US00/35095

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., Köhler, G. et al. (1975) *Nature* 256:495-497; Köhler, 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.* 6:109-120.)

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

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

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

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such 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 techniques may be used to assess the affinity of antibodies for TRICH. Affinity is expressed as an association

WO 01/46258

PCT/US00/35095

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 for TRICH. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular TRICH epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the TRICH-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of TRICH, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of TRICH-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, *supra*, and Coligan et al. *supra*.)

In another embodiment of the invention, the polynucleotides encoding TRICH, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding TRICH. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding TRICH. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retroviruses and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, *supra*; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other

WO 01/46258

PCT/US00/35095

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; Boarko, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

- 5 In another embodiment of the invention, polynucleotides encoding TRICH may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency
- 10 (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. Sonla (1997) *Nature* 389:239-242)), (ii)
- 15 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*
- 20 *brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in TRICH expression or regulation causes disease, the expression of TRICH from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

- In a further embodiment of the invention, diseases or disorders caused by deficiencies in
- 25 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-
- 30 217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J.-L. and H. Récapon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of TRICH include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF,

WO 01/46258

PCT/US00/35095

PTE1-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. U.S.A. 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 ecodycson-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/tapamycin 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.

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. U.S.A. 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; Döll, 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

WO 01/46258

PCT/US00/35095

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. U.S.A.* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver
5 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
(Csete, M.E. et al. (1995) *Transplantation* 27:263-268). Potentially useful adenoviral vectors are
10 described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"),
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. Sonja (1997) *Nature* 18:389:239-242, both
incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver
15 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
20 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
25 the control of the appropriate promoter for purposes including human gene therapy. Also taught by this
patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22.
For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev.
Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus
sequences, the generation of recombinant virus following the transfection of multiple plasmids
30 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,

WO 01/46258

PCT/US00/35095

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 application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of TRICH into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, *Molecular and Immunologic Approaches*, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding 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

WO 01/46258

PCT/US00/35095

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.

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

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

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary

WO 01/46258

PCT/US00/35095

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
5 may comprise, for example, an intact or permeabilized cell, or an *in vitro* cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding 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
10 forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a *Schizosaccharomyces pombe* gene expression
15 system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) *Nucleic Acids Res.* 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) *Biochem. Biophys. Res. Commun.* 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide
20 sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Braice, 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.
25 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
30 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

WO 01/46258

PCT/US00/35095

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, 5 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 10 formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle 15 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 20 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 25 administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example TRICH or fragments thereof, antibodies of TRICH, and agonists, antagonists or inhibitors of TRICH, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose 30

WO 01/46258

PCT/US00/35095

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

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

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

WO 01/46258

PCT/US00/35095

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.

5 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.

10 In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding TRICH or closely related molecules may be used to identify nucleic acid sequences which encode TRICH. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding TRICH, allelic variants, or related

15 sequences. Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the TRICH encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:28-54 or from genomic sequences including promoters, enhancers, and introns of the TRICH gene.

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

25 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

WO 01/46258

PCT/US00/35095

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,

5 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,

10 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 disease, and Fanconi disease; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's

15 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

20 Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tubercous sclerosis, cerebellarretinal 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

25 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,

30 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

35 syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, pheochromocytoma, and

WO 01/46258

PCT/US00/35095

- myopathies including encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, ophthalmoplegia, and acid maltase deficiency (AMD, also known as Pompe's disease); and an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma,
- 5 atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's
- 10 thyroiditis, hyper eosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal
- 15 circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma. 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.
- 20 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
- 25 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.
- 30 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

WO 01/46258

PCT/US00/35095

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.

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

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

Additional diagnostic uses for oligonucleotides designed from the sequences encoding TRICH may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced *in vitro*. Oligomers will preferably contain a fragment of a polynucleotide encoding TRICH, or a fragment of a polynucleotide complementary to the polynucleotide encoding TRICH, and
20 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
25 substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding TRICH are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal
30 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 methods, termed *in silico* SNP

WO 01/46258

PCT/US00/35095

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

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

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

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

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seibamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of

WO 01/46258

PCT/US00/35095

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

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

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with *in vitro* model systems and preclinical evaluation of
10 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 (Nowaysir, 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
15 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
20 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
25 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
30 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.

WO 01/46258

PCT/US00/35095

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 separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

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

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

WO 01/46258

PCT/US00/35095

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Sialon, 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 chromosomal 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.)

WO 01/46258

PCT/US00/35095

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

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

In another embodiment of the invention, TRICH, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between TRICH and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with TRICH, or fragments thereof, and washed. Bound TRICH is then detected by methods well known in the art. Purified TRICH can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding TRICH specifically compete with a test compound for binding TRICH. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with TRICH.

WO 01/46258

PCT/US00/35095

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.

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

The disclosures of all patents, applications and publications, mentioned above and below, in particular U.S. Ser. No. 60/172,000, U.S. Ser. No. 60/176,083, U.S. Ser. No. 60/177,332, U.S. Ser. No. 60/178,572, U.S. Ser. No. 60/179,758, and U.S. Ser. No. 60/181,625, are expressly incorporated by reference herein.

EXAMPLES

15 I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIHESQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)⁺ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSOFT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-5.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the

WO 01/46258

PCT/US00/35095

appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPIAROSE CL2B, or SEPIAROSE 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.,

- 5 PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pCDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY (Incyte Genomics, Palo Alto CA). Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

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

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-
20 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.
25 Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (PE Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI
30 PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (PE 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

WO 01/46258

PCT/US00/35095

sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DDMO, PRODOM, and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLJMP, and HMMR. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences which were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DDMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and

WO 01/46258

PCT/US00/35095

polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:28-54. 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

5 Putative transporters and ion channels were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpr1 and gbhg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) *J. Mol. Biol.* 268:78-94, and Burge, C. and S. Karlin (1998) *Curr. Opin. Struct. Biol.* 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode transporters and ion channels, the encoded polypeptides were analyzed by querying against PFAM models for transporters and ion channels. Potential transporters and ion channels were also identified by homology to Incyte cDNA sequences that had been annotated as transporters and ion channels. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpr1 public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data "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

WO 01/46258

PCT/US00/35095

the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpro 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 FRICH Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:28-54 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:28-54 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

WO 01/46258

PCT/US00/35095

of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

Map locations are represented by ranges, or intervals, or 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 G6u6thon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GenoMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel (1995) *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIPSEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

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

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter

WO 01/46258

PCT/US00/35095

of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding TRICH are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system: exocrine glands; genitalia, female; genitalia, male; germ cells; hematologic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding TRICH. cDNA sequences and cDNA library/tissue information are found in the LITSEQ 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 Mg²⁺, (NH₄)₂SO₄, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer

WO 01/46258

PCT/US00/35095

pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with *Cvi*I choleia virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

WO 01/46258

PCT/US00/35095

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:28-54 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments.

- 5 Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [γ - 32 P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^3 counts per
- 10 minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

- The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16
- 15 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

X. Microarrays

- 20 The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing. See, e.g., Baldeschwieler, *supra.*), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schna (1999), *supra.*). Suggested
- 25 substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schna, M. et al. (1995) *Science* 270:467-470; Shalon, D. et al. (1996) *Genome Res.* 6:639-645; Marshall, A. and J. Hodgson (1998) *Nat. Biotechnol.*
- 30 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

WO 01/46258

PCT/US00/35095

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

Microarray Preparation

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

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

WO 01/46258

PCT/US00/35095

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

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

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

Hybridization

Hybridization reactions contain 9 µl of sample mixture consisting of 0.2 µg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and 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 µl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

Detection

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

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

WO 01/46258

PCT/US00/35095

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

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

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated 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).

20 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

30 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 (tac)* hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory

WO 01/46258

PCT/US00/35095

element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express TRICH upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of TRICH in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant *Autographica californica* nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding TRICH by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases.

10 Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, TRICH is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from TRICH at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10 and 16). Purified TRICH obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, and XVIII, where applicable.

25 XIII. Functional Assays

TRICH function is assessed by expressing the sequences encoding TRICH at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μ g of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μ g of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the

WO 01/46258

PCT/US00/35095

recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of

5 fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies;

10 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 Ormszol, M.G. (1994) *Flow Cytometry*, Oxford, New York NY.

The influence of TRICH on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding TRICH and either CD64 or CD64-GFP. CD64 and

15 CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding TRICH and other genes of interest can be analyzed by northern analysis or

20 microarray techniques.

XIV. Production of TRICH Specific Antibodies

TRICH substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

25 Alternatively, the TRICH amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. 11.)

30 Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (PE Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-TRICH

WO 01/46258

PCT/US00/35095

activity by, for example, binding the peptide or TRICH to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring TRICH Using Specific Antibodies

Naturally occurring or recombinant TRICH is substantially purified by immunoaffinity chromatography using antibodies specific for TRICH. An immunoaffinity column is constructed by covalently coupling anti-TRICH antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing TRICH are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of TRICH (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/TRICH binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and TRICH is collected.

XVI. Identification of Molecules Which Interact with TRICH

Molecules which interact with TRICH may include transporter substrates, agonists or antagonists, modulatory proteins such as G $\beta\gamma$ proteins (Reimann, *supra*) or proteins involved in TRICH localization or clustering such as MAGUKs (Craven, *supra*). TRICH, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled TRICH, washed, and any wells with labeled TRICH complex are assayed. Data obtained using different concentrations of TRICH are used to calculate values for the number, affinity, and association of TRICH with the candidate molecules.

Alternatively, proteins that interact with TRICH are isolated using the yeast 2-hybrid system (Fields, S. and O. Song (1989) *Nature* 340:245-246). TRICH, or fragments thereof, are expressed as fusion proteins with the DNA binding domain of Gal4 or LexA and potential interacting proteins are expressed as fusion proteins with an activation domain. Interactions between the TRICH fusion protein and the reconstitutes 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 Niehhammer, 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

WO 01/46258

PCT/US00/35095

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 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 pBF, and 0.5-5 ng of mRNA is injected into mature stage IV oocytes. Injected oocytes are incubated at 18°C for 1-5 days. Inside-out macropatches are excised into an intracellular solution containing 116 mM K-gluconate, 4 mM KCl, and 10 mM Hepes (pH 7.2). The intracellular solution is supplemented with varying concentrations of the TRICH mediator, such as cAMP, cGMP, or Ca^{+2} (in the form of $CaCl_2$), where appropriate. Electrode resistance is set at 2-5 M Ω and electrodes are filled with the intracellular solution lacking mediator. Experiments are performed at room temperature from a holding potential of 0 mV. Voltage ramps (2.5 s) from -100 to 100 mV are acquired at a sampling frequency of 500 Hz. Current measured is proportional to the activity of TRICH in the assay.

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

WO 01/46258

PCT/US00/35095

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, 50µg/ml gentamycin, pH 7.8) to allow expression of TRICH. Oocytes are then transferred to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM 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.

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 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

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 membrane (Vellicelebi, 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 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 oxonyl dyes such as DiBAC₄ (Molecular Probes). DiBAC₄ equilibrates between the extracellular solution and cellular sites according to the cellular membrane potential. The dye's fluorescence intensity is 20-fold greater when bound to hydrophobic intracellular sites, allowing detection of DiBAC₄ entry into the cell (Gonzalez, J.E. and P.A. Negulescu (1998) *Curr. Opin. Biotechnol.* 9:624-631). Candidate agonists or antagonists may be

WO 01/46258

PCT/US00/35095

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
5 be apparent to those skilled in the art without departing from the scope and spirit of the invention.
Although the invention has been described in connection with certain embodiments, it should be
understood that the invention as claimed should not be unduly limited to such specific embodiments.
Indeed, various modifications of the described modes for carrying out the invention which are obvious
to those skilled in molecular biology or related fields are intended to be within the scope of the following
10 claims.

WO 01/46258

PCT/US00/35095

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
1416107	1	1415107CDD1	28	1416107CBI
1684513	2	1682313CDD1	29	1684513CBI
2446438	3	2444638CDD1	30	2446438CBI
2817822	4	2817822CDD1	31	2817822CBI
4009329	5	4009329CDD1	32	4009329CBI
6618083	6	6618083CDD1	33	6618083CBI
7472002	7	7472002CDD1	34	7472002CBI
1811682	8	1811682CDD1	35	1811682CBI
3232922	9	3232922CDD1	36	3232922CBI
3558382	10	3558382CDD1	37	3558382CBI
4250093	11	4250093CDD1	38	4250093CBI
7086893	12	7086893CDD1	39	7086893CBI
5674114	13	5674114CDD1	40	5674114CBI
1254635	14	1254635CDD1	41	1254635CBI
1670955	15	1670955CDD1	42	1670955CBI
1859560	16	1859560CDD1	43	1859560CBI
5530154	17	5530154CDD1	44	5530154CBI
1391115	18	1391115CDD1	45	1391115CBI
1702940	19	1702940CDD1	46	1702940CBI
1703342	20	1703342CDD1	47	1703342CBI
1727529	21	1727529CDD1	48	1727529CBI
2289333	22	2289333CDD1	49	2289333CBI
2720334	23	2720334CDD1	50	2720334CBI
3038153	24	3038153CDD1	51	3038153CBI
3460979	25	3460979CDD1	52	3460979CBI
7472200	26	7472200CDD1	53	7472200CBI
	27		54	

WO 01/46258

PCT/US00/35095

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability Score	GenBank Homolog
1	1416107CD1	G7018605	1.9e-302	Glucose transporter [Rattus norvegicus] [Bjerskov, M. et al. (2000) J. Biol. Chem. 275:4607-4612]
2	1682513CD1	G5263196	1.4e-153	Stretch-inhibitable nonselective channel (SIC) [Rattus norvegicus] (Cloning of a stretch-inhibitable nonselective cation channel. J. Biol. Chem. 1999 Mar 5;274(10):5330-5335)
3	2484438CD1	G4589141	0	Vanilloid receptor-like protein 1 [Homo sapiens] (A capsaicin-receptor homologue with a high threshold for pungent sensation, very low affinity to capsaicin) [Homo sapiens] (GenBank accession number U18548.1)
5	4009139CD1	G3873983	1.9e-64	Facilitative glucose transporter family member GLUT9 [Homo sapiens] (Phay, J. E. et al. (2000) Genomics 66:217-220)
6	6613083CD1	G9230651	4.7e-268	Facilitative glucose transporter family member GLUT9 [Homo sapiens] (Phay, J. E. et al. (2000) Genomics 66:217-220)
7	7472020CD1	G433940	0	Aorta cMG channel (aMG) [Oryctolagus cuniculus] (Primary structure and functional expression of a cyclic nucleotide-gated channel from rabbit aorta. FEBS Lett. 1993 Aug 23;329(1-2):134-138)
8	1812692CD1	G3928756	4.5e-48	Transient receptor potential channel 7 [Homo sapiens] (Ragamine, X. et al. (1998) Molecular cloning of a novel putative Ca ²⁺ channel protein (TRPC7) highly expressed in brain. FEBS Lett. 1998 Jul 11;441(1-2):113-118)
9	3329392CD1	G3874275	3.2e-70	Similarity to yeast low-affinity glucose transporter HXT4 [Caenorhabditis elegans]
10	3358383CD1	G3004482	1.4e-163	Putative integral membrane transport protein [Rattus norvegicus] (Schomig, E. et al. (1998) Molecular cloning and characterization of two novel transport proteins from rat kidney. FEBS Lett. 425:73-86)
11	4250091CD1	G3880445	5.7e-16	VM106R.1 (similar to K ⁺ channel tetramerization domain) [Caenorhabditis elegans]
12	70064803CD1	G3874275	7.0e-84	Similarity to yeast low-affinity glucose transporter HXT4 [Caenorhabditis elegans]
13	70356768CD1	G182298	4.1e-54	Glucose transporter (GLUT5) [Rattus norvegicus] (Kavans, Y. et al. (1990) Human facilitative glucose transporters: Isolation, functional characterization, and gene localization of cDNAs encoding an isoform (GLUT5) expressed in small intestine, kidney, muscle, and adipose tissue and an unusual glucose transporter pseudogene-like sequence (GLUT6). J. Biol. Chem. 265:13276-13282)

Table 2 (cont.)

Peptide SEQ ID NO:	Incites Peptide ID	GenBank ID NO:	Probability Score	GenBank Homolog
14	5674114CD1	55771352	1.3e-238	Inward rectifier potassium channel Kir2.4 [Homo sapiens; (Topart, C. et al. (1998) Kir2.4: a novel K ⁺ inward rectifier channel associated with motoneurons of cranial nerve nuclei. <i>J. Neurosci.</i> 18:4096-4105)]
15	1254635CD1	53953533	1.7e-210	Inwardly rectifying potassium channel Kir3.1 [Mus musculus] (Mourl, T. et al. (1998) Assignment of mouse inwardly rectifying potassium channel Kir3.1 to the distal region of chromosome 16 encodes a transient receptor (Homo sapiens) [Cazen, L. et al. (2000) <i>J. Biol. Chem.</i> 275:32027-32036]
17	1859560CD1	55834394	1.4e-101	Sulfate transporter [Drosophila melanogaster]
18	5510156CD1	54803004	3.5e-20	UDP-N-acetylglucosamine transporter [Homo sapiens] (Ishida, M. et al. (1999) Molecular cloning and functional expression of the human golgi UDP-N-acetylglucosamine transporter. <i>J. Biochem.</i> 125:68-77).
19	139115CD1	58131858	1.5e-49	Putative thymic stromal co-transporter, TSC09 [Mus musculus] (M2123.9. et al. (2000) Immunol. 14:315-319)
20	1702840CD1	55735224	2.5e-143	Calcium channel alpha 2 delta 3 [Homo sapiens]
21	1705342CD1	56003526	5.1e-08	Calcium channel alpha 2 subunit [Edelloura candida]
22	1727529CD1	54523890	0.0	M222 [Homo sapiens]
23	2288333CD1	54339333	5.6e-35	Putative amino acid transport protein (Arabidopsis thaliana)
24	2720354CD1	53875242	4.1e-38	Similar to mitochondrial carrier protein [Caenorhabditis elegans]
26	3460979CD1	51931644	4.7e-08	Membrane protein PMI precursor isolog (putative major facilitator, superfamily transporter) [Arabidopsis thaliana]
27	7472200CD1	52811254	2.8e-21	Anion chloride-sensitive Na ⁺ channel [Drosophila melanogaster] (Adams, C.M. et al. (1998) <i>J. Cell Biol.</i> 140:133-152)

Table 3

SEQ ID NO.	Inverte Polyptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential glycosylation sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	14416107CD1	477	S59 T2 T281 S430 T205	N349	Sugar transporter domain: A29-E474 Sugar transport protein signatures: V108-L174, L291-S350, G41-L51, L428-V143, Q267-F277, V375-L396, S368-T400 Glycophorin transporter signatures: E457-Y278, V375-S398, G439-Y459 Transmembrane domains: L259-A279, L291-K313, L320-Y339, Y438-F457 Glucose transporter signature: P19-N339	MOTIFS HMMER-PFAM BLIKFS- BLOCKS PROFILEScan SIGNALS- PRINTS HMMER
2	1562513CD1	498	S47 T131 S286 T367 S463 T67 S315 S382	N278 N411 N429	Transmembrane domains: S103-L119, V174-F189, L200-F216, F274-S283, F319-S333 Ankyrin repeat: K162-G194, F208-S243 Q393-F328 Transmembrane domains: L286-F405, I463-Y486, F538-S557, L623-I642	MOTIFS BLIKFS- PRINTS HMMER
3	2446438CD1	764	T64 T139 S23 S67 T105 S268 S339 S348 S353 S464 S468 S657 T652 S697 V720 T101 T115 S325 T414 Y110 Y227 T338 S157 R33 T71 S23 T50 S114 T163 T238 S258 S321 T750 S271 S273 S468 S514 S62 T132	S570	Potassium channel signature: K76-K95 Signal peptide: M1-C29 Sodium/calcium exchanger protein domain: T11-Q252, I431-P576 Transmembrane domains: T101-F121, T166-L199, L234-Y251, F392-E402, F392-R393, D393-H394	MOTIFS BLIKFS- PRINTS HMMER HMMER-PFAM HMMER
4	2817822CD1	255	S157 R33 T71 S23 T50 S114 T163 T238 S258 S321 T750 S271 S273 S468 S514 S62 T132			MOTIFS BLIKFS- PRINTS HMMER
5	4009329CD1	584	S271 S273 S468 S514 S62 T132	N60 N135		MOTIFS BLIKFS- PRINTS HMMER HMMER-PFAM HMMER

WO 01/46258

PCT/US00/35095

Table 3 (cont.)

SEQ ID NO:	Incyte ID poly(amide) ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
6	5618083CD1	416	M11 S4 S164 S274 T374 S16 S226 S289	M45 N61 N410	Signal peptide: M1-Q37 Sugar transporter domain: K30-Q416 Sugar transport proteins signatures: S39-Q49 S133-Q138 S298-Q308 Glucose transporter signatures: V288-Q309, I356-Q376 Transmembrane domains: V43-V59, I355-L475	NOTIFS SFCAN HMMER HMMER-PPAM BLIMPS- BLOCKS ELOCSS BLIMPS- SCAN BLIMPS- SCAN PALINUS HMMER
7	7472002CD1	664	S402 S40 S46 S93 T107 T213 T37 T381 T422 S466 T524 T591 S506 T624 T7 S35 T124 T208 S418 T448 Y648	N111 M379	Transmembrane region, cyclic nucleotide gated channel: Y215-Y440 Cytosolic nucleotide binding domain: G469-N525, S480-S476, G478-V501, G516-L525 Transmembrane domain: Y350-I375	NOTIFS HMMER-PPAM BLIMPS- BLOCKS
8	1812692CD1	242	T95 S35 S37 T124 S204 S221 S2 S9 S20 T21 S86 Y16	N15 N84	Protein malastatin chromosome transmembrane PD018035: Y117-E227	BLAST- PRODOM
9	3232992CD1	398	T107 S38 T41 S100 S139 T41 E303 S377 S393	M16L	Transmembrane domains: A217-Q242, L247-P264, L350-F360 Sugar transporter-proteins signature: E45-G94, V30-E96	HMMER NOTIFS BLIMPS- BLOCKS ProfileScan
10	3383833CD1	353	S137 S352 S409 T58 S60 S109 M133 S337 T433 S527 S187 S201 M226 S282 T123 T409	M39 M56 N62 M102 M107 M473	Transmembrane domains: F205-A222, M470-Y493, I590-T519 GLPT family of transporters: V151-D168	HMMER NOTIFS BLIMPS- BLOCKS PRODOM
					Organic transporter protein, renal anion transporter, cationic kidney specific, human PDB1520: M102-L144	BLAST- PRODOM

Table 3 (cont.)

SEQ ID NO:	Protein ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
11	4250091CD1	213	S2 S91 S172 S184 T17 T22 T137 S210 Y89	M189	Potassium channel signature: Q48-T67	MOTIFS BLIMPS- PRIMTS HMMER
12	70064803CD1	476	T365 S11 S364 S451 S222 T361 S390 S466		Transmembrane domains: V222-R239, G327-V350, M413-F432 Sugar transport proteins signature: L153-G202, V136-L204	MOTIFS BLIMPS- PRIMTS HMMER PROFILINSCAN
13	70355766CD1	246	S100 S233 S118 S215	M34 M50	Signal peptide: M1-Q27 Sugar transport proteins signature: L127-G176, A112-V178, T28-I38, M129-M147, M133-R158	MOTIFS HMMER- BLIMPS- PRIMTS HMMER BLAST- DOMO
14	5674114CD1	436	S11 T60 S154 S340 S362 T263 S276 S422 Y47	M195	Transmembrane domain: M163-L181 Sugar transport proteins: DM001351P46408; A113-C229 Inward rectifier potassium channel domain: V53-L394, R72-L116, P126-Q169, C170-V199, A204-Q238, D256-Y346, T358-E388 Transmembrane domain: M88-L114 Inward rectifier potassium channel: DM00481P2189; L134-F336 Inward rectifier potassium channel: PM001103; V53-Q372 KIR2.4 protein: A373-P436 PD063375; M1-P52	MOTIFS HMMER- BLIMPS- PRIMTS HMMER BLAST- DOMO HMMER BLAST- DOMO HMMER BLAST- DOMO HMMER BLAST- DOMO

WO 01/46258

PCT/US00/35095

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
15	L254635CD1	453	S208 T16 T99 S416 S23 T209 T216 T250 S375		Signal peptide: M1-R12 Transmembrane domains: F109-A131, V-82-I200 Inward rectifier potassium channel domain: I77-R199 Inward rectifier potassium ion channel family PD001403: K74-K83 Voltage-gated inward rectifier potassium channel BR9, KIR5.1, transmembrane PD03175, M36-E73 Inward rectifier potassium channel: PD09448 P52185[27-380], K62-E379 Inward rectifier potassium channel signature: P145-Q188, S188-R213, S223-R257, K310-C360, S371-M381 Transmembrane domains: S271-M381 F159-V179 L37-L58, A74-I93, L134-V154, Sensitive cotransporter, chloride, sodium: DM01178 SC06803 1-428: S32-L153	SRScan HMMER HMMER-PP2M BLAST-PRODOM BLAST-PRODOM BLAST-DOKO BLIIPS-PP2M HMMER BLAST-DOKO
16	L670595CD1	299	S18 T111 S263 S33 T187	M193 N208		

WO 01/46258

PCT/US00/35095

Table 3 (cont.)

SEQ ID NO.	Invertebrate Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences Domains and Motifs	Analytical Methods and Databases
17	1859360CD1	406	T96 T116 S298 S571 T572 S595 T245	N294	Transmembrane domains: L45-I63, T396-T421, M17-A468 Sulfate transporter family domain: M17-A468 Sulfate transporters protein signature: A24-E107, E23-E16 Sulfate transporter family domain: M17-A468 Transmembrane, permease P001255: M17-L465 Sulfate transport protein, transmembrane, permease P001121: L30-G143 Sulfate transporter: P001255 S64926 59-591: F30-W423 Sulfate transporters motif: P77-R88	HMMER HMMER-PFAM BLIPS- BLIPS- BLIPS- PRODOM
18	5530164CD1	324	S2 S139 Y115	N99 N100 N101 N232	Transmembrane domains: M182-P200, P242-I261, Y283-F302 Sugar transporter motif: L75-S91 Glucose transporter signature: M182-I202	HMMER
19	139116CD1	445	S64 S82 S77 S217 S424 S438 T150 S237 S443 Y23	N22 N30 N37 N127 N213 N233	Transmembrane domains: M182-P200, P242-I261, Y283-F302 Sugar transporter motif: L75-S91 Glucose transporter signature: M182-I202	HMMER
20	1702840CD1	337	T30 S167 T208 S306		Apolipoprotein L precursor, lipid transport, glycoprotein, signal, transmembrane domains: T30-T118, M142-F161, F170-I185, T102-T155	MOTIFS PAINS- PAINS- PRODOM
21	1703342CD1	273	T3 S63 T222 S446 S490 I40 S58 S519 S224		Ion transport protein domain: L95-L269 (Score: -132.1, E-value: 0.72)	HMMER HMMER-PFAM

WO 01/46258

PCT/US00/35095

Table 3 (cont.)

SEQ ID NO.	Incyte ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
22	1727529CD1	710	S31 S102 S115 T135 S304 S32 S218 S430 S431 W34 S573 S619 Y13	R29 N69 N155 M197 K298 K393 N405 M416 M678	Transmembrane domains: C38-Y38, V241-L266, W309-V336, F356-W375, F440-L458, T499-I522, L538-F618, L645-V663 ABC 3 transport family: S228-Q427 (Source: S32.P, E-value: 2.1) Am1, L1, L2, L3, L4, L5, L6, L7, L8, L9, L10, L11, L12, L13, L14, L15, L16, L17, L18, L19, L20, L21, L22, L23, L24, L25, L26, L27, L28, L29, L30, L31, L32, L33, L34, L35, L36, L37, L38, L39, L40, L41, L42, L43, L44, L45, L46, L47, L48, L49, L50, L51, L52, L53, L54, L55, L56, L57, L58, L59, L60, L61, L62, L63, L64, L65, L66, L67, L68, L69, L70, L71, L72, L73, L74, L75, L76, L77, L78, L79, L80, L81, L82, L83, L84, L85, L86, L87, L88, L89, L90, L91, L92, L93, L94, L95, L96, L97, L98, L99, L100, L101, L102, L103, L104, L105, L106, L107, L108, L109, L110, L111, L112, L113, L114, L115, L116, L117, L118, L119, L120, L121, L122, L123, L124, L125, L126, L127, L128, L129, L130, L131, L132, L133, L134, L135, L136, L137, L138, L139, L140, L141, L142, L143, L144, L145, L146, L147, L148, L149, L150, L151, L152, L153, L154, L155, L156, L157, L158, L159, L160, L161, L162, L163, L164, L165, L166, L167, L168, L169, L170, L171, L172, L173, L174, L175, L176, L177, L178, L179, L180, L181, L182, L183, L184, L185, L186, L187, L188, L189, L190, L191, L192, L193, L194, L195, L196, L197, L198, L199, L200, L201, L202, L203, L204, L205, L206, L207, L208, L209, L210, L211, L212, L213, L214, L215, L216, L217, L218, L219, L220, L221, L222, L223, L224, L225, L226, L227, L228, L229, L230, L231, L232, L233, L234, L235, L236, L237, L238, L239, L240, L241, L242, L243, L244, L245, L246, L247, L248, L249, L250, L251, L252, L253, L254, L255, L256, L257, L258, L259, L260, L261, L262, L263, L264, L265, L266, L267, L268, L269, L270, L271, L272, L273, L274, L275, L276, L277, L278, L279, L280, L281, L282, L283, L284, L285, L286, L287, L288, L289, L290, L291, L292, L293, L294, L295, L296, L297, L298, L299, L300, L301, L302, L303, L304, L305, L306, L307, L308, L309, L310, L311, L312, L313, L314, L315, L316, L317, L318, L319, L320, L321, L322, L323, L324, L325, L326, L327, L328, L329, L330, L331, L332, L333, L334, L335, L336, L337, L338, L339, L340, L341, L342, L343, L344, L345, L346, L347, L348, L349, L350, L351, L352, L353, L354, L355, L356, L357, L358, L359, L360, L361, L362, L363, L364, L365, L366, L367, L368, L369, L370, L371, L372, L373, L374, L375, L376, L377, L378, L379, L380, L381, L382, L383, L384, L385, L386, L387, L388, L389, L390, L391, L392, L393, L394, L395, L396, L397, L398, L399, L400, L401, L402, L403, L404, L405, L406, L407, L408, L409, L410, L411, L412, L413, L414, L415, L416, L417, L418, L419, L420, L421, L422, L423, L424, L425, L426, L427, L428, L429, L430, L431, L432, L433, L434, L435, L436, L437, L438, L439, L440, L441, L442, L443, L444, L445, L446, L447, L448, L449, L450, L451, L452, L453, L454, L455, L456, L457, L458, L459, L460, L461, L462, L463, L464, L465, L466, L467, L468, L469, L470, L471, L472, L473, L474, L475, L476, L477, L478, L479, L480, L481, L482, L483, L484, L485, L486, L487, L488, L489, L490, L491, L492, L493, L494, L495, L496, L497, L498, L499, L500, L501, L502, L503, L504, L505, L506, L507, L508, L509, L510, L511, L512, L513, L514, L515, L516, L517, L518, L519, L520, L521, L522, L523, L524, L525, L526, L527, L528, L529, L530, L531, L532, L533, L534, L535, L536, L537, L538, L539, L540, L541, L542, L543, L544, L545, L546, L547, L548, L549, L550, L551, L552, L553, L554, L555, L556, L557, L558, L559, L560, L561, L562, L563, L564, L565, L566, L567, L568, L569, L570, L571, L572, L573, L574, L575, L576, L577, L578, L579, L580, L581, L582, L583, L584, L585, L586, L587, L588, L589, L590, L591, L592, L593, L594, L595, L596, L597, L598, L599, L600, L601, L602, L603, L604, L605, L606, L607, L608, L609, L610, L611, L612, L613, L614, L615, L616, L617, L618, L619, L620, L621, L622, L623, L624, L625, L626, L627, L628, L629, L630, L631, L632, L633, L634, L635, L636, L637, L638, L639, L640, L641, L642, L643, L644, L645, L646, L647, L648, L649, L650, L651, L652, L653, L654, L655, L656, L657, L658, L659, L660, L661, L662, L663, L664, L665, L666, L667, L668, L669, L670, L671, L672, L673, L674, L675, L676, L677, L678, L679, L680, L681, L682, L683, L684, L685, L686, L687, L688, L689, L690, L691, L692, L693, L694, L695, L696, L697, L698, L699, L700, L701, L702, L703, L704, L705, L706, L707, L708, L709, L710, L711, L712, L713, L714, L715, L716, L717, L718, L719, L720, L721, L722, L723, L724, L725, L726, L727, L728, L729, L730, L731, L732, L733, L734, L735, L736, L737, L738, L739, L740, L741, L742, L743, L744, L745, L746, L747, L748, L749, L750, L751, L752, L753, L754, L755, L756, L757, L758, L759, L760, L761, L762, L763, L764, L765, L766, L767, L768, L769, L770, L771, L772, L773, L774, L775, L776, L777, L778, L779, L780, L781, L782, L783, L784, L785, L786, L787, L788, L789, L790, L791, L792, L793, L794, L795, L796, L797, L798, L799, L800, L801, L802, L803, L804, L805, L806, L807, L808, L809, L810, L811, L812, L813, L814, L815, L816, L817, L818, L819, L820, L821, L822, L823, L824, L825, L826, L827, L828, L829, L830, L831, L832, L833, L834, L835, L836, L837, L838, L839, L840, L841, L842, L843, L844, L845, L846, L847, L848, L849, L850, L851, L852, L853, L854, L855, L856, L857, L858, L859, L860, L861, L862, L863, L864, L865, L866, L867, L868, L869, L870, L871, L872, L873, L874, L875, L876, L877, L878, L879, L880, L881, L882, L883, L884, L885, L886, L887, L888, L889, L890, L891, L892, L893, L894, L895, L896, L897, L898, L899, L900, L901, L902, L903, L904, L905, L906, L907, L908, L909, L910, L911, L912, L913, L914, L915, L916, L917, L918, L919, L920, L921, L922, L923, L924, L925, L926, L927, L928, L929, L930, L931, L932, L933, L934, L935, L936, L937, L938, L939, L940, L941, L942, L943, L944, L945, L946, L947, L948, L949, L950, L951, L952, L953, L954, L955, L956, L957, L958, L959, L960, L961, L962, L963, L964, L965, L966, L967, L968, L969, L970, L971, L972, L973, L974, L975, L976, L977, L978, L979, L980, L981, L982, L983, L984, L985, L986, L987, L988, L989, L990, L991, L992, L993, L994, L995, L996, L997, L998, L999, L1000, L1001, L1002, L1003, L1004, L1005, L1006, L1007, L1008, L1009, L1010, L1011, L1012, L1013, L1014, L1015, L1016, L1017, L1018, L1019, L1020, L1021, L1022, L1023, L1024, L1025, L1026, L1027, L1028, L1029, L1030, L1031, L1032, L1033, L1034, L1035, L1036, L1037, L1038, L1039, L1040, L1041, L1042, L1043, L1044, L1045, L1046, L1047, L1048, L1049, L1050, L1051, L1052, L1053, L1054, L1055, L1056, L1057, L1058, L1059, L1060, L1061, L1062, L1063, L1064, L1065, L1066, L1067, L1068, L1069, L1070, L1071, L1072, L1073, L1074, L1075, L1076, L1077, L1078, L1079, L1080, L1081, L1082, L1083, L1084, L1085, L1086, L1087, L1088, L1089, L1090, L1091, L1092, L1093, L1094, L1095, L1096, L1097, L1098, L1099, L1100, L1101, L1102, L1103, L1104, L1105, L1106, L1107, L1108, L1109, L1110, L1111, L1112, L1113, L1114, L1115, L1116, L1117, L1118, L1119, L1120, L1121, L1122, L1123, L1124, L1125, L1126, L1127, L1128, L1129, L1130, L1131, L1132, L1133, L1134, L1135, L1136, L1137, L1138, L1139, L1140, L1141, L1142, L1143, L1144, L1145, L1146, L1147, L1148, L1149, L1150, L1151, L1152, L1153, L1154, L1155, L1156, L1157, L1158, L1159, L1160, L1161, L1162, L1163, L1164, L1165, L1166, L1167, L1168, L1169, L1170, L1171, L1172, L1173, L1174, L1175, L1176, L1177, L1178, L1179, L1180, L1181, L1182, L1183, L1184, L1185, L1186, L1187, L1188, L1189, L1190, L1191, L1192, L1193, L1194, L1195, L1196, L1197, L1198, L1199, L1200, L1201, L1202, L1203, L1204, L1205, L1206, L1207, L1208, L1209, L1210, L1211, L1212, L1213, L1214, L1215, L1216, L1217, L1218, L1219, L1220, L1221, L1222, L1223, L1224, L1225, L1226, L1227, L1228, L1229, L1230, L1231, L1232, L1233, L1234, L1235, L1236, L1237, L1238, L1239, L1240, L1241, L1242, L1243, L1244, L1245, L1246, L1247, L1248, L1249, L1250, L1251, L1252, L1253, L1254, L1255, L1256, L1257, L1258, L1259, L1260, L1261, L1262, L1263, L1264, L1265, L1266, L1267, L1268, L1269, L1270, L1271, L1272, L1273, L1274, L1275, L1276, L1277, L1278, L1279, L1280, L1281, L1282, L1283, L1284, L1285, L1286, L1287, L1288, L1289, L1290, L1291, L1292, L1293, L1294, L1295, L1296, L1297, L1298, L1299, L1300, L1301, L1302, L1303, L1304, L1305, L1306, L1307, L1308, L1309, L1310, L1311, L1312, L1313, L1314, L1315, L1316, L1317, L1318, L1319, L1320, L1321, L1322, L1323, L1324, L1325, L1326, L1327, L1328, L1329, L1330, L1331, L1332, L1333, L1334, L1335, L1336, L1337, L1338, L1339, L1340, L1341, L1342, L1343, L1344, L1345, L1346, L1347, L1348, L1349, L1350, L1351, L1352, L1353, L1354, L1355, L1356, L1357, L1358, L1359, L1360, L1361, L1362, L1363, L1364, L1365, L1366, L1367, L1368, L1369, L1370, L1371, L1372, L1373, L1374, L1375, L1376, L1377, L1378, L1379, L1380, L1381, L1382, L1383, L1384, L1385, L1386, L1387, L1388, L1389, L1390, L1391, L1392, L1393, L1394, L1395, L1396, L1397, L1398, L1399, L1400, L1401, L1402, L1403, L1404, L1405, L1406, L1407, L1408, L1409, L1410, L1411, L1412, L1413, L1414, L1415, L1416, L1417, L1418, L1419, L1420, L1421, L1422, L1423, L1424, L1425, L1426, L1427, L1428, L1429, L1430, L1431, L1432, L1433, L1434, L1435, L1436, L1437, L1438, L1439, L1440, L1441, L1442, L1443, L1444, L1445, L1446, L1447, L1448, L1449, L1450, L1451, L1452, L1453, L1454, L1455, L1456, L1457, L1458, L1459, L1460, L1461, L1462, L1463, L1464, L1465, L1466, L1467, L1468, L1469, L1470, L1471, L1472, L1473, L1474, L1475, L1476, L1477, L1478, L1479, L1480, L1481, L1482, L1483, L1484, L1485, L1486, L1487, L1488, L1489, L1490, L1491, L1492, L1493, L1494, L1495, L1496, L1497, L1498, L1499, L1500, L1501, L1502, L1503, L1504, L1505, L1506, L1507, L1508, L1509, L1510, L1511, L1512, L1513, L1514, L1515, L1516, L1517, L1518, L1519, L1520, L1521, L1522, L1523, L1524, L1525, L1526, L1527, L1528, L1529, L1530, L1531, L1532, L1533, L1534, L1535, L1536, L1537, L1538, L1539, L1540, L1541, L1542, L1543, L1544, L1545, L1546, L1547, L1548, L1549, L1550, L1551, L1552, L1553, L1554, L1555, L1556, L1557, L1558, L1559, L1560, L1561, L1562, L1563, L1564, L1565, L1566, L1567, L1568, L1569, L1570, L1571, L1572, L1573, L1574, L1575, L1576, L1577, L1578, L1579, L1580, L1581, L1582, L1583, L1584, L1585, L1586, L1587, L1588, L1589, L1590, L1591, L1592, L1593, L1594, L1595, L1596, L1597, L1598, L1599, L1600, L1601, L1602, L1603, L1604, L1605, L1606, L1607, L1608, L1609, L1610, L1611, L1612, L1613, L1614, L1615, L1616, L1617, L1618, L1619, L1620, L1621, L1622, L1623, L1624, L1625, L1626, L1627, L1628, L1629, L1630, L1631, L1632, L1633, L1634, L1635, L1636, L1637, L1638, L1639, L1640, L1641, L1642, L1643, L1644, L1645, L1646, L1647, L1648, L1649, L1650, L1651, L1652, L1653, L1654, L1655, L1656, L1657, L1658, L1659, L1660, L1661, L1662, L1663, L1664, L1665, L1666, L1667, L1668, L1669, L1670, L1671, L1672, L1673, L1674, L1675, L1676, L1677, L1678, L1679, L1680, L1681, L1682, L1683, L1684, L1685, L1686, L1687, L1688, L1689, L1690, L1691, L1692, L1693, L1694, L1695, L1696, L1697, L1698, L1699, L1700, L1701, L1702, L1703, L1704, L1705, L1706, L1707, L1708, L1709, L1710, L1711, L1712, L1713, L1714, L1715, L1716, L1717, L1718, L1719, L1720, L1721, L1722, L1723, L1724, L1725, L1726, L1727, L1728, L1729, L1730, L1731, L1732, L1733, L1734, L1735, L1736, L1737, L1738, L1739, L1740, L1741, L1742, L1743, L1744, L1745, L1746, L1747, L1748, L1749, L1750, L1751, L1752, L1753, L1754, L1755, L1756, L1757, L1758, L1759, L1760, L1761, L1762, L1763, L1764, L1765, L1766, L1767, L1768, L1769, L1770, L1771, L1772, L1773, L1774, L1775, L1776, L1777, L1778, L1779, L1780, L1781, L1782, L1783, L1784, L1785, L1786, L1787, L1788, L1789, L1790, L1791, L1792, L1793, L1794, L1795, L1796, L1797, L1798, L1799, L1800, L1801, L1802, L1803, L1804, L1805, L1806, L1807, L1808, L1809, L1810, L1811, L1812, L1813, L1814, L1815, L1816, L1817, L1818, L1819, L1820, L1821, L1822, L1823, L1824, L1825, L1826, L1827, L1828, L1829, L1830, L1831, L1832, L1833, L1834, L1835, L1836, L1837, L1838, L1839, L1840, L1841, L1842, L1843, L1844, L1845, L1846, L1847, L1848, L1849, L1850, L1851, L1852, L1853, L1854, L1855, L1856, L1857, L1858, L1859, L1860, L1861, L1862, L1863, L1864, L1865, L1866, L1867, L1868, L1869, L1870, L1871, L1872, L1873, L1874, L1875, L1876, L1877, L1878, L1879, L1880, L1881, L1882, L1883, L1884, L1885, L1886, L1887, L1888, L1889, L1890, L1891, L1892, L1893, L1894, L1895, L1896, L1897, L1898, L1899, L1900, L1901, L1902, L1903, L1904, L1905, L1906, L1907, L1908, L1909, L1910, L1911, L1912, L1913, L1914, L1915, L1916, L1917, L1918, L1919, L1920, L1921, L1922, L1923, L1924, L1925, L1926, L1927, L1928, L1929, L1930, L1931, L1932, L1933, L1934, L1935, L1936, L1937, L1938, L1939, L1940, L1941, L1942, L1943, L1944, L1945, L1946, L1947, L1948, L1949, L1950, L1951, L1952, L1953, L1954, L1955, L1956, L1957, L1958, L1959, L1960, L1961, L1962, L1963, L1964, L1965, L1966, L1967, L1968, L1969, L1970, L1971, L1972, L1973, L1974, L1975, L1976, L1977, L1978, L1979, L1980, L1981, L1982, L1983, L1984, L1985, L1986, L1987, L1988, L1989, L1990, L1991, L1992, L1993, L1994, L1995, L1996, L1997, L1998, L1999, L2000, L2001, L2002, L2003, L2004, L2005, L2006, L2007, L2008, L2009, L2010, L2011, L2012, L2013, L2014, L2015, L2016, L2017, L2018, L2019, L2020, L2021, L2022, L2023, L2024, L2025, L2026, L2027, L2028, L2029, L2030, L2031, L2032, L2033, L2034, L2035, L2036, L2037, L2038, L2039, L2040, L2041, L2042, L2043, L2044, L2045, L2046, L2047, L2048, L2049, L2050, L2051, L2052, L2053, L2054, L2055, L2056, L2057, L2058, L2059, L2060, L2061, L2062, L2063, L2064, L2065, L2066, L2067, L2068, L2069, L2070, L2071, L2072, L2073, L2074, L2075, L2076, L2077, L2078, L2079, L2080, L2081, L2082, L2083, L2084, L2085, L2086, L2087, L2088, L2089, L2090, L2091, L2092, L2093, L2094, L2095, L2096, L2097, L2098, L2099, L2100, L2101, L2102, L2103, L2104, L2105, L2106, L2107, L2108, L2109, L2110, L2111, L2112, L2113, L2114, L2115, L2116, L2117, L2118, L2119, L2120, L2121, L2122, L2123, L2124, L2125, L2126, L2127, L2128, L2129, L2130, L2131, L2132, L2133, L2134, L2135, L2136, L2137, L2138, L2139, L2140, L2141, L2142, L2143, L2144, L2145, L2	

WO 01/46258

PCT/US00/35095

Table 3 (cont.)

SEQ ID NO:	Invertebrate Polyptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences Domains and Motifs	Analytical Methods and Databases
25	3038193CD1	345	T204 T251 S57 S243 T263 T308 S340	N246	Transmembrane domains: L67-L95, I134-I156, I224-F242 Y44-P417 Sodium bile acid symporter family: (Score: 7.0, E-value: 9.0e-6)	HMMER
26	3460979CD1	521	S115 T184 S75 W93 S100 S126 S128 S134 S148 S183 S213 S256 S363 S389 S430 S510 T471 S180 S235 T247 S222 K306 S55 T92 T148 T298 S423 S468 S20 S52 S82 S96 T184 S208 S252 S393	N70 N169 N211	Transmembrane domains: L365-L284, I335-I361 Protein precursor PM1, transmembrane, signal PD014174: G219-E517 (p-value: 7.1e-07)	HMMER-PPAM
27	747200CD1	555	K306 S55 T92 T148 T298 S423 S468 S20 S52 S82 S96 T184 S208 S252 S393	N132 N175 N311 N361 N421	Amiloride-sensitive sodium channel alpha-subunit signature: PR01078 V102-R118, Y342-Q353, Q353-P370, Q388-R404, G455-E471 Q388-R404, G455-E471 Transmembrane domain: V452-F475 ASC, F38-L475 Amiloride-sensitive sodium channel proteins s001206: E37-I67, Y342-F368, L427-I472	BLAST-PRINTS PRODOM HMMER HMMER-PPAM BLAST-PRINTS BLOCKS

WO 01/46258

PCT/US80/35095

Table 4

Polynucleotide SEQ ID NO:	Inverte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	Position	Position
28	1416107CB1	2880	1-109,	91541704	116	609
			1901-2080,	5813453H1 (ADREINR01)	319	870
			1363-1446	5605280H1 (UPREDI707)	820	1447
				8H1845R1 (HRAINOT02)	859	1479
				1416107F6 (HRAINOT12)	1348	1896
				1416107W6 (HRAINOT12)	1579	2080
				7182660AV1	1551	1974
				7483905V1 (BRAYDIN03)	1578	2050
				8360412H1 (UPREDI707)	163	452
				7182660AV1	166	450
29	1682513CB1	2138	1-1535,	6833453R6 (ADREINR01)	1	312
			1560-1581	70207988V1	1	489
				70211506V1	394	872
				70211116V1	489	985
				70210573V2	852	1468
				70207907V1	988	1512
				70210540V2	1357	1948
				2866122T6 (RIDANNOT2)	1548	2108
				70211461V1	1597	2128
				5973322H2 (COFONP701)	1	311
30	2446638CB1	2829	1-85,	5329454H1 (INREDFR03)	250	812
			800-2202,	8259008H1 (ECLIFR03)	344	996
			999-1820	70207907V1	394	872
				70207907V1	489	985
				70207907V1	1459	2092
				17623585V6 (COLANP427)	1632	2194
				15620885V6 (SPLINOT04)	2178	2723
				2514470F6 (LIPTINOT04)	2303	2825
				5502510F6 (BRAYDIN07)	1	439
				7027134V1	183	768
31	2817822CB1	1718	1-71,	7027134V1	183	768
			509-514	7027134V1	431	930
				70271651V1	891	1453
				2817822T6 (RESINOT14)	981	1538
				70272460V1	1094	1718

WO 01/46258

PCT/US00/35095

Table 4 (cont.)

Polynucleotide SEQ ID No:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
32	400922CB1	2000	1-962	646629HH (FLACRHS01)	1	640
				678042JL (OVARLTK01)	582	1260
				630786ZH (NERDTRK03)	725	1364
				6781259H (OVARLTK01)	972	1819
				7253109J (PROSTME03)	1514	1842
33	6618083CB1	2216	1-86, 1201-2216	6759035J (BERONDR01)	1515	2000
				6722668H (BERONDR01)	604	1177
				7096752V1	586	1203
				7079481V1	1050	1650
				7078781V1	1381	1984
				7079126V1	1695	2216
				92121300.v113.gsc.2.nt.edit	1	1995
				1812692F6 (PROSTME12)	564	584
				549524F6 (PROSTME7)	1	488
				92525933	823	904
36	3222992CB1	3179	1-1846, 2106-2865, 1-1846	5070813F6 (PROSTME21)	2415	3087
				22400086 (FANCRTO1)	283	504
				6925334J (SINWNR01)	1	515
				7027068H (SINWNR02)	289	249
				6838277F6 (SINWNR01)	1752	233
				70255653D1	1392	1870
				180740385 (SINWNR13)	476	1013
				7005598D1	1317	1837
				717082442 (BRSTYMG01)	710	1372
				6555653M1 (BRSTYMG02)	1859	2372
				91444660	992	1484
				91009386	768	1254
				02719505 (SPZAFER01)	1674	1986
				91505781	552	1252
				3358383F6 (PROSTME16)	1341	1691
622195601	585	1252				
622195701	1	727				

WO 01/46258

PCT/US00/35095

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
38	4250091CB1	3294	1-920,	971557D	2823	3294
			1991-2034,	70759965V1	1854	2477
			2488-2760,	4250091F6 (BRADIC01)	1	532
			1165-1030	5715643H1 (PANCNOT16)	2539	3235
				70792723V1	444	1032
				96645666 (BRSTNOT05)	2852	3293
				70759467V1	1199	3832
				70786828V1 (BRANW02)	904	3302
				6536445H1 (BRANW02)	2740	3507
				70784825V1 (BRANW02)	1940	3507
				70784825V1	1333	3937
				275864826 (THEIAGS68)	1540	3043
39	70064603CB1	2043	1-22,	681002471 (SKRNOR01)	1361	2019
			544-1285	167618206 (BLANW05)	576	1280
				210976266 (BRANW03)	1181	1797
				70503885V1	501	1183
				7177480H2 (BRADIC01)	1	534
40	70356768CB1	1915	1241-1265,	70450108V1	509	1081
			853-897,	70451575V1	1072	1730
			1-143,	145830756 (PANCNOT2)	1	534
			1450-1532,	70431567V1	368	1078
			698-820	70445089V1	1384	1915
				70439322V1	1050	1720
				67321851 (OVAR01)	1678	307
				8292284H1 (SAR01)	509	3209
				8776218H1 (OVAR01)	184	344
				85586563 v113 (S16 nt)	1	1311
42	1256631CB1	1730	1-106,	2613664F6 (SGPNT02)	855	1177
			1711-1730,	SMBC01015V1	75	573
			567-635,	2853342F6 (KERNOT20)	1	515
			698-900	2614317G6 (GLANC01)	1126	1730
43	1670595CB1	1147		SC8A01493V1	548	724
				3323244T6 (PPIRNOT03)	860	1627
			SC8A02831V1	368	1147	
			SC1A04658V1	1	641	

WO 01/46258

PCT/US00/35095

Table 4 (cont.)

Polynucleotide SEQ ID NO.	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position			
43	185956081	2745	1-820,	515522HL (PTINOR01)	2367	2745			
			865-2059	824186A1 (PROSOT06)	1159	1748			
				1359644F6 (BRATIN08)	503	973			
				651256J1 (ADRETI03)	30	732			
				725524HL (FIBRIAC01)	814	1359			
				21272397 (ZIDRNG05)	1732	2476			
				225066HL (TORVOR02)	1493	1795			
				325066HL (TORVOR02)	1493	1795			
				832645HL (PR318R08)	2056	2740			
				867711HL (MGSHET02)	2107	1180			
				185556W6 (SSOSNG18)	2347	2745			
				432038HL (BRADDT02)	1925	2202			
			45	5550164CB1	3204	1341-2064,	432038HL (BRADDT02)	178	811
1-548,	717513HL (BRSTPM01)	530				812			
572-1097	3217236HL (WESTMT07)	2452				3175			
	2552002HL (LGMSTU06)	1781				2492			
	70090155V1	2452				3175			
	63302487J (SINNR01)	542				1211			
	6729730HL (COLLUP02)	1214				1809			
	2350920F6 (BRSTPM13)	929				1416			
	5125234HL (BRANON05)	2505				3204			
	808282HL (BRANON04)	1452				2890			
	62852376 (COMNOR04)	1452				2890			
	70077020V1 (COLELY01)	538				318			
46	1391150HL	2763				1-770,	70077020V1 (COLELY01)	1900	3538
			1345-1389,	7136223HL (COLELY01)	1	580			
			1455-1693	2468105F6 (EYENR02)	827	1663			
				70379483V1	2145	2763			
				70122316V1	1424	1971			
				70122316V1	1451	1976			
				351558A1 (EYENR01)	2056	3556			
				70480521V1	1077	1639			
				4335401F6 (KIDCTM01)	1	516			
				70466195V1	559	1157			
				70466476V1	494	1102			
			47	1702540CB1	1639	1-246,			
						1536-1639			

WO 01/46258

PCT/US00/35095

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incycle Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
48	1703342CB1	1600	1-812	3348562H1 (BRALPT024)	1	283
				285125R1 (POSTHET02)	1041	1599
				7071066H1 (BRALPT002)	250	854
				6494627H1 (SONRPT01)	1270	1600
				6679036V1 (LONRPT03)	360	1094
				927391H1 (KIDNPT05)	2085	2380
				6041361L1	691	1237
				880335V1 (COLNCR03)	1229	2383
				60211344H1	262	607
				6786334H1 (COLNCR03)	1089	1774
49	1727525CB1	3380	1-569, 1228-1654	3249033F5 (SEMVCT03)	1	626
				3556495H1 (SRGNCV02)	1934	2398
				2522115T6 (LONRPT01)	1322	1862
				6872898	848	1328
				1435229F1 (FANRPT08)	2187	2725
				3553901H1 (SYNCR01)	2570	2855
				2508452H1 (CORUTUR01)	1	114
				2771704H1 (COLNCR03)	1	2078
				6929443H1 (BRALPT01)	1815	2078
				91665164	2554	3038
50	2289333CB1	3018	1-611, 2497-3534	2493333R5 (BRALNCR01)	1234	1708
				88792695	205	302
				3579272H1 (LONRPT03)	612	3062
				3836245K1 (SRALPT05)	2624	2880
				4220788K5 (FANRPT07)	613	938
				1994713T6 (BRSPPT03)	1949	2451
				2040380K6 (BIFONR02)	463	821
				2720354F6 (LONRPT10)	450	1046
				6942433R1 (FUTETUR01)	885	1437
				61211903H1 (BRALNCR05)	1630	2340
51	2720354CB1	2608	1-2058	6556226H1 (BRALNCR02)	1713	2406
				6940912H1 (FUTETUR01)	1	462
				91327466	325	972
				8856181J1 (SRALNCR01)	1062	1666
				8197603H1 (FUTONR01)	2184	2600

WO 01/46258

PCT/US00/35095

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragments	Sequence Fragments	5' Position:	3' Position:
52	3018193CB1	3804	3392-3457, 1169-1264, 1-829, 2771-2483, 2319-1363	044564H1 (TBDYK0101) 50144686 (BRSTUT03) 52028232 1328531H1 (SINWPS01) 274322876 (PERSF0904) 4970206H1 (KIDORG01) 5685247H1 (GOLANG02) 3058193F6 (PERSM016) 3058193F6 (PERSM016) 5477440H1 (PROSTW01) 70865191V1 3154662H1 (TLYMTX02) 2257401R6 (OVARTU01) 94768882 2257401T6 (OVARTU01) 2257401T6 (OVARTU01)	2311 1733 3459 473 3235 1023 72 124 1843 2411 1 2660 1421 2856	2952 2282 3804 651 3804 1303 430 1710 3498 2832 3157 1815 3520
53	3460979CB1	1894	1-36 1746-1894	3460979F6 (PANCYU02) 3460979F6 (PANCYU02) 7161336H1 (PLACNOR01) 6860922A1 (COLENG03) 7057496H1 (BRALNG02)	519 1000 582 3 1206	945 1500 1193 557 1894
54	7472200CB1	1668	1-1668	GNN1.66554405_006	1	1668

WO 01/46258

PCT/US00/35095

Table 5

Polymerase II Seq. ID NO:	Insert Project ID	Representative Library
28	1415107CB1	UTRBDT07
29	1685113CB1	SFLN0T11
30	2445438CB1	BCLLN0R03
31	2817822CB1	BRATN076
32	4009329CB1	OVAD1R02
33	6618083CB1	HELAUT01
34	1618682CB1	HELAUT01
35	3238924CB1	PROSTN12
37	3238933CB1	PROSTN12
38	4250091CB1	PANCN0T01
39	70054403CB1	BRALT03
40	7035818CB1	BRALT03
41	1554658CB1	BRALT03
42	1554658CB1	BRALT03
43	1670555CB1	BRALT03
44	1670555CB1	BRALT03
45	5530164CB1	SGA0N0T01
46	139115CB1	BRAD1R03
47	1702940CB1	SIMN0T18
48	1703342CB1	BRAVT04
49	1727529CB1	FOST1PT02
50	2289333CB1	PROSN0T18
51	2720354CB1	LUNG0T06
52	3038183CB1	PROSTN12
53	3260979CB1	LIVPR008
54		COLPR003

Table 6

Library	Vector	Library Description
UTREB107	P1WCY	Library was constructed using RNA isolated from diseased endometrial tissue removed from a female during endometrial biopsy. Pathology indicated in phase endometrium with missing beta 3, Type II defects.
SELAN011	P1WCY	Library was constructed using RNA isolated from diseased spleen tissue removed from a 14-year-old Asian male during a total splenectomy. Pathology indicated changes consistent with idiopathic thrombocytopenic purpura. The patient presented with bruising. Patient medications included Vincristine.
MCLD7X803	P1WCY	Library was constructed from a pool of two dendritic cell libraries. Starting libraries were constructed using RNA isolated from untreated and treated dendritic cells from umbilical cord blood CD34+ precursor cells removed from a male. The cells were derived with granulocyte/macrophage colony stimulating factor (GM-CSF), tumor necrosis factor (TNF), and interleukin-3 (IL-3) in the presence of irradiated feeder cells under conditions adapted from Soares et al. (1994) Proc Natl Acad Sci USA 91:5228 and Ronaldo et al. (1996) Genome Res. 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
BRAIT0707	P1WCY	Library was constructed using RNA isolated from left frontal lobe tumor tissue removed from the brain of a 32-year-old Caucasian male during excision of a cerebral meningioma lesion. Pathology indicated low grade desmoplastic neuronal neoplasm. The patient presented with nausea, vomiting, and headache. Patient history included alcohol, tobacco use, and marijuana use twice a week for six years. Family history included
OVARD1002	P1WCY	Library was constructed using RNA isolated from a 39-year-old Caucasian female during total diseased ovarian tissue removed from a 39-year-old Caucasian female during total abdominal hysterectomy, bilateral salpingo-oophorectomy, dilation and curettage, partial colectomy, incidental appendectomy, and temporary colostomy. Pathology indicated the right and left adnexa, mesentery and muscularis propria of the sigmoid colon were extensively involved by endometriosis. Endometriosis also involved the anterior and posterior serosal surfaces of the uterus and the cul-de-sac. The endometrium was proliferative. Pathology for the associated tumor tissue indicated multiple (3) foci of moderately differentiated adenocarcinoma. The patient presented with abdominal pain and infertility. Subsequent to hysterectomy, the patient underwent laparoscopic oophorectomy and cholecystectomy and exploratory laparotomy. Patient medications included Metaxalone, Naproxen. Family history included hyperlipidemia in the mother, benign hypertension, hyperlipidemia, atherosclerotic coronary artery disease, coronary artery bypass graft, depressive disorder, brain cancer, and Type II diabetes. The library was normalized under conditions adapted from Soares et al. (1994) Proc. Natl. Acad. Sci. USA 91:5228 and Ronaldo et al. (1996) Genome Res. 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
RELAN0101	P1WCY	Library was constructed from RNA isolated from an untreated HeLa cell line, derived from cervical squamous metaplasia removed from a 31-year-old Black female.

WO 01/46258

PCT/US00/35095

Table 6 (cont.)

Library	Vector	Library Description
BRAL1003	PSPO3H1	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 17-year-old Caucasian female during excision of a cerebral meningioma. Pathology indicated a grade 4 fibrillary giant and small-cell astrocytoma. Family history included benign hypertension and cerebrovascular disease.
HNT2AGT01	PELOESCRIP1	Library was constructed at Stratstone (STR937213). Using RNA isolated from the HNT2 cell line derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor. Cells were treated with retinoic acid for 5 weeks and with mitotic inhibitors for two weeks and allowed to mature for an additional 3 weeks in conditioned medium.
OVAR2IR01	PCDNA2.1	Library was constructed using RNA isolated from right ovary tissue removed from a 45-year-old Caucasian female during bilateral oophorectomy. Pathology indicated oophorectomy, vaginal hyperthecosis, and incidental appendectomy. Pathology indicated stromal hyperthecosis of the right and left ovaries. Pathology for the matched tumor tissue indicated a dermoid cyst (benign cystic teratoma) in the left ovary. Multiple (3) intramural leiomyomata were identified. The cervix showed squamous metaplasia. Patient history included metrorrhagia, female stress incontinence, alopecia, depressive disorder, pneumonia, normal delivery, and deficiency anemia. Family history included benign hypertension, atherosclerotic deficiency anemia, and primary biliary cirrhosis.
PANCR0T01	PELOESCRIP1	Library was constructed using RNA isolated from pancreas tissue removed from an old Caucasian male who died from head trauma.
PROST0T12	P1N1CY	Library was constructed using RNA isolated from prostate tumor tissue removed from a 55-year-old Caucasian male during a radical prostatectomy. Pathology indicated adenocarcinoma (Gleason grade 2+2). Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA).
PROST0T16	P1N1CY	Library was constructed using RNA isolated from prostate tumor tissue removed from a 55-year-old Caucasian male. Pathology indicated adenocarcinoma, Gleason grade 5+4. Adenofibromatous hyperplasia was also present. The patient presented with elevated PSA. Family history included hypertension and diabetes.
THPLA2S08	ESPOW11	This subtracted THP-1 promonocyte cell line library was constructed using 5.76 x 1e6 clones from a 5-aza-2'-deoxycytidine (A2) treated THP-1 cell library. Starting RNA was made from THP-1 promonocyte cells treated for three days with 0.6 micromolar A2. The hybridization probe for subtraction was derived from a similarly constructed library. Made from RNA isolated from untreated THP-1 cells. 5.76 million clones from the A2-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. Subtracted clones were screened for expression of the human promonocyte gene at the A2C1R (393119, 1954) and Bonaids et al. (1986) locus. AAVC THP 202 is a human promonocyte line derived from peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (ref: Int. J. Cancer 26 (1980):171).

Table 6 (cont.)

Library	Vector	Library Description
BRADIN024	PINCY	Library was constructed using RNA isolated from right frontal brain tumor tissue removed from a 50-year-old Caucasian male during a cerebral meningioma lesion excision. Pathology indicated meningioma. Family history included colon cancer and cerebrovascular disease.
BRADIN03	PINCY	This normalized brain tissue library was constructed from 6.7 million independent clones from a brain tissue library. Starting RNA was made from RNA isolated from a diseased hypothalamus tissue removed from a 57-year-old Caucasian male who died from a cerebrovascular accident. Patient history included Huntington's disease and cerebral palsy. This library was constructed from 6.7 million independent clones from Soares et al. PNAS (1994) 91:9228 and Ronaldo et al. Genome Research 6 (1996):791, except that a significantly longer (48 hours/round) reannealing hybridization was used. The library was linearized and recircularized to select for insert containing clones.
LUNGFET03	PINCY	Library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation.
MEANN0701	PSPORT1	Library was constructed using RNA isolated from tumorous neuroangioma tissue removed from a 9-year-old Caucasian male during a soft tissue excision of the chest wall. Pathology indicated a ganglioneuroma.
BRANT04	PSPORT1	Library was constructed using RNA isolated from a pool of human astrocytes stimulated for 4 to 6 hours with a combination of cytokines including IL-1. The RNA was pooled for poly(A) RNA isolation and library construction.
EOSIN0203	PMUESCRIPT	Library was constructed using RNA isolated from peripheral blood cells adhered from a 48-year-old Caucasian male. Patient history included hyperosinophilia. The cell population was determined to be greater than 7% eosinophils by Wright's staining.
LIVR0803	PINCY	This normalized liver tissue library was constructed from 5.7 million independent clones from a pooled liver tissue library. Starting RNA was isolated from pooled liver tissue removed from a 4-year-old Hispanic male who died from anoxia and a 16 week female fetus who died after 16-weeks gestation from encephaly. Sarcoglycins were the genes of interest. Starting RNA was isolated from a pool of RNA isolated from the liver of a 4-year-old female patient who died from a congenitally ligated esophagus in valve prolapse in the mother of the fetus. The library was normalized in 2 rounds using conditions adapted from Soares et al. Proc. Natl. Acad. Sci. USA (1994) 91:9228 and Ronaldo et al. (1996) Genome Research 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.

Table 6 (cont.)

Library	Vector	Library Description
LUNG70706	pINCY	Library was constructed using RNA isolated from apical lung tumor tissue removed from an 80-year-old Caucasian female during a segmental lung resection. Pathology indicated a metastatic granulosa cell tumor. Patient history included pelvic soft tissue tumor and chemotherapy for one year. Family history included tuberculosis, lung cancer, and atherosclerotic coronary artery disease.
PROSB0118	pINCY	Library was constructed using RNA isolated from diseased prostate tissue removed from a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, and gastrectomy. Pathology indicated adenofibromatous hyperplasia; this tissue was associated with a grade 2 transitional cell carcinoma. Patient history included atherosclerotic coronary artery disease, and type II diabetes.
PROST0823	pINCY	This subtracted prostate tumor library was constructed using 1 million clones from a pooled prostate tumor library that was subjected to 2 rounds of subtractive hybridization with 1 million clones from a pooled prostate tissue library. The starting library for subtraction was constructed by pooling equal numbers of clones from 4 prostate tumor libraries using mRNA isolated from prostate tumor removed from Caucasian males at ages 58 (A), 61 (B), 66 (C), and 68 (D) during prostatectomy with lymph node excision. Pathology indicated adenocarcinoma. Patient history included prostate hyperplasia, renal failure, osteoarthritis, renal artery stenosis, benign HTN, thrombocytopenia, hyperlipidemia, tobacco/alcohol and hepatitis C (carrier) in donor B; elevated PSA, induration, and kidney calculus in donor C; and elevated PSA, induration, hypercholesterolemia, and kidney calculus in donor D. The hybridization probe for subtraction was constructed by pooling equal numbers of cDNA clones from 3 prostate tissue libraries derived from prostate tissue, prostate epithelial cells, and fibroblasts from prostate stroma from 3 different donors. Subtractive hybridization conditions were based on the methodologies of Swaroop et al. (1991), Muzikar, Rada, and Pflanz and Swadlow et al. (1990), and Swadlow et al. (1991).
CELE08003	pINCY pCDNA2.1	Library was constructed using RNA isolated from small intestine tissue obtained from a 59-year-old male. Library was constructed using RNA isolated from colon epithelium tissue removed from a 11-year-old Caucasian female who died from a motor vehicle accident.

WO 01/46258

PCT/US00/35095

Table 7

Program	Description	Reference	Parameter Threshold
ABIEFACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABIPARAMCEL.TDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Perceel, Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastn, blastx, blastp, blastm, 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.	ESY: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, tfastx, and tsearch.	Pearson, W.R. and 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.	ESY: fasta E value=1.0E-6 Assembled ESY: 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
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Arwood, 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. 235:1301-1334; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nussliel, Cambridge Univ. Press, pp. 1-350.	PFAM hmr: Probability value= 1.0E-5 or less Signal peptide hmr: Score= 0 or greater

WO 01/46258

PCT/US00/35095

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.	Grisikov, M. et al. (1988) CABIOS 4:61-66; Grisikov, M. et al. (1989) Methode Enzymol. 183:146-159; Poirsch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores-GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Twing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phil's Revised Assembly Program including SWAT and CrossMatch programs based on efficient implementation of the Smith-Waterman algorithm, used 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.	Score= 120 or greater; Match length= 56 or greater
Cloned SPScan	A graphical tool for viewing and editing Phrap assemblies. A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Gordon, D. et al. (1998) Genome Res. 8:195-202. Nikison, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Pearson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Pearson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMEMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sommadossi, R.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Poirsch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 5, page M51-59, Genetics Computer Group, Madison, WI.	

106

WO 01/46258

PCT/US00/35095

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
 - 5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27,
 - b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27,
 - c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and
 - 10 d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-15 27.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
- 20 5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:28-54.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
- 25 7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
- 30 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 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
 - 35 b) recovering the polypeptide so expressed.

WO 01/46258

PCT/US00/35095

10. An isolated antibody which specifically binds to a polypeptide of claim 1.
11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:
- 5 a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54,
b) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54,
c) a polynucleotide sequence complementary to a),
d) a polynucleotide sequence complementary to b), and
10 e) an RNA equivalent of a)-d).
12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.
- 15 13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization
20 complex is formed between said probe and said target polynucleotide or fragments thereof, and
b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.
- 25 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
30 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 an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

35

WO 01/46258

PCT/US00/35095

17. A composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-27.
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.
19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:
- 10 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.
- 15 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:
- 20 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
b) detecting antagonist activity in the sample.
23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.
- 25 24. A method for treating a disease or condition associated with overexpression of functional TRICH, comprising administering to a patient in need of such treatment a composition of claim 23.
- 30 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
- 35 b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a

WO 01/46258

PCT/US00/35095

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:

- 5 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
10 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
15 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
20 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;
b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at
25 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
30 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.

WO 01/46258

PCT/US00/35095

<110> INCYTE GENOMICS, INC.
 BAUGHN, Mariah R.
 BURFORD, Neil
 AU-YOUNG, Janice
 LU, Dying Aina M.
 YANG, Junming
 REDDY, Rupa
 LAI, Presti
 HILFMAN, Jennifer L.
 AZIMZAL, Yalda
 YUE, Henry
 NGUYEN, Dannie B.
 YAO, Monique C.
 GANDHI, Ameel R.
 TRNG, Y. Tom
 KHAN, Farrah A.

<120> TRANSPORTERS AND ION CHANNELS

<130> PI-0005 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/172,000; 60/176,063; 60/177,332; 60/178,572; 60/179,758;
 60/181,625

<151> 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;
 2000-02-10

<160> 54

<170> PERL Program

<210> 1

<211> 477

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> incyte ID No: L416107C01

<400> 1

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

WO 01/46258

PCT/US00/35095

```

185          190          195
Leu Leu Leu Met Cys Phe Met Pro Glu Thr Pro Arg Phe Leu Leu
200          205          210
Thr Gln His Arg Arg Gln Glu Ala Met Ala Ala Leu Arg Phe Leu
215          220          225
Trp Gly Ser Glu Gln Gly Trp Glu Asp Pro Pro Ile Gly Ala Glu
230          235          240
Gln Ser Phe His Leu Ala Leu Leu Arg Gln Pro Gly Ile Tyr Lys
245          250          255
Pro Phe Ile Ile Gly Val Ser Leu Met Ala Phe Gln Gln Leu Ser
260          265          270
Gly Val Asn Ala Val Met Phe Tyr Ala Glu Thr Ile Phe Glu Glu
275          280          285
Ala Lys Phe Lys Asp Ser Ser Ser Leu Ala Ser Val Val Val Gly Val
290          295          300
Ile Gln Val Leu Phe Thr Ala Val Ala Ala Leu Ile Met Asp Arg
305          310          315
Ala Gly Arg Arg Leu Leu Leu Val Leu Ser Gly Val Val Met Val
320          325          330
Phe Ser Thr Ser Ala Phe Gly Ala Tyr Phe Lys Leu Thr Gln Cys
335          340          345
Gly Pro Gly Asn Ser Ser His Val Ala Ile Ser Ala Pro Val Ser
350          355          360
Ala Gln Pro Val Asp Ala Ser Val Gly Leu Ala Trp Leu Ala Val
365          370          375
Gly Ser Met Cys Leu Phe Ile Ala Gly Phe Ala Val Gly Trp Gly
380          385          390
Pro Ile Pro Trp Leu Leu Met Ser Glu Ile Phe Pro Leu His Val
395          400          405
Lys Gly Val Ala Thr Gly Ile Cys Val Leu Thr Asn Trp Leu Met
410          415          420
Ala Phe Leu Val Thr Lys Glu Phe Ser Ser Leu Met Glu Val Leu
425          430          435
Arg Pro Tyr Gly Ala Phe Trp Leu Ala Ser Ala Phe Cys Ile Phe
440          445          450
Ser Val Leu Phe Thr Leu Phe Cys Val Pro Glu Thr Lys Gly Lys
455          460          465
Thr Leu Glu Gln Ile Thr Ala His Phe Glu Gly Arg
470          475

```

<210> 2
<211> 498
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> IncyLe ID No: 1682513CD1

```

<400> 2
Met Arg Arg Gln Asp Ser Arg Gly Asn Thr Val Leu His Ala Leu
1          5          10          15
Val Ala Ile Ala Asp Asn Thr Arg Glu Asn Thr Lys Phe Val Thr
20          25          30
Lys Met Tyr Asp Leu Leu Leu Lys Cys Ala Arg Leu Phe Pro
35          40          45
Asp Ser Asn Leu Glu Ala Val Leu Asn Asn Asp Gly Leu Ser Pro
50          55          60
Leu Met Met Ala Ala Lys Thr Gly Lys Ile Gly Asn Arg His Glu
65          70          75
Met Leu Ala Val Glu Pro Ile Asn Glu Leu Leu Arg Asp Lys Trp
80          85          90
Arg Lys Phe Gly Ala Val Ser Phe Tyr Ile Asn Val Val Ser Tyr
95          100          105
Leu Cys Ala Met Val Ile Phe Thr Leu Thr Ala Tyr Tyr Gln Pro
110          115          120
Leu Glu Gly Thr Pro Pro Tyr Pro Tyr Arg Thr Thr Val Asp Tyr

```

WO 01/46258

PCT/US00/35095

```

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

```

```

<210> 3
<211> 764
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2445438CD1

```

```

<400> 3
Met Thr Ser Pro Ser Ser Ser Pro Val Phe Arg Leu Glu Thr Leu
1      5      10      15
Asp Ala Gly Gln Glu Asp Gly Ser Glu Ala Asp Arg Gly Lys Leu
20      25      30
Asp Phe Gly Ser Gly Leu Pro Pro Met Glu Ser Gln Phe Gln Gly

```

WO 01/46258

PCT/US00/35095

	35		40		45										
Glu	Asp	Arg	Lys	Phe	Ala	Pro	Gln	Ile	Arg	Val	Asn	Leu	Asn	Tyr	Glu
	50														60
Arg	Lys	Gly	Thr	Gly	Ala	Ser	Gln	Pro	Asp	Pro	Asn	Arg	Phe	Asp	
	65														75
Arg	Asp	Arg	Leu	Phe	Asn	Ala	Val	Ser	Arg	Gly	Val	Pro	Glu	Asp	
	80														90
Leu	Ala	Gly	Leu	Pro	Glu	Tyr	Leu	Ser	Lys	Thr	Ser	Lys	Tyr	Leu	
	95														105
Thr	Asp	Ser	Glu	Tyr	Thr	Glu	Gly	Ser	Thr	Gly	Lys	Thr	Cys	Leu	
	110														120
Met	Lys	Ala	Val	Leu	Asn	Leu	Lys	Asp	Gly	Val	Asn	Ala	Cys	Ile	
	125														135
Leu	Pro	Leu	Leu	Gln	Ile	Asp	Arg	Asp	Ser	Gly	Asn	Pro	Gln	Pro	
	140														150
Leu	Val	Asn	Ala	Gln	Cys	Thr	Asp	Asp	Tyr	Tyr	Arg	Gly	His	Ser	
	155														165
Ala	Leu	His	Ile	Ala	Ile	Glu	Lys	Arg	Ser	Leu	Gln	Cys	Val	Lys	
	170														180
Leu	Leu	Val	Glu	Asn	Gly	Ala	Asn	Val	His	Ala	Arg	Ala	Cys	Gly	
	185														195
Arg	Phe	Phe	Gln	Lys	Gly	Gln	Gly	Thr	Cys	Phe	Tyr	Phe	Gly	Glu	
	200														210
Leu	Pro	Leu	Ser	Leu	Ala	Ala	Cys	Thr	Lys	Gln	Trp	Asp	Val	Val	
	215														225
Ser	Tyr	Leu	Leu	Glu	Asn	Pro	His	Gln	Pro	Ala	Ser	Leu	Gln	Ala	
	230														240
Thr	Asp	Ser	Gln	Gly	Asn	Thr	Val	Leu	His	Ala	Leu	Val	Met	Ile	
	245														255
Ser	Asp	Asn	Ser	Ala	Glu	Asn	Ile	Ala	Leu	Val	Thr	Ser	Met	Tyr	
	260														270
Asp	Gly	Leu	Leu	Gln	Ala	Gly	Ala	Arg	Leu	Cys	Pro	Thr	Val	Gln	
	275														285
Leu	Glu	Asp	Ile	Arg	Asn	Leu	Gln	Asp	Leu	Thr	Pro	Leu	Lys	Leu	
	290														300
Ala	Ala	Lys	Glu	Gly	Lys	Ile	Glu	Ile	Phe	Arg	His	Ile	Leu	Gln	
	305														315
Arg	Glu	Phe	Ser	Gly	Leu	Ser	His	Leu	Ser	Arg	Lys	Phe	Thr	Glu	
	320														330
Trp	Cys	Tyr	Gly	Pro	Val	Arg	Val	Ser	Leu	Tyr	Asp	Leu	Ala	Ser	
	335														345
Val	Asp	Ser	Cys	Glu	Glu	Asn	Ser	Val	Leu	Glu	Ile	Ile	Ala	Phe	
	350														360
His	Cys	Lys	Ser	Pro	His	Arg	His	Arg	Met	Val	Val	Leu	Glu	Pro	
	365														375
Leu	Asn	Lys	Leu	Leu	Gln	Ala	Lys	Trp	Asp	Leu	Leu	Ile	Pro	Lys	
	380														390
Phe	Phe	Leu	Asn	Phe	Leu	Cys	Asn	Leu	Ile	Tyr	Met	Phe	Ile	Phe	
	395														405
Thr	Ala	Val	Ala	Tyr	His	Gln	Pro	Thr	Leu	Lys	Lys	Gln	Ala	Ala	
	410														420
Pro	His	Leu	Lys	Ala	Glu	Val	Gly	Asn	Ser	Met	Leu	Leu	Thr	Gly	
	425														435
His	Ile	Leu	Ile	Leu	Leu	Gly	Gly	Ile	Tyr	Leu	Leu	Val	Gly	Gln	
	440														450
Leu	Trp	Tyr	Phe	Trp	Arg	Arg	His	Val	Phe	Ile	Trp	Ile	Ser	Phe	
	455														465
Ile	Asp	Ser	Tyr	Phe	Glu	Ile	Leu	Phe	Leu	Phe	Gln	Ala	Leu	Leu	
	470														480
Thr	Val	Val	Ser	Gln	Val	Leu	Cys	Phe	Leu	Ala	Ile	Glu	Trp	Tyr	
	485														495
Leu	Pro	Leu	Leu	Val	Ser	Ala	Leu	Val	Leu	Gly	Trp	Leu	Asn	Leu	
	500														510
Leu	Tyr	Tyr	Thr	Arg	Gly	Phe	Gln	His	Thr	Gly	Ile	Tyr	Ser	Val	
	515														525
Met	Ile	Gln	Lys	Val	Ile	Leu	Arg	Asp	Leu	Leu	Arg	Phe	Leu	Leu	
	530														540

WO 01/46258

PCT/US00/35095

Ile Tyr Leu Val Phe Leu Phe Gly Phe Ala Val Ala Leu Val Ser
 545 550 555
 Leu Ser Gln Glu Ala Trp Arg Pro Glu Ala Pro Thr Gly Pro Asn
 560 565 570
 Ala Thr Glu Ser Val Gln Pro Met Glu Gly Gln Glu Asp Glu Gly
 575 580
 Asn Gly Ala Gln Tyr Arg Gly Ile Leu Glu Ala Ser Leu Glu Leu
 590 595
 Phe Lys Phe Thr Ile Gly Met Gly Glu Leu Ala Phe Gln Glu Gln
 600 605
 Leu His Phe Arg Gly Met Val Leu Leu Leu Leu Ala Tyr Val
 610 615
 Leu Leu Thr Tyr Ile Leu Leu Leu Asn Met Leu Ile Ala Leu Met
 620 625
 Ser Glu Thr Val Asn Ser Val Ala Thr Asp Ser Trp Ser Ile Trp
 630 635
 Lys Leu Gln Lys Ala Ile Ser Val Leu Glu Met Glu Asn Gly Tyr
 640 645
 Trp Trp Cys Arg Lys Lys Gln Arg Ala Gly Val Met Leu Thr Val
 650 655
 Gly Thr Lys Pro Asp Gly Ser Pro Asp Glu Arg Trp Cys Phe Arg
 660 665
 Val Glu Glu Val Asn Trp Ala Ser Trp Glu Gln Thr Leu Pro Thr
 670 675
 Leu Cys Glu Asp Pro Ser Gly Ala Gly Val Pro Arg Thr Leu Glu
 680 685
 Asn Pro Val Leu Ala Ser Pro Pro Lys Glu Asp Glu Asp Gly Ala
 690 695
 Ser Glu Glu Asn Tyr Val Pro Val Gln Leu Leu Gln Ser Asn
 700 705
 710 715
 720 725
 730 735
 740 745
 750 755

<210> 4
 <211> 255
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2817822CD1

<400> 4
 Met Trp Gln Gly Cys Ala Val Glu Arg Pro Val Gly Arg Met Thr
 1 5 10 15
 Ser Gln Thr Pro Leu Pro Gln Ser Pro Arg Pro Arg Arg Pro Thr
 20 25 30
 Met Ser Thr Val Val Glu Leu Asn Val Gly Gly Glu Phe His Thr
 35 40 45
 Thr Thr Leu Gly Thr Leu Arg Lys Phe Pro Gly Ser Lys Leu Ala
 50 55 60
 Glu Met Phe Ser Ser Leu Ala Lys Ala Ser Thr Asp Ala Glu Gly
 65 70 75
 Arg Phe Phe Ile Asp Arg Pro Ser Thr Tyr Phe Arg Pro Ile Leu
 80 85 90
 Asp Tyr Leu Arg Thr Gly Gln Val Pro Thr Gln His Ile Pro Glu
 95 100 105
 Val Tyr Arg Glu Ala Gln Phe Tyr Glu Ile Lys Pro Leu Val Lys
 110 115 120
 Leu Leu Glu Asp Met Pro Gln Ile Phe Gly Glu Gln Val Ser Arg
 125 130 135
 Lys Gln Phe Leu Leu Gln Val Pro Gly Tyr Ser Glu Asn Leu Glu
 140 145 150
 Leu Met Val Arg Leu Ala Arg Ala Glu Ala Ile Thr Ala Arg Lys
 155 160 165
 Ser Ser Val Leu Val Cys Leu Val Glu Thr Glu Glu Gln Asp Ala
 170 175 180
 Tyr Tyr Ser Glu Val Leu Cys Phe Leu Gln Asp Lys Lys Met Phe
 185 190 195

WO 01/46258

PCT/US00/35095

Lys Ser Val Val Lys Phe Gly Pro Trp Lys Ala Val Leu Asp Asn 200 205 210
 Ser Asp Leu Met His Cys Leu Glu Met Asp Ile Lys Ala Gln Gly 215 220 225
 Tyr Lys Val Phe Ser Lys Phe Tyr Leu Thr Tyr Pro Thr Lys Arg 230 235 240
 Asn Glu Phe His Phe Asn Ile Tyr Ser Phe Thr Phe Thr Trp Trp 245 250 255

<210> 5
 <211> 584
 <212> FRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 40C9329CD1

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

WO 01/46258

PCT/US00/35095

```

350          355          360
Leu His Leu Val Ile Ser Pro Leu Val Val Val Leu Thr Leu Gln
365
Ser Gly Thr Tyr Gly Val Tyr Glu Ile Gly Gly Leu Val Pro Val
380
Trp Val Val Val Val Ile Ala Gly Thr Ala Leu Ala Ser Val Thr
395
Phe Phe Ala Thr Ser Asp Ser Gln Pro Pro Arg Leu His Trp Leu
410
Phe Ala Phe Leu Gly Phe Leu Thr Ser Ala Leu Trp Ile Asn Ala
425
Ala Ala Thr Glu Val Val Asn Ile Leu Arg Ser Leu Gly Val Val
440
Phe Arg Leu Ser Asn Thr Val Leu Gly Leu Thr Leu Leu Ala Trp
455
Gly Asn Ser Ile Gly Asp Ala Phe Ser Asp Phe Thr Leu Ala Arg
470
Gln Gly Tyr Pro Arg Met Ala Phe Ser Ala Cys Phe Gly Gly Ile
485
Ile Phe Asn Ile Leu Val Gly Val Gly Leu Gly Cys Leu Leu Gln
500
Ile Ser Arg Ser His Thr Glu Val Lys Leu Glu Pro Asp Gly Leu
515
Leu Val Trp Val Leu Ala Gly Ala Leu Gly Leu Ser Leu Val Phe
530
Ser Leu Val Ser Val Pro Leu Gln Cys Phe Gln Leu Ser Arg Val
545
Tyr Gly Phe Cys Leu Leu Leu Phe Tyr Leu Asn Phe Leu Val Val
560
Ala Leu Leu Ile Glu Phe Gly Val Ile His Leu Lys Ser Met
575
580

```

```

<210> 6
<211> 416
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 6618063CD1

```

```

<400> 6
Met Lys Leu Ser Lys Lys Asp Arg Gly Glu Asp Glu Glu Ser Asp
1 10 15
Ser Ala Lys Lys Lys Leu Asp Trp Ser Cys Ser Leu Leu Val Ala
20 25 30
Ser Leu Ala Gly Ala Phe Gly Ser Ser Phe Leu Tyr Gly Tyr Asn
35 40 45
Leu Ser Val Val Asn Ala Pro Thr Pro Tyr Ile Lys Ala Phe Tyr
50 55 60
Asn Glu Ser Trp Glu Arg Arg His Gly Arg Pro Ile Asp Pro Asp
65 70 75
Thr Leu Thr Leu Leu Trp Ser Val Thr Val Ser Ile Phe Ala Ile
80 85 90
Gly Gly Leu Val Gly Thr Leu Ile Val Lys Met Ile Gly Lys Val
95 100 105
Leu Gly Arg Lys His Thr Leu Leu Ala Asn Asn Gly Phe Ala Ile
110 115 120
Ser Ala Ala Leu Leu Met Ala Cys Ser Leu Gln Ala Gly Ala Phe
125 130 135
Glu Met Leu Ile Val Gly Arg Phe Ile Met Gly Ile Asp Gly Gly
140 145 150
Val Ala Leu Ser Val Leu Pro Met Tyr Leu Ser Glu Ile Ser Pro
155 160 165
Lys Glu Ile Arg Gly Ser Leu Gly Gln Val Thr Ala Ile Phe Ile
170 175 180
Cys Ile Gly Val Phe Thr Gly Gln Leu Leu Gly Leu Pro Glu Leu

```

WO 01/46258

PCT/US00/35095

```

185
Leu Gly Lys Glu Ser Thr Trp Pro Tyr Leu Phe Gly Val Ile Val 195
200
Val Pro Ala Val Val Gln Leu Leu Ser Leu Pro Phe Leu Pro Asp 210
215
Ser Pro Arg Tyr Leu Leu Leu Glu Lys His Asn Glu Ala Arg Ala 225
230
Val Lys Ala Phe Gln Thr Phe Leu Gly Lys Ala Asp Val Ser Gln 240
245
Glu Val Glu Glu Val Leu Ala Glu Ser His Val Gln Arg Ser Ile 255
260
Arg Leu Val Ser Val Leu Glu Leu Leu Arg Ala Pro Tyr Val Arg 270
275
Trp Gln Val Val Thr Val Ile Val Thr Met Ala Cys Tyr Gln Leu 285
290
Cys Gly Leu Asn Ala Ile Trp Phe Tyr Thr Asn Ser Ile Phe Gly 300
305
Lys Ala Gly Ile Pro Leu Ala Lys Ile Pro Tyr Val Thr Leu Ser 315
320
Thr Gly Gly Ile Glu Thr Leu Ala Ala Val Phe Ser Gly Leu Val 330
335
Ile Glu His Leu Gly Arg Arg Pro Leu Leu Ile Gly Gly Phe Gly 345
350
Leu Met Gly Leu Phe Phe Gly Thr Leu Thr Ile Thr Leu Thr Leu 360
365
Gln Asp His Ala Pro Trp Val Pro Tyr Leu Ser Ile Val Gly Ile 375
380
Leu Ala Ile Ile Ala Ser Phe Cys Ser Gly Pro Ala Val Phe Pro 390
395
Glu Glu Thr Val Asn Val Ser Ile Val Ser Glu 405
410

```

```

<210> 7
<211> 664
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7472002C01

```

```

<400> 7
Met Thr Glu Lys Thr Asn Gly Val Lys Ser Ser Pro Ala Asn Asn 15
1
His Asn His His Ala Pro Pro Ala Ile Lys Ala Asn Gly Lys Asp 20
25
Esp His Arg Thr Ser Ser Arg Pro His Ser Ala Ala Asp Asp Asp 30
35
Thr Ser Ser Glu Leu Gln Arg Leu Ala Asp Val Asp Ala Pro Gln 40
45
Gln Gly Arg Ser Gly Phe Arg Arg Ile Val Arg Leu Val Gly Ile 50
55
Ile Arg Glu Trp Ala Asn Lys Asn Phe Arg Glu Glu Glu Pro Arg 60
65
Pro Asp Ser Phe Leu Glu Arg Phe Arg Gly Pro Glu Leu Gln Thr 70
75
Val Thr Thr Gln Glu Gly Asp Gly Lys Gly Asp Lys Asp Gly Glu 80
85
Asp Lys Gly Thr Lys Lys Phe Glu Leu Phe Val Leu Asp Pro 90
95
Ala Gly Asp Trp Tyr Tyr Cys Trp Leu Phe Val Ile Ala Met Pro 100
105
Val Leu Tyr Asn Trp Cys Leu Leu Val Ala Arg Ala Cys Phe Ser 110
115
Asp Leu Gln Lys Gly Tyr Tyr Leu Val Trp Leu Val Leu Asp Tyr 120
125
Val Ser Asp Val Val Tyr Ile Ala Asp Leu Phe Ile Arg Leu Arg 130
135
140
145
150
155
160
165
170
175
180

```

WO 01/46258

PCT/US00/35095

Thr	Gly	Phe	Leu	185	190	195								
			Glu	Gln	Gly	Leu	Val	Lys	Asp	Thr	Lys	Lys		
			200									210		
Leu	Arg	Asp	Asn	Tyr	Ile	His	Thr	Leu	Gln	Phe	Lys	Leu	Asp	Val
			215											225
Ala	Ser	Ile	Ile	Pro	Thr	Asp	Leu	Ile	Tyr	Phe	Ala	Val	Asp	Ile
			230											240
His	Ser	Pro	Glu	Val	Arg	Phe	Asn	Arg	Leu	Leu	His	Phe	Ala	Arg
			245											255
Met	Phe	Clu	Phe	Phe	Asp	Arg	Thr	Clu	Thr	Arg	Thr	Asn	Tyr	Pro
			260											270
Asn	Ile	Phe	Arg	Ile	Ser	Asn	Leu	Val	Leu	Tyr	Ile	Leu	Val	Ile
			275											285
Ile	His	Trp	Asn	Ala	Cys	Ile	Tyr	Tyr	Ala	Ile	Ser	Lys	Ser	Ile
			290											300
Gly	Phe	Gly	Val	Asp	Thr	Trp	Val	Tyr	Pro	Asn	Ile	Thr	Asp	Pro
			305											315
Glu	Tyr	Gly	Tyr	Leu	Ala	Arg	Glu	Tyr	Ile	Tyr	Cys	Leu	Tyr	Trp
			320											330
Ser	Thr	Leu	Thr	Leu	Thr	Thr	Ile	Gly	Glu	Thr	Pro	Pro	Pro	Val
			335											345
Lys	Asp	Glu	Glu	Tyr	Leu	Phe	Val	Ile	Phe	Asp	Phe	Leu	Ile	Gly
			350											360
Val	Leu	Ile	Phe	Ala	Thr	Ile	Val	Gly	Asn	Val	Gly	Ser	Met	Ile
			365											375
Ser	Asn	Met	Asn	Ala	Thr	Arg	Ala	Glu	Pro	Gln	Ala	Lys	Ile	Asp
			380											390
Ala	Val	Lys	His	Tyr	Met	Gln	Phe	Arg	Lys	Val	Ser	Lys	Gly	Met
			395											405
Glu	Ala	Lys	Val	Ile	Arg	Trp	Phe	Asp	Tyr	Leu	Trp	Thr	Asn	Lys
			410											420
Lys	Thr	Val	Asp	Glu	Arg	Glu	Ile	Leu	Lys	Asn	Leu	Pro	Ala	Lys
			425											435
Leu	Arg	Ala	Glu	Ile	Ala	Ile	Asn	Val	His	Leu	Ser	Thr	Leu	Lys
			440											450
Lys	Val	Arg	Ile	Phe	His	Asp	Cys	Glu	Ala	Gly	Leu	Leu	Val	Glu
			455											465
Leu	Val	Leu	Lys	Leu	Arg	Pro	Gln	Val	Phe	Ser	Pro	Gly	Asp	Tyr
			470											480
Ile	Cys	Arg	Lys	Gly	Asp	Ile	Gly	Lys	Glu	Met	Tyr	Ile	Ile	Lys
			485											495
Glu	Gly	Lys	Leu	Ala	Val	Val	Ala	Asp	Asp	Gly	Val	Thr	Gln	Tyr
			500											510
Ala	Leu	Leu	Ser	Ala	Gly	Ser	Cys	Phe	Gly	Glu	Ile	Ser	Ile	Leu
			515											525
Asn	Ile	Lys	Gly	Ser	Lys	Met	Gly	Asn	Arg	Arg	Thr	Ala	Asn	Ile
			520											530
Arg	Ser	Leu	Gly	Tyr	Ser	Asp	Leu	Phe	Cys	Leu	Ser	Lys	Asp	Asp
			545											555
Leu	Met	Glu	Ala	Val	Thr	Glu	Tyr	Pro	Asp	Ala	Lys	Lys	Val	Leu
			560											570
Glu	Glu	Arg	Gly	Arg	Glu	Ile	Leu	Met	Lys	Glu	Gly	Leu	Leu	Asp
			575											585
Glu	Asn	Glu	Val	Ala	Thr	Ser	Met	Glu	Val	Asp	Val	Gln	Glu	Lys
			590											600
Leu	Gly	Gln	Leu	Glu	Thr	Asn	Met	Glu	Thr	Leu	Tyr	Thr	Arg	Phe
			605											615
Gly	Arg	Leu	Leu	Ala	Clu	Tyr	Thr	Gly	Ala	Gln	Gln	Lys	Leu	Lys
			620											630
Gln	Arg	Ile	Thr	Val	Leu	Glu	Thr	Lys	Met	Lys	Gln	Asn	Asn	Glu
			625											645
Asp	Asp	Tyr	Leu	Ser	Asp	Gly	Met	Asn	Ser	Pro	Glu	Leu	Ala	Ala
			635											650
Ala	Asp	Glu	Pro											655

<210> 8

WO 01/46258

PCT/US00/35095

<211> 242
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1812692CD1

```

<400> 8
Met Ser Phe Arg Ala Ala Arg Leu Ser Met Arg Asn Arg Arg Asn
1 5 10 15
Asp Thr Leu Asp Ser Thr Arg Thr Leu Tyr Ser Ser Ala Ser Arg
20 25 30
Ser Thr Asp Leu Ser Tyr Ser Glu Ser Asp Leu Val Asn Phe Ile
35 40 45
Gln Ala Asn Phe Lys Lys Arg Glu Cys Val Phe Phe Thr Lys Asp
50 55 60
Ser Lys Ala Thr Glu Asn Val Cys Lys Cys Gly Tyr Ala Gln Ser
65 70 75
Gln His Met Glu Gly Thr Gln Ile Asn Gln Ser Glu Lys Trp Asn
80 85 90
Tyr Lys Lys His Thr Lys Glu Phe Pro Thr Asp Ala Phe Gly Asp
95 100 105
Ile Gln Phe Glu Thr Leu Gly Lys Lys Gly Lys Tyr Ile Arg Leu
110 115 120
Ser Cys Asp Thr Asp Ala Glu Ile Leu Tyr Glu Leu Leu Thr Gln
125 130 135
His Trp His Leu Lys Thr Pro Asn Leu Val Ile Ser Val Thr Gly
140 145 150
Gly Ala Lys Asn Phe Ala Leu Lys Pro Arg Met Arg Lys Ile Phe
155 160 165
Ser Arg Leu Ile Tyr Ile Ala Gln Ser Lys Gly Ala Trp Ile Leu
170 175 180
Thr Gly Gly Thr His Tyr Gly Leu Met Lys Tyr Ile Gly Glu Val
185 190 195
Val Arg Asp Asn Thr Ile Ser Arg Ser Ser Glu Glu Asn Ile Val
200 205 210
Ala Ile Gly Ile Ala Ala Trp Gly Met Val Ser Asn Arg Asp Thr
215 220 225
Leu Ile Arg Asn Cys Asp Ala Glu Val Pro Val Gly Gln Glu Glu
230 235 240
Val Cys

```

<210> 9
<211> 398
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3222992CD1

```

<400> 9
Met Val Ala Ala Pro Ile Phe Gly Tyr Leu Gly Asp Arg Phe Asn
1 5 10 15
Arg Lys Val Ile Leu Ser Cys Gly Ile Phe Phe Trp Ser Ala Val
20 25 30
Thr Phe Ser Ser Ser Phe Ile Pro Gln Gln Tyr Phe Trp Leu Leu
35 40 45
Val Leu Ser Arg Gly Leu Val Gly Ile Gly Glu Ala Ser Tyr Ser
50 55 60
Thr Ile Ala Pro Thr Ile Ile Gly Asp Leu Phe Thr Lys Asn Thr
65 70 75
Arg Thr Leu Met Leu Ser Val Phe Tyr Phe Ala Ile Pro Leu Gly
80 85 90
Ser Gly Leu Gly Tyr Ile Thr Gly Ser Ser Val Lys Gln Ala Ala

```

WO 01/46258

PCT/US00/35095

```

          95          100          105
Gly Asp Trp His Trp Ala Leu Arg Val Ser Pro Val Leu Gly Met
110          115          120
Ile Thr Gly Thr Leu Ile Leu Ile Leu Val Pro Ala Thr Lys Arg
125          130          135
Gly His Ala Asp Gln Leu Gly Asp Gln Leu Lys Ala Arg Thr Ser
140          145          150
Trp Leu Arg Asp Met Lys Ala Leu Ile Arg Asn Arg Ser Tyr Val
155          160          165
Phe Ser Ser Leu Ala Thr Ser Ala Val Ser Phe Ala Thr Gly Ala
170          175          180
Leu Gly Met Trp Ile Pro Leu Tyr Leu His Arg Ala Gln Val Val
185          190          195
Gln Lys Thr Ala Glu Thr Cys Asn Ser Pro Pro Cys Gly Ala Lys
200          205          210
Asp Ser Leu Ile Phe Gly Ala Ile Thr Cys Phe Thr Gly Phe Leu
215          220          225
Gly Val Val Thr Gly Ala Gly Ala Thr Arg Trp Cys Arg Leu Lys
230          235          240
Thr Gln Arg Ala Asp Pro Leu Val Cys Ala Val Gly Met Leu Gly
245          250          255
Ser Ala Ile Phe Ile Cys Leu Ile Phe Val Ala Ala Lys Ser Ser
260          265          270
Ile Val Gly Ala Tyr Ile Cys Ile Phe Val Gly Glu Thr Leu Leu
275          280          285
Phe Ser Asn Trp Ala Ile Thr Ala Asp Ile Leu Met Tyr Val Val
290          295          300
Ile Pro Thr Arg Arg Ala Thr Ala Val Ala Leu Gln Ser Phe Thr
305          310          315
Ser His Leu Leu Gly Asp Ala Gly Ser Pro Tyr Leu Ile Gly Phe
320          325          330
Ile Ser Asp Leu Ile Arg Gln Ser Thr Lys Asp Ser Pro Leu Trp
335          340          345
Glu Phe Leu Ser Leu Gly Tyr Ala Leu Met Leu Cys Pro Phe Val
350          355          360
Val Val Leu Gly Gly Met Phe Phe Leu Ala Thr Ala Leu Phe Phe
365          370          375
Val Ser Asp Arg Ala Arg Ala Glu Gln Gln Val Asn Gln Leu Ala
380          385          390
Met Pro Pro Ala Ser Val Lys Val
395

```

<210> 10

<211> 553

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3358383CD1

<400> 10

```

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

```

WO 01/46258

PCT/US00/35095

```

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

```

```

<210> 11
<211> 213
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature

```

WO 01/46258

PCT/US00/35095

<223> Incyte ID No: 4250091CD1

<400> 11

```

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

```

<210> 12

<211> 476

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 70C64803CD1

<400> 12

```

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

```

WO 01/46258

PCT/US00/35095

170 175 180
Ala Asp Gln Arg Ser Arg Met Leu Ser Ile Phe Tyr Phe Ala Ile
185 190 195
Pro Val Gly Ser Gly Leu Gly Tyr Ile Ala Gly Ser Lys Val Lys
200 205 210
Asp Met Ala Gly Asp Trp His Trp Ala Leu Arg Val Thr Pro Gly
215 220 225
Leu Gly Val Val Ala Val Leu Leu Leu Phe Leu Val Val Arg Glu
230 235 240
Pro Pro Arg Gly Ala Val Glu Arg His Ser Asp Leu Pro Pro Leu
245 250 255
Asn Pro Thr Ser Trp Trp Ala Asp Leu Arg Ala Leu Ala Arg Asn
260 265 270
Leu Ile Phe Gly Leu Ile Thr Cys Leu Thr Gly Val Leu Gly Val
275 280 285
Gly Leu Gly Val Glu Ile Ser Arg Arg Leu Arg His Ser Asn Pro
290 295 300
Arg Ala Asp Pro Leu Val Cys Ala Thr Gly Leu Leu Gly Ser Ala
305 310 315
Pro Phe Leu Phe Leu Ser Leu Ala Cys Ala Arg Gly Ser Ile Val
320 325 330
Ala Thr Tyr Ile Phe Ile Phe Ile Gly Gln Thr Leu Leu Ser Met
335 340 345
Asn Thr Ala Ile Val Ala Asp Ile Leu Leu Tyr Val Val Ile Pro
350 355 360
Thr Arg Arg Ser Thr Ala Glu Ala Phe Gln Ile Val Leu Ser His
365 370 375
Leu Leu Gly Asp Ala Gly Ser Pro Tyr Leu Ile Gly Leu Ile Ser
380 385 390
Asp Arg Leu Arg Arg Asn Trp Pro Pro Ser Phe Leu Ser Glu Phe
395 400 405
Arg Ala Leu Gln Phe Ser Leu Met Leu Cys Ala Phe Val Gly Ala
410 415 420
Leu Gly Gly Ala Ala Phe Leu Gly Thr Ala Ile Phe Ile Glu Ala
425 430 435
Asp Arg Arg Arg Ala Gln Leu His Val Gln Gly Leu Leu His Glu
440 445 450
Ala Gly Ser Thr Asp Asp Arg Ile Val Val Pro Gln Arg Gly Arg
455 460 465
Ser Thr Arg Val Pro Val Ala Ser Val Leu Ile
470 475

<210> 13
<211> 266
<212> FRT
<213> Homo sapiens

<220>
<221> misc_faature
<223> Incyte ID No: 70356768CD1

<400> 13
Met Leu His Ala Leu Leu Arg Ser Arg Met Ile Gln Gly Arg Ile
1 10 15
Leu Leu Leu Thr Ile Cys Ala Ala Gly Ile Gly Gly Thr Phe Gln
20 25 30
Phe Gly Tyr Asn Leu Ser Ile Ile Asn Ala Pro Thr Leu His Ile
35 40 45
Gln Glu Phe Thr Asn Glu Thr Trp Gln Ala Arg Thr Gly Glu Pro
50 55 60
Leu Pro Asp His Leu Val Leu Leu Met Trp Ser Leu Ile Val Ser
65 70 75
Leu Tyr Pro Leu Gly Gly Leu Phe Gly Ala Leu Leu Ala Gly Pro
80 85 90
Leu Ala Ile Thr Leu Gly Arg Lys Lys Ser Leu Leu Val Asn Asn
95 100 105
Ile Phe Val Val Ser Ala Ala Ile Leu Phe Gly Phe Ser Arg Lys

WO 01/46258

PCT/US00/35095

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

<210> 14
<211> 436
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5674114CD1

<400> 14
Met Gly Leu Ala Arg Ala Leu Arg Arg Leu Ser Gly Ala Leu Asp
1 10 15
Ser Gly Asp Ser Arg Ala Gly Asp Glu Glu Ala Gly Pro Gly
20 25 30
Leu Cys Arg Asn Gly Trp Ala Pro Ala Pro Val Gln Ser Pro Val
35 40 45
Gly Arg Arg Arg Gly Arg Phe Val Lys Lys Asp Gly His Cys Asn
50 55 60
Val Arg Phe Val Asn Leu Gly Gly Gln Gly Ala Arg Tyr Leu Ser
65 70 75
Asp Leu Phe Thr Thr Cys Val Asp Val Arg Trp Arg Trp Met Cys
80 85 90
Leu Leu Phe Ser Cys Ser Phe Leu Ala Ser Trp Leu Leu Phe Gly
95 100 105
Leu Ala Phe Trp Leu Ile Ala Ser Leu His Gly Asp Leu Ala Ala
110 115 120
Pro Pro Pro Pro Ala Pro Cys Phe Ser His Val Ala Ser Phe Leu
125 130 135
Ala Ala Phe Leu Phe Ala Leu Glu Thr Gln Thr Ser Ile Gly Tyr
140 145 150
Gly Val Arg Ser Val Thr Glu Glu Cys Pro Ala Ala Val Ala Ala
155 160 165
Val Val Leu Gln Cys Ile Ala Gly Cys Val Leu Asp Ala Phe Val
170 175 180
Val Gly Ala Val Met Ala Lys Met Ala Lys Pro Lys Lys Arg Asn
185 190 195
Glu Thr Leu Val Phe Ser Glu Asn Ala Val Val Ala Leu Arg Asp
200 205 210
His Arg Leu Cys Leu Met Trp Arg Val Gly Asn Leu Arg Arg Ser
215 220 225
His Leu Val Glu Ala His Val Arg Ala Gln Leu Leu Gln Pro Arg
230 235 240
Val Thr Pro Glu Gly Glu Tyr Ile Pro Leu Asp His Gln Asp Val
245 250 255
Asp Val Gly Phe Asp Gly Gly Thr Asp Arg Ile Phe Leu Val Ser
260 265 270
Pro Ile Thr Ile Val His Glu Ile Asp Ser Ala Ser Pro Leu Tyr

WO 01/46258

PCT/US00/35095

275 280 285
 Glu Leu Gly Arg Ala Glu Leu Ala Arg Ala Asp Phe Glu Leu Val
 290 295 300
 Val Ile Leu Glu Gly Met Val Glu Ala Thr Ala Met Thr Thr Gln
 305 310 315
 Cys Arg Ser Ser Tyr Leu Pro Gly Glu Leu Leu Trp Gly His Arg
 320 325 330
 Phe Glu Pro Val Leu Phe Gln Arg Gly Ser Gln Tyr Glu Val Asp
 335 340 345
 Tyr Arg His Phe His Arg Thr Tyr Glu Val Pro Gly Thr Pro Val
 350 355 360
 Cys Ser Ala Lys Glu Leu Asp Glu Arg Ala Glu Gln Ala Ser His
 365 370 375
 Ser Leu Lys Ser Ser Phe Pro Gly Ser Leu Thr Ala Phe Cys Tyr
 380 385 390
 Glu Asn Glu Leu Ala Leu Ser Cys Cys Gln Glu Glu Asp Glu Asp
 395 400 405
 Asp Glu Thr Glu Glu Gly Asn Gly Val Glu Thr Glu Asp Gly Ala
 410 415 420
 Ala Ser Pro Arg Val Leu Thr Pro Thr Leu Ala Leu Thr Leu Pro
 425 430 435
 Pro

<210> 15

<211> 433

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1254635CD1

<400> 15

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

WO 01/46258

PCT/US00/35095

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

<210> 16
 <211> 299
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1670595CD1

<400> 16
 Met Ala Ser Glu Ser Ser Pro Leu Leu Ala Tyr Arg Leu Leu Gly
 1 5 10 15
 Glu Glu Gly Val Ala Leu Pro Ala Asn Gly Ala Gly Gly Pro Gly
 20 25 30
 Gly Ala Ser Ala Arg Lys Leu Ser Thr Phe Leu Gly Val Val Val
 35 40 45
 Pro Thr Val Leu Ser Met Phe Ser Ile Val Val Phe Leu Arg Ile
 50 55 60
 Gly Phe Val Val Gly His Ala Gly Leu Leu Gln Ala Leu Ala Met
 65 70 75
 Leu Leu Val Ala Tyr Phe Ile Leu Ala Leu Thr Val Leu Ser Val
 80 85 90
 Cys Ala Ile Ala Thr Asn Gly Ala Val Gln Gly Gly Gly Ala Tyr
 95 100 105
 Cys Ile Leu Gln His Arg Trp Thr Gly Met Pro Gln Gly Pro Val
 110 115 120
 Gly Ser Gly Ser Cys Pro Arg Ala Thr Ala Trp Asn Leu Leu Tyr
 125 130 135
 Gly Ser Leu Leu Leu Gly Leu Val Gly Gly Val Cys Thr Leu Gly
 140 145 150
 Ala Gly Leu Tyr Ala Arg Ala Ser Phe Leu Thr Phe Leu Leu Val
 155 160 165
 Ser Gly Ser Leu Ala Ser Val Leu Ile Ser Phe Val Ala Val Gly
 170 175 180
 Pro Arg Asp Ile Arg Leu Thr Pro Arg Pro Gly Pro Asn Gly Ser
 185 190 195
 Ser Leu Pro Pro Arg Phe Gly His Phe Thr Gly Phe Asn Ser Ser

WO 01/46258

PCT/US00/35095

Thr	Leu	Lys	Asp	Asn	Leu	Gly	Ala	Gly	Tyr	Ala	Glu	Asp	Tyr	Thr	200
															205
Thr	Gly	Ala	Val	Met	Asn	Phe	Ala	Ser	Val	Phe	Ala	Val	Leu	Phe	210
															215
Asn	Gly	Arg	His	His	Gly	Trp	Gly	Gln	His	Val	Arg	Gly	Ala	Gln	220
															225
Gly	Pro	Gln	Pro	Gly	Asp	Pro	Ser	Gly	His	Asp	Arg	Arg	Arg	Arg	230
															235
Leu	His	Leu	Leu	Arg	Leu	Cys	Pro	Ala	Phe	Leu	Ser	Leu	Gln	Pro	240
															245
Pro	Phe	Thr	Gly	Ala	Leu	Met	Leu	Gly	Ala	Arg	Pro	Pro	Leu		250
															255

<210> 17
 <211> 506
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 18595600p1

Met	Pro	Ser	Ser	Val	Thr	Ala	Leu	Gly	Gln	Ala	Arg	Ser	Ser	Gly	<400> 17
															15
Pro	Gly	Met	Ala	Pro	Ser	Ala	Cys	Cys	Cys	Ser	Pro	Ala	Ala	Leu	20
															25
Gln	Arg	Arg	Leu	Pro	Ile	Leu	Ala	Trp	Leu	Pro	Ser	Tyr	Ser	Leu	30
															35
Gln	Trp	Leu	Lys	Met	Asp	Phe	Val	Ala	Gly	Leu	Ser	Val	Gly	Leu	40
															45
Thr	Ala	Ile	Pro	Gln	Ala	Leu	Ala	Tyr	Ala	Glu	Val	Ala	Gly	Leu	50
															55
Pro	Pro	Gln	Tyr	Gly	Leu	Tyr	Ser	Ala	Phe	Met	Gly	Cys	Phe	Val	60
															65
Tyr	Phe	Phe	Leu	Gly	Thr	Ser	Arg	Asp	Val	Thr	Leu	Gly	Pro	Thr	70
															75
Ala	Ile	Met	Ser	Leu	Leu	Val	Ser	Phe	Tyr	Thr	Phe	His	Glu	Pro	80
															85
Ala	Tyr	Ala	Val	Leu	Leu	Ala	Phe	Leu	Ser	Gly	Cys	Ile	Gln	Leu	90
															95
Ala	Met	Gly	Val	Leu	Arg	Leu	Gly	Phe	Leu	Leu	Asp	Phe	Ile	Ser	100
															105
Tyr	Pro	Val	Ile	Lys	Gly	Phe	Thr	Ser	Ala	Ala	Ala	Val	Thr	Ile	110
															115
Gly	Phe	Gly	Gln	Ile	Lys	Asn	Leu	Leu	Gly	Leu	Gln	Asn	Ile	Pro	120
															125
Arg	Pro	Phe	Phe	Leu	Gln	Val	Tyr	His	Thr	Phe	Leu	Arg	Ile	Ala	130
															135
Glu	Thr	Arg	Val	Gly	Asp	Ala	Val	Leu	Gly	Leu	Val	Cys	Met	Leu	140
															145
Leu	Leu	Leu	Val	Leu	Lys	Leu	Met	Arg	Asp	His	Val	Pro	Pro	Val	150
															155
His	Pro	Glu	Met	Pro	Pro	Gly	Val	Arg	Leu	Ser	Arg	Gly	Leu	Val	160
															165
Trp	Ala	Ala	Thr	Thr	Ala	Arg	Asn	Ala	Leu	Val	Val	Ser	Phe	Ala	170
															175
Ala	Leu	Val	Ala	Tyr	Ser	Phe	Glu	Val	Thr	Gly	Tyr	Gln	Pro	Phe	180
															185
Ile	Leu	Thr	Gly	Glu	Thr	Ala	Glu	Gly	Leu	Pro	Pro	Val	Arg	Ile	190
															195
Pro	Pro	Phe	Ser	Val	Thr	Thr	Ala	Asn	Gly	Thr	Ile	Ser	Phe	Thr	200
															205
Glu	Met	Val	Gln	Asp	Met	Gly	Ala	Gly	Leu	Ala	Val	Val	Pro	Leu	210
															215
Met	Gly	Leu	Leu	Glu	Ser	Ile	Ala	Val	Ala	Lys	Ala	Phe	Ala	Ser	220
															225

WO 01/46258

PCT/US00/35095

320
 Gln Asn Asn Tyr Arg Ile Asp Ala Asn Gln Glu Leu Leu Ala Ile
 335
 Gly Leu Thr Asn Met Leu Gly Ser Leu Val Ser Ser Tyr Pro Val
 350
 Thr Gly Ser Phe Gly Arg Thr Ala Val Asn Ala Gln Ser Gly Val
 365
 Cys Thr Pro Ala Gly Gly Leu Val Thr Gly Val Leu Val Leu Leu
 380
 Ser Leu Asp Tyr Leu Thr Ser Leu Phe Tyr Tyr Ile Pro Lys Ser
 395
 Ala Leu Ala Ala Val Ile Ile Met Ala Val Ala Pro Leu Phe Asp
 410
 Thr Lys Ile Phe Arg Thr Leu Trp Arg Val Lys Arg Leu Asp Leu
 425
 Leu Pro Leu Cys Val Thr Phe Leu Leu Cys Phe Trp Glu Val Gln
 440
 Tyr Gly Ile Leu Ala Gly Ala Leu Val Ser Leu Leu Met Leu Leu
 455
 His Ser Ala Ala Arg Pro Glu Thr Lys Val Ser Glu Gly Pro Val
 470
 Leu Val Leu Gln Pro Ala Ser Gly Leu Ser Phe Pro Ala Met Glu
 485
 Ala Leu Arg Glu Glu Ile Leu Ser Arg Ala Leu Glu Val Ser Pro
 500
 Pro Arg Cys Leu Val Leu Glu Cys Thr His Val Cys Ser Ile Asp
 515
 Tyr Thr Val Val Leu Gly Leu Gly Glu Leu Leu Gln Asp Phe Gln
 530
 Lys Gln Gly Val Ala Leu Ala Phe Val Gly Leu Gln Val Pro Val
 545
 Leu Arg Val Leu Leu Ser Ala Asp Leu Lys Gly Phe Gln Tyr Phe
 560
 Ser Thr Leu Glu Glu Ala Glu Lys His Leu Arg Gln Glu Pro Gly
 575
 Thr Gln Pro Tyr Asn Ile Arg Glu Asp Ser Ile Leu Asp Gln Lys
 590
 Val Ala Leu Leu Lys Ala
 605

<210> 18
 <211> 324
 <212> PPT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyta ID No: 5530164C91

<400> 18
 Met Ser Val Glu Asp Gly Gly Met Pro Gly Leu Gly Arg Pro Arg
 1
 5
 10
 15
 Gln Ala Arg Trp Thr Leu Met Leu Leu Leu Ser Thr Ala Met Tyr
 20
 25
 30
 Gly Ala His Ala Pro Leu Leu Ala Leu Cys His Val Asp Gly Arg
 35
 40
 45
 Val Pro Phe Arg Pro Ser Ser Ala Val Leu Leu Thr Glu Leu Thr
 50
 55
 60
 Lys Leu Leu Leu Cys Ala Phe Ser Leu Leu Val Gly Trp Gln Ala
 65
 70
 75
 Trp Pro Gln Gly Pro Pro Pro Tyr Arg Gln Ala Ala Pro Phe Ala
 80
 85
 90
 Leu Ser Ala Leu Leu Tyr Gly Ala Asn Asn Leu Val Ile Tyr
 95
 100
 105
 Leu Gln Arg Tyr Met Asp Pro Ser Thr Tyr Gln Val Leu Ser Asn
 110
 115
 120
 Leu Lys Ile Gly Ser Thr Ala Val Leu Tyr Cys Leu Cys Leu Arg

WO 01/46258

PCT/US00/35095

```

125
His Arg Leu Ser Val Arg Gln Gly Leu Ala Leu Leu Leu Leu Met 135
140
Ala Ala Gly Ala Cys Tyr Ala Ala Gly Gly Leu Gln Val Pro Gly 150
155
Asn Thr Leu Pro Ser Pro Pro Pro Ala Ala Ala Ala Ser Pro Met 165
170
Pro Leu His Ile Thr Pro Leu Gly Leu Leu Leu Leu Ile Leu Tyr 180
185
Cys Leu Ile Ser Gly Leu Ser Ser Val Tyr Thr Gln Leu Leu Met 195
200
Lys Arg Gln Arg Leu Pro Leu Ala Leu Gln Asn Leu Phe Leu Tyr 210
215
Thr Phe Gly Val Leu Leu Asn Leu Gly Leu His Ala Gly Gly Gly 225
230
Ser Gly Pro Gly Leu Leu Glu Gly Phe Ser Gly Trp Ala Ala Leu 240
245
Val Val Leu Ser Gln Ala Leu Asn Gly Leu Leu Met Ser Ala Val 255
260
Met Lys His Gly Ser Ser Ile Thr Arg Leu Phe Val Val Ser Cys 270
275
Ser Leu Val Val Asn Ala Val Leu Ser Ala Val Leu Leu Arg Leu 285
290
Gln Leu Thr Ala Phe Phe Leu Ala Thr Leu Leu Ile Gly Leu 300
305
Ala Met Arg Leu Tyr Tyr Gly Ser Arg 310
320

```

```

<210> 19
<211> 445
<212> PRT
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 139115CD1

```

```

<400> 19
Met Thr Leu Thr Gly Pro Leu Thr Thr Gln Tyr Val Tyr Arg Arg 15
1
Ile Trp Glu Glu Thr Gly Asn Tyr Thr Phe Ser Ser Asp Ser Asn 20
20
Ile Ser Glu Cys Glu Lys Asn Lys Ser Ser Pro Ile Phe Ala Phe 30
35
Gln Glu Glu Val Gln Lys Lys Val Ser Arg Phe Asn Leu Gln Met 45
50
Asp Ile Ser Gly Leu Ile Pro Gly Leu Val Ser Thr Phe Ile Leu 60
65
Leu Ser Ile Ser Asp His Tyr Gly Arg Lys Phe Pro Met Ile Leu 75
80
Ser Ser Val Gly Ala Leu Ala Thr Ser Val Trp Leu Cys Leu Leu 90
95
Cys Tyr Phe Ala Phe Pro Phe Gln Leu Leu Ile Ala Ser Thr Phe 105
110
Ile Gly Ala Phe Cys Gly Asn Tyr Thr Thr Phe Trp Gly Ala Cys 120
125
Phe Ala Tyr Ile Val Asp Gln Cys Lys Gln His Lys Gln Lys Thr 135
140
Ile Arg Ile Ala Ile Ile Asp Phe Leu Leu Gly Leu Val Thr Gly 150
155
Leu Thr Gly Leu Ser Ser Gly Tyr Phe Ile Arg Glu Leu Gly Phe 165
170
Glu Trp Ser Phe Leu Ile Ile Ala Val Ser Leu Ala Val Asn Leu 180
185
Ile Tyr Ile Leu Phe Phe Leu Gly Asp Pro Val Lys Glu Cys Ser 195
200
Ser Gln Asn Val Thr Met Ser Cys Ser Glu Gly Phe Lys Asn Leu 210

```

WO 01/46258

PCT/US00/35095

215 220 225
 Phe Tyr Arg Thr Tyr Met Leu Phe Lys Asn Ala Ser Gly Lys Arg
 230 235 240
 Arg Phe Leu Leu Cys Leu Leu Leu Phe Thr Val Ile Thr Tyr Phe
 245 250 255
 Phe Val Val Ile Gly Ile Ala Pro Ile Phe Ile Leu Tyr Glu Leu
 260 265 270
 Asp Ser Pro Leu Cys Trp Asn Glu Val Phe Ile Gly Tyr Gly Ser
 275 280 285
 Ala Leu Gly Ser Ala Ser Phe Leu Thr Ser Phe Leu Gly Ile Trp
 290 295 300
 Leu Phe Ser Tyr Cys Met Glu Asp Ile His Met Ala Phe Ile Gly
 305 310 315
 Ile Phe Thr Thr Met Thr Gly Met Ala Met Thr Ala Phe Ala Ser
 320 325 330
 Thr Thr Leu Met Met Phe Leu Ala Arg Val Pro Phe Leu Phe Thr
 335 340 345
 Ile Val Pro Phe Ser Val Leu Arg Ser Met Leu Ser Lys Val Val
 350 355 360
 Arg Ser Thr Glu Gln Gly Thr Leu Phe Ala Cys Ile Ala Phe Leu
 365 370 375
 Glu Thr Leu Gly Gly Val Thr Ala Val Ser Thr Phe Asn Gly Ile
 380 385 390
 Tyr Ser Ala Thr Val Ala Trp Tyr Pro Gly Phe Thr Phe Leu Leu
 395 400 405
 Ser Ala Gly Leu Leu Leu Leu Pro Ala Ile Ser Leu Cys Val Val
 410 415 420
 Lys Cys Thr Ser Trp Asn Glu Gly Ser Tyr Glu Leu Leu Ile Gln
 425 430 435
 Glu Glu Ser Ser Glu Asp Ala Ser Asp Arg
 440 445

<210> 20
 <211> 337
 <212> PBT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1702940CD1

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

WO 01/46258

PCT/US00/35095

Lys	Glu	Phe	Val	Gly	Gly	Asn	Thr	Pro	Asn	Val	Leu	Thr	Leu	Val	185
Asp	Asn	Trp	Tyr	Gln	Val	Thr	Gln	Gly	Ile	Gly	Arg	Asn	Ile	Arg	190
Ala	Ile	Arg	Arg	Ala	Arg	Ala	Asn	Pro	Gln	Leu	Gly	Ala	Tyr	Ala	195
Pro	Pro	Pro	His	Val	Ile	Gly	Arg	Ile	ser	Ala	Glu	Gly	Gly	Glu	200
Gln	Val	Glu	Arg	Val	Val	Glu	Gly	Pro	Ala	Gln	Ala	Met	Ser	Arg	205
Gly	Thr	Met	Ile	Val	Gly	Ala	Ala	Thr	Gly	Gly	Ile	Leu	Leu	Leu	210
Leu	Asp	Val	Val	Ser	Leu	Ala	Tyr	Glu	Ser	Lys	His	Leu	Leu	Glu	215
Gly	Ala	Lys	Ser	Glu	Ser	Ala	Glu	Glu	Leu	Lys	Lys	Arg	Ala	Gln	220
Glu	Leu	Glu	Gly	Lys	Leu	Asn	Phe	Leu	Thr	Lys	Ile	His	Glu	Met	225
Leu	Gln	Pro	Gly	Gln	Asp	Gln									230
															235

<210> 21
 <211> 273
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc feature
 <223> Incyte ID No: 1703342CD1

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

WO 01/46258 PCT/US00/35095

260 265 270

Glu Val Asn

<210> 22
 <211> 710
 <212> PRU
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1727529CD1

<400> 22
 Met Gly Gly Lys Gln Arg Asp Glu Asp Asp Glu Ala Tyr Gly Lys
 1 5 10 15
 Pro Val Lys Tyr Asp Pro Ser Phe Arg Gly Pro Ile Lys Asn Arg
 20 25 30
 Ser Cys Thr Asp Val Ile Cys Cys Val Leu Phe Leu Leu Phe Ile
 35 40 45
 Leu Gly Tyr Ile Val Val Gly Ile Val Ala Trp Leu Tyr Gly Asp
 50 55
 Pro Arg Gln Val Leu Tyr Pro Arg Asn Ser Thr Gly Ala Tyr Cys
 65 70 75
 Gly Met Gly Glu Asn Lys Asp Lys Pro Tyr Leu Leu Tyr Phe Asn
 80 85 90
 Ile Phe Ser Cys Ile Leu Ser Ser Asn Ile Ile Ser Val Ala Glu
 95 100 105
 Asn Gly Leu Gln Cys Pro Thr Pro Gln Val Cys Val Ser Ser Cys
 110 115 120
 Pro Glu Asp Pro Trp Thr Val Gly Lys Asn Glu Phe Ser Gln Thr
 125 130 135
 Val Gly Glu Val Phe Tyr Thr Lys Asn Arg Asn Phe Cys Leu Pro
 140 145 150
 Gly Val Pro Trp Asn Met Thr Val Ile Thr Ser Leu Gln Glu Glu
 155 160 165
 Leu Cys Pro Ser Phe Leu Leu Pro Ser Ala Pro Ala Leu Gly Arg
 170 175 180
 Cys Phe Pro Trp Thr Asn Ile Thr Pro Pro Ala Leu Pro Gly Ile
 185 190 195
 Thr Asn Asp Thr Thr Ile Gln Gln Gly Ile Ser Gly Leu Ile Asp
 200 205 210
 Ser Leu Asn Ala Arg Asp Ile Ser Val Lys Ile Phe Glu Asp Phe
 215 220 225
 Ala Gln Ser Trp Tyr Trp Ile Leu Val Ala Leu Gly Val Ala Leu
 230 235 240
 Val Leu Ser Leu Leu Phe Ile Leu Leu Leu Arg Leu Val Ala Gly
 245 250 255
 Pro Leu Val Leu Val Leu Ile Leu Gly Val Leu Gly Val Leu Ala
 260 265 270
 Tyr Gly Ile Tyr Tyr Cys Trp Glu Glu Tyr Arg Val Leu Arg Asp
 275 280 285
 Lys Gly Ala Ser Ile Ser Gln Leu Gly Phe Thr Thr Asn Leu Ser
 290 295 300
 Ala Tyr Gln Ser Val Gln Glu Thr Trp Leu Ala Ala Leu Ile Val
 305 310 315
 Leu Ala Val Leu Glu Ala Ile Leu Leu Leu Val Leu Ile Phe Leu
 320 325 330
 Arg Gln Arg Ile Arg Ile Ala Ile Ala Leu Leu Lys Glu Ala Ser
 335 340 345
 Lys Ala Val Gly Gln Met Met Ser Thr Met Phe Tyr Pro Leu Val
 350 355 360
 Thr Phe Val Leu Leu Ile Cys Ile Ala Tyr Trp Ala Met Thr
 365 370 375
 Ala Leu Tyr Leu Ala Thr Ser Gly Gln Pro Gln Tyr Val Leu Trp
 380 385 390
 Ala Ser Asn Ile Ser Ser Pro Gly Cys Glu Lys Val Pro Ile Asn

WO 01/46258

PCT/US00/35095

```

395          400          405
Thr Ser Cys Asn Pro Thr Ala His Leu Val Asn Ser Ser Cys Pro
410          415          420
Gly Leu Met Cys Val Phe Gln Gly Tyr Ser Ser Lys Gly Leu Ile
425          430          435
Gln Arg Ser Val Phe Asn Leu Gln Ile Tyr Gly Val Leu Gly Leu
440          445          450
Phe Trp Thr Leu Asn Trp Val Leu Ala Leu Gly Gln Cys Val Leu
455          460          465
Ala Gly Ala Phe Ala Ser Phe Tyr Trp Ala Phe His Lys Pro Gln
470          475          480
Asp Ile Pro Thr Phe Pro Leu Ile Ser Ala Phe Ile Arg Thr Leu
485          490          495
Arg Tyr His Thr Gly Ser Leu Ala Phe Gly Ala Leu Ile Leu Thr
500          505          510
Leu Val Gln Ile Ala Arg Val Ile Leu Glu Tyr Ile Asp His Lys
515          520          525
Leu Arg Gly Val Gln Asn Pro Val Ala Arg Cys Ile Met Cys Cys
530          535          540
Phe Lys Cys Cys Leu Trp Cys Leu Glu Lys Phe Ile Lys Phe Leu
545          550          555
Asn Arg Asn Ala Tyr Ile Met Ile Ala Ile Tyr Gly Lys Asn Phe
560          565          570
Cys Val Ser Ala Lys Asn Ala Phe Met Leu Leu Met Arg Asn Ile
575          580          585
Val Arg Val Val Val Leu Asp Lys Val Thr Asp Leu Leu Leu Phe
590          595          600
Phe Gly Lys Leu Leu Val Val Gly Gly Val Gly Val Leu Ser Phe
605          610          615
Phe Phe Phe Ser Gly Arg Ile Pro Gly Leu Gly Lys Asp Phe Lys
620          625          630
Ser Pro His Leu Asn Tyr Tyr Trp Leu Pro Ile Met Thr Ser Ile
635          640          645
Leu Gly Ala Tyr Val Ile Ala Ser Gly Phe Phe Ser Val Phe Gly
650          655          660
Met Cys Val Asp Thr Leu Phe Leu Cys Phe Leu Glu Asp Leu Glu
665          670          675
Arg Asn Asn Gly Ser Leu Asp Arg Pro Tyr Tyr Met Ser Lys Ser
680          685          690
Leu Leu Lys Ile Leu Gly Lys Lys Asn Glu Ala Pro Pro Asp Asn
695          700          705
Lys Lys Arg Lys Lys
710

```

```

<210> 23
<211> 476
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> Incyte ID No: 2289333CD1

```

```

<400> 23
Glu Gln Asn Phe Asp Gly Thr Ser Asp Glu Glu His Glu Gln Glu
1          5          10          15
Leu Leu Pro Val Gln Lys His Tyr Gln Leu Asp Asp Gln Glu Gly
20          25          30
Ile Ser Phe Val Gln Thr Leu Met His Leu Leu Lys Gly Asn Ile
35          40          45
Gly Thr Gly Leu Leu Gly Leu Pro Leu Ala Ile Lys Asn Ala Gly
50          55          60
Ile Val Leu Gly Pro Ile Ser Leu Val Phe Ile Gly Ile Ile Ser
65          70          75
Val His Cys Met His Ile Leu Val Arg Cys Ser His Phe Leu Cys
80          85          90
Leu Arg Phe Lys Lys Ser Thr Leu Gly Tyr Ser Asp Thr Val Ser

```

WO 01/46258

PCT/US00/35095

```

95
Phe Ala Met Glu Val Ser Pro Trp Ser Cys Leu Gln Lys Gln Ala 100 105
110
Ala Trp Gly Arg Ser Val Val Asp Phe Phe Leu Val Ile Thr Gln 115 120
125
Leu Gly Phe Cys Ser Val Tyr Ile Val Phe Leu Ala Glu Asn Val 130 135
140
Lys Gln Val His Glu Gly Phe Leu Glu Ser Lys Val Phe Ile Ser 145 150
153
Asn Ser Thr Asn Ser Ser Asn Pro Cys Glu Arg Arg Ser Val Asp 160 165
170
Leu Arg Ile Tyr Met Leu Cys Phe Leu Pro Phe Ile Ile Leu Leu 175 180
185
Val Phe Ile Arg Glu Leu Lys Asn Leu Phe Val Leu Ser Phe Leu 190 195
200
Ala Asn Val Ser Met Ala Val Ser Leu Val Ile Ile Tyr Gln Tyr 205 210
215
Val Val Arg Asn Met Pro Asp Pro His Asn Leu Pro Ile Val Ala 220 225
230
Gly Trp Lys Lys Tyr Pro Leu Phe Phe Gly Thr Ala Val Phe Ala 235 240
245
Phe Glu Gly Ile Gly Val Val Leu Pro Leu Glu Asn Gln Met Lys 250 255
260
Glu Ser Lys Arg Phe Pro Gln Ala Leu Asn Ile Gly Met Gly Ile 265 270
275
Val Thr Thr Leu Tyr Val Thr Leu Ala Thr Leu Gly Tyr Met Cys 280 285
290
Phe His Asp Glu Ile Lys Gly Ser Ile Thr Leu Asn Leu Pro Gln 295 300
305
Asp Val Trp Leu Tyr Gln Ser Val Lys Ile Leu Tyr Ser Phe Cys 310 315
320
Ile Phe Val Thr Tyr Ser Ile Gln Phe Tyr Val Pro Ala Glu Ile 325 330
335
Ile Ile Pro Gly Ile Thr Ser Lys Phe His Thr Lys Trp Lys Gln 340 345
350
Ile Cys Glu Phe Gly Ile Arg Ser Phe Leu Val Ser Ile Thr Cys 355 360
365
Ala Gly Ala Ile Leu Ile Pro Arg Leu Asp Ile Val Ile Ser Phe 370 375
380
Val Gly Ala Val Ser Ser Ser Thr Leu Ala Leu Ile Leu Pro Pro 385 390
395
Leu Val Glu Ile Leu Thr Phe Ser Lys Glu His Tyr Asn Ile Trp 400 405
410
Met Val Leu Lys Asn Ile Ser Ile Ala Phe Thr Gly Val Val Gly 415 420
425
Phe Leu Leu Gly Thr Tyr Ile Thr Val Glu Glu Ile Ile Tyr Pro 430 435
440
Thr Pro Lys Val Val Ala Gly Thr Pro Gln Ser Pro Phe Leu Asn 445 450
455
Leu Asn Ser Thr Cys Leu Thr Ser Gly Leu Lys 460 465
470

```

```

<210> 24
<211> 237
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2720354CD1

```

```

<400> 24
Met Gly Leu Thr Phe Ile Asn Ala Leu Val Phe Gly Val Gln Gly 1 5 10 15
Asn Thr Leu Arg Ala Leu Gly His Asp Ser Pro Leu Asn Gln Phe 20 25 30
Leu Ala Gly Ala Ala Ala Gly Ala Ile Gln Cys Val Ile Cys Cys

```

WO 01/46258

PCT/US00/35095

```

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

```

<210> 25

<211> 345

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte JD No: 3038193CD1

<400> 25

```

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

```

WO 01/46258

PCT/US00/35095

215
 Phe Ile Ile Phe Ser Ile Gln Leu Ser Phe Met Leu Leu Thr Phe Gly
 230
 Ile Phe Ser Thr Arg Asn Asn Ser Gly Phe Thr Pro Ala Asp Thr
 245
 Val Ala Ile Ile Phe Cys Ser Thr His Lys Ser Leu Thr Leu Gly
 260
 Ile Pro Met Leu Lys Ile Val Phe Ala Gly Tyr Glu His Leu Ser
 275
 Leu Ile Ser Val Pro Leu Leu Ile Tyr His Pro Ala Gln Ile Leu
 290
 Leu Gly Ser Val Leu Val Pro Thr Ile Lys Ser Trp Met Val Ser
 305
 Arg Gln Lys Lys Leu Leu Gln Thr Arg Gly Pro Leu Ala Asn Leu
 320
 Asn Asn Pro Glu Gly Leu Glu Tyr Leu Ser Ile Lys Phe Gly His
 335

<210> 26
 <211> 521
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3460979CD1

<400> 26
 Met Ala Ala Leu Ala Pro Val Gly Ser Pro Ala Ser Arg Gly Pro
 1
 Arg Leu Ala Ala Gly Leu Arg Leu Leu Pro Met Leu Gly Leu Leu
 20
 Gln Leu Leu Ala Glu Pro Gly Leu Gly Arg Val His His Leu Ala
 35
 Leu Lys Asp Asp Val Arg His Lys Val His Leu Asn Thr Phe Gly
 50
 Phe Phe Lys Asp Gly Tyr Met Val Val Asp Val Ser Ser Leu Ser
 65
 Leu Asn Glu Pro Glu Asp Lys Asp Val Thr Ile Gly Phe Ser Leu
 80
 Asp Arg Thr Lys Asn Asp Gly Phe Ser Ser Tyr Leu Asp Glu Asp
 95
 Val Asn Tyr Cys Ile Leu Lys Lys Gln Ser Val Ser Val Thr Leu
 110
 Leu Ile Leu Asp Ile Ser Arg Ser Glu Val Arg Val Lys Ser Pro
 125
 Pro Glu Ala Gly Thr Gln Leu Pro Lys Ile Ile Phe Ser Arg Asp
 140
 Glu Lys Val Leu Gly Gln Ser Gln Glu Pro Asn Val Asn Pro Ala
 155
 Ser Ala Gly Asn Gln Thr Gln Lys Thr Gln Asp Gly Gly Lys Ser
 170
 Lys Arg Ser Thr Val Asp Ser Lys Ala Met Gly Glu Lys Ser Phe
 185
 Ser Val His Asn Asn Gly Gly Ala Val Ser Phe Gln Phe Phe Phe
 200
 Asn Ile Ser Thr Asp Asp Gln Glu Gly Leu Tyr Ser Leu Tyr Phe
 215
 His Lys Cys Leu Gly Lys Glu Leu Pro Ser Asp Lys Phe Thr Phe
 230
 Ser Leu Asp Ile Glu Ile Thr Glu Lys Asn Pro Asp Ser Tyr Leu
 245
 Ser Ala Gly Glu Ile Pro Leu Pro Lys Leu Tyr Ile Ser Met Ala
 260
 Phe Phe Phe Phe Leu Ser Gly Thr Ile Trp Ile His Ile Leu Arg
 275

WO 01/46258

PCT/US00/35095

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

<210> 27

<211> 555

<212> PRT

<213> Homo sapiens

<320>

<221> misc_feature

<223> Incyte ID No: 7472200CD1

<400> 27

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

WO 01/46258

PCT/US00/35095

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

<210> 28
 <211> 2080
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1416107CS1

<400> 28
 ggcdggttcag ggcgcagagc tggccgctag gggttggccg ccgacalga ccccagggac 60
 ccagcggaaa cccagccgct tctggggcct cctggggcca ggcgcccgc cggcgcggc 120
 gtcttctcgg ccgccttcgc cgtgcccctg ggcacactca gcttgggctt cgcgctcgc 180
 tacagctccc cggccatccc tagcctgcag cgcgcgcgc ccccgccccc ggcctcggac 240
 gacgcgcgcg cctcctggtt cggggctgtc gtgacctggt gtgcgcggc ggggggagtg 300

WO 01/46258

PCT/US00/35095

```

ctgggggct ggtgggtgga ccggccggg cgaagctga gctcttget dtgctcgtg 368
ccctctgtg ccgctttgc cgtcaccac gggcccagg acygtggat gctgctggg 429
ggccgctcc tcaacggcct ggctggggt gttgctccc tagtggccc ggtctacac 480
tcgaaatcg cctaccacc agtccgggg ttgctggct cctgttgcg gctaatgct 540
gtgtgggca tctctctgg ctaaccggca ggtgggtg tggagtggc ctggctgct 600
gtgtgggct gctgcccgc ctcccctat ctgcttcca tgtgtctcat gcccgagcc 660
ccygcctcc tctgactca gcaaggcgc caggagcca tggcgcctt gccgttctg 720
tgggctccg gcaaggctg ggaagcccc ccatgggg ctgagcagag ctltcaactg 780
gctctctgc ggcagccgg catetacag ccttctaca tggggtctc cctgatgac 840
ttccagcag tctgggggt caacgcctc atgttatag nagagccat cttgagag 900
gccagttca agyagcag cctggcctc gctgctgg gttctcca ggtgttctc 960
acagctgtg ggtctctat catggacaga gcaagggca ggtgtctc ggtctgtca 1020
ggtgtgtca tgggttctc cactgctcc ttcgggctt actcaagct gaccagggc 1080
ggccttgca actctctga cgtggccatc tggggcctg tctctgaca gcccttga 1140
gcaagctgg gctggctg gctggcctg ggcagctgt gctctctat gcccgctt 1200
tgggtggct gggggcctt cctctgctc ctactgtg agtcttcc tctgactg 1260
aagggtgg gcaagcct ctgctctc accaacctgc tcatggctt tctgtgac 1320
aagggttca gcaagcctc ggaagctct agccctatg gacctctct gcttctcc 1380
gcttctgca tctcaagt ctcttctc ttgtttgt tctctgaa ccaaggaa 1440
actctggac aatctacc cactttct ggggatgac agcactca tagggatg 1500
agcaagctg tgaactcaag ctgggcca cccagagcc ctgctctgc ccaggagc 1560
cagatctcc cctctggag ccttctctc caggtctct cctctctgc atgctctcc 1620
cagcaatg ccaaggctc ggaagctca tggctctgt ccaagctct gctgtgct 1680
tggagctca gcaagcctt cagctctgc agcctgag tcaagctct atgcaaga 1740
ctaaagcag gcaagagag ggggctctt agyatttt tctctgctt aggagctt 1800
ctggagctg gctgctggc attcaatgc tctctcaag cggctgctt atcggaag 1860
aaatttctt gccaaataa cactgacac gaaatcag tcaagctct tgggtttgt 1920
gcaagctag ttgaaagg gtttatccc atcaactgc aggacacct gttgcttac 1980
tgtctatgt tcaagcagc taccctca cactgagag tcaattctg ctactctct 2040
ggctcagtt cctgggtca tcaagcaca atctgtgt 2080

```

<210> 29
 <211> 2128
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1682513cs1

```

<400> 29
ctggccctag gtagctgccc ctgtctgtg ctgcccgcac caaccagccc cacattgca 60
actacctgac ggagaacccc caaagaagg cggccatgcy ggcagagac tggcgagca 120
acacagtct gcatggctg gttggcattg ctgacaacac cctgtagaac acaagtttg 180
ttaccaagat gtaagacctg ctgctgctca agtctgccc cctcttccc gacagcaac 240
tggagctcgt gctcaaacac gaagcctctt cggcctcat gatggctgc aagcgggca 300
agattggaa ccgcaagag atgctgctg tggagccat caatgactg ctgagagca 360
gtgtggcga gttggggcc gttctctctt acataacgt ggtctctac ctgtgtgca 420
tgytcactt cactctacc cctcaclacc agccctgga gggcacacc cctgcaact 480
accgcaaac ggtggactac ctgctgtgg ctggggaggt cattaagctc ttaactggg 540
tctgttctt ctcaacac atcaagact tgttcatgaa gaatgccc gtagtgaat 600
ctctctctat tgaagctcc ttccagctgc tctacttcat ctacttgct ctggtgatc 660
tctcagcag cctctacctg gcaaggtatg agcctactt ggcctgatg gtctttgca 720
tggctctggt ctgagaaat gcccttact tcaaccgtg gctgagctg accgggacc 780
atagcatcat gatcagaag atctctctca aggaacctt ccaatctg ctgctctact 840
tgtctctat gatggctac gctcaagccc tggctctctt cctgaaaccg tgtgcaaca 900
tgaagtgct caatggggac cagacaact gcaagctgca caetacccc tctgctgct 960
caagcagac ctcaagccc tctctctgg acctgttaa gctgacacc ggcattggc 1020
acctggagat gctgagcag accaagctc cctgttctt catcaactg ctggtgacc 1080
acatcaact cactttgtg ctgtctctca acatgctat tgcctcatg ggcagagac 1140
tggccaggt ctcaaggg agcaagcaca tctggaagct gaagtgggc accaacatc 1200
tggacattga gctctctct cccgtatctc tgaagaaatc cttccgctt ggggagatg 1260
tcaagctgg gcaagctcg gacggcctc ctgacggag gtagtcttc aggtgagt 1320
aggtgactg gctctctgg aacagaccl tgggctcat caacagag ccggcaaga 1380
atgagacct cagatattt ggtctctgc ataccgtgg cgcctcgc agggatcct 1440
ggtctctgtt gttaccacc gttgtggaac tgaacaagaa ctogaaccc gacgagctg 1500
tgytctctt gcaagcagc gggaccccc gctgctgct ccaacagag ggttaccoc 1560

```

WO 01/46258

PCT/US00/35095

```

gcaagtgag gactgatgac gccccegetct agggactgca gcccagcccc agtctctctg 1620
ccactcatt tctatgtccag ccgcatttca gcaatgctct ctggggtgtc ccccacacc 1680
ctgtcttgg cccagaggcg agggaccagt ggaagtcca gggaggcccc agaccctgt 1740
ggtcccttg ctctgctcc ccacctggg ggggggtc cggccccct gcttctctc 1800
tataggatca cctaaagccaa gcccagagcc cctccacctc agggcccagc ccttctctc 1860
ccattattta ttgctcttgc tctcaggaag cgaagtgcacc cctgccccag ctggaacctg 1920
gcaagagct tggagcccg ttcaagtgc actgcccgc cagccccag cctcagctg 1980
gctctgact gcatggcca ccatttttg caagtggca gctttcaag gggctgggc 2040
cctggctgt gggccatgcc ttctgtgtg tctgtatgt ctgggattg ccgtgtctc 2100
atcaatgitt atcaattgaa aaaaaaaa

```

```

<210> 30
<211> 2825
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2446438CB1

```

```

<400> 30
cgttggcac gtaattggc tgcagctgtg tccagatggt cagtctctgg tggctagct 60
gctctgcaag gggagcgtta agctccggtt ctccaccctg ccgctgggc agtctggctg 120
aggtgcaacc agagaccaga acctgcttgc tggagcttag tgcctagagc tggggaggga 180
ggttcccgcc ctctctctct gtcagagccg gcaagccctc ccggtctcnc ttctccccc 240
agccctctct actgagaagc tccggatcc cagcagccgc cagccctgg cctcagcctg 300
cggggctcca gtcagggcaa caccagcgc cagctggggg gaagacagga ccttgcacat 360
ctccctctgc acagaggtcc tggctggacc gacagccctc ctctctctag gatgacctca 420
ccctccagct ctccagttt cagctggag acattagatg caggcccaga agatggctct 480
gagggcagca gaggaaact ggaatttggg acgaggctgc ctccctgga gtcacagttc 540
caggggag agggaaatt cgcctctcag ataaagctca acctcaacta ccaaaaggga 600
accagtgca gtcagccgga tccaaaccga ktggaccgag atcgctctt caatgcggtc 660
tcccgggtg tcccaggga tctgcttgg obtccagagt acctgagcaa gaccagcaag 720
taactcaag actggaata cacagaggg tccacagta agcagtgct gatgaagct 780
gtctgaacc ttaaggaggg ggtcaatgcc tgcattctgc cactgtgca gatcagccg 840
gactctgca atctccagcc cctgtaaat gccagtgca cagatgacta ttaccagggc 900
cacagcctc tgcacatgc cattgagag aggagctcgc agtctgtgaa gctcctgtc 960
gagcaaggg ccaatctgca tgcctgggct tgcggcctct tcttccaaa gggccagggg 1020
actgtcttt attctggtga gctccctcag tcttggccg ctggaacca ccagtggat 1080
tggtaagct acctctgga gaaccacac cagcccctca gctgagcc cactgctcc 1140
cagggcaaca cactctgca tgcctagtg aluatctcg acaactcagc tgaagacatt 1200
gcactgtga ccaqcabtga tgaagggtc ctccaaagct gggcccgcct ctgcccctac 1260
gtcagcttg aggacatcc caactgag gatctcagc ctotgaagct gggcccaag 1320
gagggcaaga tggagatttt caggcacatc ctgcagcgg agtttccagg actgagccac 1380
cttcccgaa agttcaaccg gtggtctat gggcctgtcc ggggtgctct gtagtacctg 1440
gcttctgtg accgtctgga ggaagactca gtgctggaga tcaatgctct tcaatgcaag 1500
agcccacc gacaacgaa gtgtgtttg gaggccctga acaactgct ccaggcgaaa 1560
tgggtctgct tcaaccaca gttctcttca aactctctg tgaatctgat ctacagtctc 1620
atctcaagc ctgttgccta ccaatagct accctgaaga agcagccgc cctcaccctg 1680
aaagccaggg ttgaaactc catgctctg accggccaca tcttctctct gctagggggg 1740
atctactcc cctgtggcca gctgtgttac ttctggcgg gccacgtgtt catctggatc 1800
tctctatag acagctactt tgaatcctc ttctgttcc aggcctctct caagctggct 1860
tcccaggtc tgttttctc gccatcagc tggtaactgc acctgctgt gctctgctg 1920
gtcttggct ggtgaaact gcttactat acagctgct tccagcacac aggcactac 1980
aggtctatga tccagaggt catctctggt gaccctctg cctctctct gactactta 2040
gcttctctt tggcttgcg tctagccctg gtgagctga gccaggggc tgggcccc 2100
gaagctctca cagggccca tgcacagag tcaatgctgc ccatggaggg ccagggagac 2160
gagggcaagc gggccagta cggggctatc ctggaagct ccttggagct ctccaattc 2220
accatggca tggggagct ggccttccag gacagctgc acttccggg catggtgctg 2280
ctgtctctc tggcctaagt gctgctcaac taccctcgc tgotcaaat gctatctgc 2340
ctcatagcc agaccctca cagtgctcc actgacagct gggactctg gaagctgca 2400
aaagctatct ctgtcttga gatggagaa ggtatctgt ggtgcaggaa gaagcagcgg 2460
gcaggtgga tctgacagt tggactaag ccagctggca gccccagta gpcctgtgc 2520
ttcaggtg agaggtgaa ctggcttca tgggagcaga cgtctccac cctgtgtgag 2580
gaccctcag gggcaggtg cctccgaact ctccaganc cttctctggc ttcccctcc 2640
aagggagatg aggatgtgc ctctgaggaa aactatgtc ccgtccagct cctccagctc 2700
aactgaggg ccagatgag cagggagca yaggacagat ttcccaacc 2760

```

WO 01/46258

PCT/US00/35095

atctgtgtgc tetgggtgccc cagtgaattc tgggtggcaaa tatatatattt cactaaaaaa 2820
aaaaa 2825

<210> 31
<211> 1718
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2817822CB1

<400> 31
gcctcggctg tcccaactag gggcgggag ccaggggccc tcccgctgpc ccaagtgale 60
tgcctctagg agggctggc agtggggagg ccagtgggca ggatgagag ccagaccact 120
ctgccccagt ccccccggc cagggggcca agatgtcta ctgtgtgga cctgaacctc 180
gggggtgagt tccacaccac cccctgggt accctgggga aglittcggg ctcaaacctg 240
gcagagatgt tctctagctt agcccaaggcc tccacggcag cggagggcgg cttcttcatc 300
gaaccgcccc gaactattt cagaccatc ctggactacc tgggcactgg ccaagtgcoc 360
accacgcaca tccctaaagt gtaccgtgag gctcagttct accgaatcaa gcccttggtc 420
aagctctctg agaacatgoc acagatcttt ggtgagcagg tgtctgggaa gcagtttttg 480
ctgcaagtgc cgggtctaac cgggaacctg gagctctggt tggcctctgg acgtgcagaa 540
gccataacag cagggaagtc cagctctgctt gctgtcttgg tggcaactga gggacagat 600
gcataattat cagaggctct gtgtttctct gaggataaga agtgttcca gtcctgtgtc 660
aagtltgggc cctggaagcc ggtcctagac aacagcagcc tcatgcactg cctggaagt 720
acatlaagc cccaggggta caagctatc tccaaqtctt acctgacgta ccccaaccaa 780
agaaagcaat tccattttaa catitttcca ttcacctcca cctggtggtg atcctcagga 840
gcagagactg ttatgnaatc tggcgtggct tatgaaatta aaagtggcca tcaagccat 900
ttctttttaa ttcccaaac atcaggcaat tccagggtt ggtctagagt cttgccacta 960
aaatattgac actcgtttbaa ggaacttcca ctccattgca actgagtgcca ctatatattg 1020
ctagcaactg ccagctactt ccttttccaa gccctatgta tctccagag cctctctctg 1080
aagttcaata accagaccaa gtaagaatgt tccaacatg cgttggcaag agatgtcaga 1140
tgacaacagc aacafacaag atactgtgaa tctagatgtt ctgacctaaa gatgtagctc 1200
acatagcccc agcttggggt ccaatccatc tgtccctggc atgtccttcc atgtagtagg 1260
tgccttccctg atcccctttg ccagatgctg tgggtgctaa caccacagag ctgtcctctt 1320
ctctagagtg gaggttttca aagtgcatca ttagcattac ctgtgaactt gctggaata 1380
caaatcctca gcccceact ccagcctact gaatcagaat cctctggggt tggcacagca 1440
ttctgattta ccaaaccttc caagtatct tgatgattc taactttgag accatctcta 1500
gaaaagatt gctacctctt gtatggaggt acaaaagact gactcttacc atcaaggaac 1560
tctctttccc agagctctcc atgggallca gctgaagcca gctctctctc gacagcaact 1620
tctactcag ttttttctt ctgtctcag ctgtctccct cactccctt cctctaaag 1680
cactccatca ataaacctc tggcagagaa aaaaaaa 1728

<210> 32
<211> 2000
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4009329CB1

<400> 32
gaccgaattg aaaccagggg gtgtctctyt tgaacttggc gccagataga gtaactgga 60
ctccagttgg aggggtctgg gagaaccata gaagaggaag ggcctgtctc tccgtggaca 120
ggccaccgga gccccagct gtttggaaat gactactgc aaaaaggaa gtaggagcta 180
aggccagcgc cccctggggc cagatggcgc gcaagaagct gaactctgoc tggcactcga 240
gtgtgttttg tgtctgcta atggcggaga cagtctctgg gaactagggc tegtctacag 300
gagctcaact tagccccag tttccagctt caggtgtgaa ccagaccoccc gttgtagact 360
gccgcaaggt gtgtggcctg aatgtctctg accgtgtgga ctctatccgg aaccaacctg 420
accgccaag tgatgggggg taccctgact acctggaag catctctgoc cacttccctc 480
ccagctccc ccctcggct gtcactctct accttctctg cctgtctctc ctgtttctga 540
ttctggaggt caaccagcc agtttttct gcccaacctt gtcggcctt tctaccacc 600
tgaactctc ccaaacctg gcagggctca cctctctgoc atttgggaat ggtgcacctg 660
acatctttag tgcctctggt gccctctctg acccgaccac acccgccctg gcccttgggg 720
cactgtttgg cgtctggctg ctgttaccac cagtgtggc cggagccatt accactctac 780
acccttccat ggcctgctcc agggcctct tcaaggaact cgtttctctc atggtggctg 840

WO 01/46258

PCT/US00/35095

```

tggtccctgac cttccctcatg ctctctccgtg gaagggctaac cctggcatgg gctctggggtt 960
acctggggctt gttatgtgttc batgtgttca ctgtgatctt ctgcaactgg atctcaaac 960
ggcaaacggag aggatctctg ttctggccca tggcaagttac tccagagato ctctcagact 1020
ccgagggaga ccgggtatct totaatacca acagctatga ctacggtgat gaggaccgco 1080
cgcgttctt ctaccaggag accacggctc agatccctgt ccgggcccct aaacctctgy 1140
atccatgaa gtggagaagg aatcagcat actggaaagc cctcaagggt tccaagctgc 1200
ctgtggagct cctctgctc ctccacagtc ccgtctggga ccggcaaac gatlaccaga 1260
actggaaagc gccctcaac tgtctgcatc tggatctcag cccctsgtt gtagtctgy 1320
cctgcaatc ggggacctat ggtgtctatg agatagcgg cctgtctccc gctgggtctg 1380
tgggtgtgat ccgagggaca gccctggctt cagtacctt tttgccaac tctgcccgc 1440
agccccccag gcttctactg ctctttgctt tctctggctt tctgcccgc gcccttgya 1500
tcaacggggc cgcacagag gtgggtgaaa tcttggctc cctgggtgtg gctctccgy 1560
tgagcaaac tgtgtctggg ctcaagctgc tggcctgggg gaaacagctt gggatgctc 1620
tctggtatt cacactggct cggcaggctt acccaggat ggcgctctcc gctcactttg 1680
ggggctcat ctccaacat ctctgggtg tggggctggg ctgctctgc cagatctccc 1740
gagcccaac agaggtgag ctggagagc ccggactgt ggtgtgggtc ctggcaggg 1800
cctggggct cagcctctgc ttctcctgg tctcagctcc attgcaatgc tccagctca 1860
cgaggtctta tggctctgc ctgctctct tctactgaa ctctctgtc gtggccctc 1920
tcattgaatt tggagtgatt cacctgaaaa gcattgtact gaagcagctt agtctgtgtg 1980
ctcactgca ggcaggagcc 2000

```

```

<210> 33
<211> 2216
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<223> Incyte ID No: 6610083CB1

```

```

<400> 33
gaaaactott cctgaaggag atgcaaggga agattcgaac tggaggaaaa cctaaaata 60
aacaaataaca acaaagttc aaaaactgaa aag-gaacca tgaagctag laaaaaggcc 120
cgaaggagag atgaagaaag tgaatcagc aaaaagaaat tggactggtc ctctctctc 180
ctcgtggcct cctctggcgg cgcctctggc tctctctctc tctaaaggta caactctct 240
gtggtgaatg ccccccccc gtaactcaag gccctttaca atgagctatg ggaagaggg 300
catggacttc caatgagacc agacactctg actttgctct ggtctgtgac tgtgtccaa 360
tctgccaatg gttgacttgt ggggagctta attgtgaga tgaatggaaa ggtctctggg 420
aggaagcaca ctctgtctgc caataatggg ttgtgacttt ctctctgact ctctgagcc 480
tgtctctctc aggcaggagc ctttgaatg ctctctctgg gacgtctca catgggata 540
gatggagcgc tgcctctcag tgtctctccc atgtaactca gtagatoto acccaaggag 600
atcctggctt ctctggggca ggtgactgcc atctttatct gcatggcgt gtlcacttgg 660
cagctctctg gccctccaga gctgctggga aaggagagta cctggccata cctgtttgga 720
gtgattgtgt tccctggcgt tgtccagctg ctgagccttc cctttctccc ggaagccca 780
cgtcactctg tcttgagaa gcaaacagag gcaagagctg tgaagcctt ccaaacgttc 840
tgggttaag cagacttct ccaagggtta gaggagctc tggctgagag ccaactgca 900
aggaactct cctgtgtgct cgtctggag ctgctgagag ctctctactt cctctggag 960
gggtcaacc tgaatgtcac catggctgc taaccgctct gtagcctca tgaactttg 1020
ttctatacca aacacatctt tggaaaagct gggatccctc tggcaaatg cccatctgc 1080
acctgagta cagggggcat cggactcttg gctcgcctct tctctggttt ggtcattgag 1140
cacctggag ggaagacct cctcattggt ggttttggc tcaatggcct cctcttggg 1200
acctctccca tcaactgac cctcaggac caagccccct gggccccca cctgagatc 1260
gtggcaatc tggcaatct cgcctcttct tgcagtgggc cagctgtttt ccaagaaga 1320
aggtcaatg tcaacttct atctgagtga aaagtgtacc tctctctcca ccaatgaca 1380
caaacagcc agatctgact ctctctgata tctgctgaaa gttcttggct aaccaaaat 1440
cactaaactt agcctctct gtttttttt tctaaagccc tcccaagact tttgcaatg 1500
atcctgarto tgttcaaggt gtttgaact gtggtctct tttgactgta gaaactgct 1560
catttccag cctttaaag ctgggctccc catcagctc tatgggactc cctgagagg 1620
agggcactg cacctccaa tccagatca cctgtagcc cctgcccctc gctctccaa 1680
tcaatctca acccctgtg ttgcccagc acctgggctt tctggttag caatgactt 1740
agcaaaaga tggaccagg ttagaagct tcaatgaaac tcaactgac agtctadag 1800
tcaaacctc agggactta cctgtctaac aaaaagctg accatgaca tgaagcaatc 1860
ttgcacttc ctcaatggac agggagagc aagctccag ggaagctacc cgggaagtg 1920
tggcaaggag atgtctagag ctgaaatgca gaaagctca tgggctctct ctctgactt 1980
aagcaaggag gatgactct ccaaggagag gccaggagcc cgcctggggc agctctctg 2040
aggaacaggl cgaatgaaq agacttgaca aggagttgaa attagttgaa gcaaacgaa 2100
gaaaccaaga gaggcagttt cctgctgcat attttatgt tgtgataac ccaaggcag 2160

```

WO 01/46258

PCT/US00/35095

tggcagggaa gtctaataaa tgaggcaaaa taaaagagct tcccttttta aaaaaa 2216

<210> 34
<211> 1995
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7472062CB1

<400> 34
atgacggaaa aaaccaatgg tgtgaagagc tccccagca ataatoacaa ccatcatgca 60
ctctccgcca tcaaggccaa tggcaagat gaccacagga caagcagcag gccacactct 120
gcagctgacg atgacacttc ctcaagaacty cagagcttgg cagacttggc tgcctccacg 180
cgggaaagga gtggctctcc caggatagtt cgcctgttgg ggtatctcag agaatgggca 240
aacaaatct tccgagagga ggaacctagg cctgactct tctcggagcg tttctgttgg 300
ctcgaacttc agactgtgac cacacagag ggggagggca aagycagcaa ggalggcag 360
gacaagagca ocaagaagaa atttgaacta tttgtcttgg acccagctgg ggattggtac 420
taactgttggc tatttgcctat tgcctgtccc gtccctttaca actggtgctt gctgttggcc 480
agagctgct tcaqtgacct acagaaaggg tactacctgg tgtggttgg tctggtatfat 540
gtctcagatg tggctctcat tgcggacttc tcatctcgat tggcgcaggg tttctctggag 600
caggggctgc tggcaaaaga taccagaana ctgcagagca actacatcca caactctcag 660
tccaagctgg atgtggcttc catcactccc actgacctga tctatcttgc agtggacatc 720
cacagccctg aggtgtgctt caaccgctgy ctgcacttgg cccgactggt tgaattcttt 780
gaccggacag agacaagcag caactacctt aacatctctc gcatcagaaa ccttgccttc 840
tacctcttgg tcatcatcca ctggaatgcc tgcctctatt atgcaatctc caaatccata 900
ggcttttggg tggacacctg ggtttaccca aacatcacty accctgagta tggctactgt 960
gctagggaaat acatctattg cctttactgg tcaacactga ctctcctac ccttggggag 1020
accaccacc ctgtaaaagga tgaaggaglac ctatcttcca tctttgactt cctgaltggc 1080
gtctctactc tttcccactc cylvggaaat gtggctcca tgccttcca cctgaltggc 1140
accgggcaag agttcccaag taagatcgat gcgtgaaac actcaatgca gttcggaaag 1200
gtcagcaagg ggatgcaagc caaggtcatt aggtggtttg actacttctg gaccaataag 1260
aagacagctg atgagggaga aattctcaag aatctgcag ccaagctcag gctgagata 1320
gcactcaatg tccacttctc cacactcaag aaagtgcgca tcttccatga ttgtgaggt 1380
ggcctgctgg tagagctggt actgaaactc cgtctccag tcttcaagtc tggggattac 1440
atttgcgcga aaggggacat cggcaaggag atgtacatca ttaaggaggg caaactgpcn 1500
gtggtgctg atgattgtgt gactcagtat gctctgtctt cggctggag ctgcttggc 1560
ggatcagta tctttaaact taagggcagt aaaaatggca atgcagcac agttaaattc 1620
cgcagcttgg gctactcaga tctctcttgc ttgtccagg atgacttat ggsagctgtg 1680
actggtacc ctgatccca gaaagctcta gaagagcggg gtcggagat cctcatyaa 1740
gagcactgc tggatcagaa cgaagtggca accagctgg aggtgcact gcaggagaa 1800
ctagggcagc tggagaccaa catggaacc tttacatctc gcttggccg cctgtggct 1860
gagtacacgg gggcccaagca gaagctcaag cagcgcata cagtcttgg aaccaaagt 1920
aaacagaca atgaaatga ctacctgtct gatggatga acagcctga gctgctgtc 1980
gctgacggc cataa 1995

<210> 35
<211> 988
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1812692CB1

<400> 35
cttgggtgaa agaaaaatct gcttgacaaa aaccgtcact taggaaaaga tgtcctttcg 60
ggcagccagc ctacagatga ggaacagaag gaatgacact ctggacagca cccggacct 120
gtactccagc gcgtctcggg gcaacagact gtcttaccgt gaaagcact tgggtaatt 180
tattcaagca aattttaaga aacgagaatg tgtctctttt accaaagatt ccaaggccac 240
ggagaatctg tgcaggtgtg gctatgcccc gagcagcac atggaaagca cccagatcaa 300
ccaagttag aatttgaact acaagaacaa caccaaagaa ttctctacag acgcttgg 360
ggatattcag tttgagacac tgggaaagaa agggagat atactgtgt cctgpcac 420
gaaagcggaa alcctttacg agctgtgtac ccagcactgg cacttgaana caccacact 480
ggctallctc glgacgggg ggcnaagaa cttagccctg aagccgcqca tggcagagat 540
ctcagccgg ctactctaca tccggcagtc caaaggtgct tggattctca cgggagcac 600

WO 01/46258

PCT/US00/35095

```

ccattatggc ctgatgnaagt acatcggggg ggtgggtgaga gataaaccca tcagcaaggg 660
tccagaggag aatattgtgg ccatctggcat agcagcttgg ggcattggtct ccaaccggga 720
cacccctatc aggaatttgg atgtctgagg accggtggga caggaggagg totgtatggt 780
cacatggaaq aaagaccatg gcatggccct gtggcctgaa ccttggggct ctgtalggaa 840
gcccagccaga tcatggggaa gtctggcctt caaggagtgc uttgggacc ttaaaggaa: 900
tgaaaacaag gatgacgtac ctaattaact gctgggaaag agttaacaat gaatgtttt 960
tgcattaaaa tgtttctca gcaatctc

```

```

<210> 36
<211> 3179
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> Misc_feature
<223> Incyte ID No: J232992CB1

```

```

<400> 36
gcccagcggc ggcgcggggc ccggggggcg cagcaggggg ctggcggtag cggttgtgc 60
ggggcccggg ggcgcggggc cgctggggtc tggcgcgggg gcatgagggt gcagacgtg 120
tggggcagcg taaggggggc cccgaccgga cccccggcac ccccgctgc 180
ggactactg caaaggggcc ccggcgctca gcaacccaaa ccggccagct tggggccggg 240
gcccggggca agccgcagcc atctcagct tggcaaaagt gctcaactac ctggaacagt 300
acccctggc agggctctct ctggacatcc agcagcaact tgggttcagg cccgagggc 360
ccggcctgct gcaatcagtg ttcactgtta gctctatggt ggttgcctcc atuttggct 420
ccctggcgca ccgctcaaac aggaaggtga ttctcagctg ggcattttc ttctgttgg 480
cctcaccctt ctccagctcc ttcattcccc agcaglatct ctggctgtgt gctctgtcc 540
ggggcctggt gggcctgggg gaggccagct actccaccat cgcctccact atcattggg 600
acctctccac caagaaacag cgtacgctca tgcctcctgt cttctacttc gccatccacc 660
tggcagctgg cctgggctac attactgact ccacgctgaa gcaggcagcc ggagactggc 720
actgggcat aggggtgtcc cctgtcttgg gcatgctac aggaacatc atctctatc 780
tggctccagc cactaaaagg ggtcagtcag accagctcgg ggcacagctc aaagccggga 840
cctcattgcl ccgagatag aaggccctga ttcgaaacc cagctacttc ttctctccc 900
tggccagctc ggtctgtctc ttgcccaagg gggccctagg catgtggalc cctgtctacc 960
tgccaccgct ccaagtgtgt cagaagacag cagagacgtg caacagcccg cctgttgggg 1020
ccaagacag cctcactctt ggggcaatca cctgctctac gggattctg cgtctggtca 1080
cggggcagg agccacgggc tgggtccgcc tgaagaccca gcccgcggac ccactggtgt 1140
gtgcccctgg catgctgggc tctgccatct ccatctgctt gatcttgggt gctgtccaa 1200
ggagcaactg agggagctal atctgtatct tcttgggga gacgtctgt ttcttaact 1260
ggccatcac tgcagacatc ctcaagtac tggctcctcc caccgggggc gccactgccc 1320
tggccttcca gacttacc tcccactgc tgggggacgc cgggagcccc taactcatt 1380
gctttatctc agacctgac cgcacagaca ctcaaggact cccgctctgg agttctctga 1440
ccttgggcta cggcctcatg ctctgcccct tctgtctggt cctggcggcc atgtctctc 1500
tgcacactgc gctctctctc gtcaagacac ggcgcagggc tgaagcagag gtgaaccagc 1560
tggcagctgc gcccgcatct gtaaagtct gagggtgtgc catgtggaca atgaagaacc 1620
cacactccca cctctctctg gagggtctct acagcttccy ggaacggcty ggtgtcccc 1680
aaactttctg tggatccac ggcacagac ccaactctc tggccagggc ctctgtgagt 1740
gccctggcal caggaggagg ctgtctctc agtaccctg gaagatgtg tgggtggag 1800
caaacagtt ggcacagttc ccagcctag gtttggggc caggyccctt gggcccaag 1860
aaagagcaag ccccagctg gtttccgggg agagcctggc ctgccaccag ctatgtgat 1920
cttgggcaag tccctgccc ccttggaaag aagggccagg gggctggact tcccaaca 1980
acttctggg caaagacaga tctgagctt tgaagactca acagaccctg gaccatacgg 2040
agagcaggtg gccacggcct cagggggca gtcccggctt tgaggctac gggaggcct 2100
gglatcagg gaccactgct cagctgggac tccgaccctg gggatattg agcaaacctg 2160
gaaatgag ctggggccc aagttcttgg gtaactctg gaggacatg tctcactgt 2220
tgggttggc tccagcctg gaggctccag atggggacty ttctgcaag ctggcctac 2280
cagggttgaa ggcctggtt gcaactgtac adcaactgtg ccccaagctt caaggtctct 2340
ggcaggtgca caaccgcca actctgcaag gctctctcc ctgccaccac ccccaaguc 2400
agggcccccac tcttccccg aggtgagct gaccttttc cagggcagg gcccaagaa 2460
cacttccag aatccatggg gaagttagca gggctccggg tgcctgagga agcagctatc 2520
cacaaagctt cctgcccag agctgaggct gaggcccggg gaggggcggc cctaacccaa 2580
accttggctg ctggcattcc ccaaggtgac cccagggccc aggcctctga tcaaccacct 2640
cccttccatg caccaccag gctgagctg ccaacttcc accagccca accccttgg 2700
ggggagctta gccccctgg tcacccactg cctgcacttc tgcctcaatc aaggtgttbc 2760
tggcggggg ttggggggg ggggtgaggc ttgtggccaa tgggggacc cccaaagacc 2820
agctgggaca atgtctctcl tgcctcttag ttaactgctg gctgtgctt cagtgtgtg 2880
taagcaggtg gaatactcac caccanaact ctgggttacc ccgagggcct gacaagggga 2940

```

WO 01/46258

PCT/US00/35095

tggggggggg	gtggcctcct	ccaaagacca	gctccacccc	ccactccagc	ctcagcgggg	3000
ccccagcat	gtttcttctg	tgtacacgaa	ccaggtccga	gtgtgtccct	ctcttccttc	3060
cygaaaccaa	actgctcctt	tattttttag	agctgctgat	tgtgaaatcc	agagtcataa	3120
gagggagcc	aaatbatatc	ctctttgtaa	tgaagaaata	aacctattta	aatcacaaa	3179

<210> 37
 <211> 1986
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3358383CB1

<400> 37	ggagatcttg	agcaaaat	ttcttacgtg	acbttagaga	aanccgctac	ctatctgacc	66
	gcaaaacgac	tgagggaac	gtttccacg	gtccctgctc	aggggggaag	acagctgtc	120
	aaagaagcag	tggggctagg	atcaaacac	alltagtqlg	acttagggaa	sgaaaaactt	180
	ttccctcttt	gaacctctct	ggatacagtc	atlttgctct	taactggaga	tcacctgttc	240
	aaactcaatg	gaccttcagg	acctctctgg	tcacgctggt	gacctgtgga	gattccagat	300
	ctctcagact	gtttttctct	caatctttgc	tglttgctaca	taacctcatt	ttatgctgga	360
	gaacttcaact	gcattcaatc	ctggccatcg	ctgtctggtc	caactctggy	acaatgacac	420
	tgtctctgac	aatgcaactg	gggacctcag	ccaagatgca	ctcttgaaaa	ctccatctcc	480
	actggaacta	aaatgagcgc	caggaagctg	ctgtctcttt	gttccctctc	agtggaagct	540
	ctctccactg	aatgggaact	tcaccaaac	aagtgcagca	gcactggacc	octgtgtgga	600
	tggctctggtg	tatgacaaga	tctcctctct	atccaccatc	gtgactgagt	gggatctggt	660
	atgtgactct	caatcaactg	cttcagtggc	taaatctgca	ttcattgctg	gaatgatggt	720
	ggggagcacc	ctaggcggtc	atctatcaga	caggtttggg	agaaggttcg	tgtctcagatg	780
	gtgttaactc	caggttgcca	ttgttggcac	ctgtgcagcc	ttggctccca	ecttctctat	840
	ttactgtctca	ctagctctct	tgctgggat	tgctgcaatg	agcctcctaa	caaatactat	900
	tatataaaba	gcccagctgg	caacacacag	atcccaggcc	atgggaatta	cattgggaat	960
	gtgcctctct	ggatctgcaat	ttatgacctt	ggcagcctg	gcttttgcca	ttcgagcctg	1020
	gcatactctc	cagctggctg	tgctctgacc	atactttgtg	atcttcttga	cttcaagttg	1080
	ctgtctagag	tctgctcgtt	gctctattat	caaccaataa	ccaggggaag	gcttaaaagg	1140
	acttagaaaa	gctgcaacaa	ggagctgcaat	gaagaaatgc	agagacaccc	taacctctga	1200
	gattttgaaa	tcaccatgca	aaaaagaact	ggagcagcca	caaaaaaaaa	aacctctctc	1260
	gtgtgaaatg	ctccacatgc	ccaacatag	taaaaggatc	tcctctctgt	cttttagcag	1320
	atttgcaaac	tttatggcct	atlttggcct	taatctccat	gtccagctcc	tggggaccaa	1380
	ttgtttctcg	tggcagctc	tcctttgggc	agctaccctc	ctggcaactc	gtgttgcaac	1440
	tgggcaactg	aaatcaatga	ccctctgagc	aaagccagatg	ctgtctctatg	acctatggc	1500
	aatctgcttt	atggccatca	tattttgtcc	ccaagaatg	cagacgctgc	gtyaggtttt	1560
	ggcaaacatg	gctctaggag	cgctggctct	gaccaatacc	cttgettttg	ccaatggaaa	1620
	tatagcaatt	ccaccataaa	tcagggcaag	agctatgggg	atcaatgcaa	octttgctaa	1680
	tatagcagga	gacctggctc	ccctcatgat	gatctcaagt	gtgtattctc	cacctctgcc	1740
	ctggatcctc	tatggagctc	tcctctctat	ctctggcttt	gotttctctc	tccttctgca	1800
	aaccagyaac	aagcctctgt	ttgacacct	ccaggtatgag	aaaaatgaga	gaaaagccc	1860
	cagagaccaa	aagcagaggt	atccggaggt	ggaggtgagc	cagttttaa	gaatccaggt	1920
	agctgactgc	cgatcaatga	gccagatgaa	gggaacaatc	agyaactctc	ctagacacta	1980
	gcaaat						1986

<210> 38
 <211> 3294
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4250091CB1

<400> 38	tgtaaagacag	gaaagggatc	tatttgatgt	ctatcttcaag	atatattggc	agttttctct	60
	agctatbttt	gtctctcctc	tgctgctctt	ctatctttga	tactgcaagt	tcctgggcaa	120
	ctcgaatttg	caaacacagc	cctggataca	ctatttacc	tacagtagtt	tcctgggcaat	186
	ctcaagctgg	ttttttttat	tcctctctcc	ctctccctgc	ataatcaagt	ataactggca	240
	ccatctatgg	ttatctggtt	atctctcaag	tgtggctgat	gctgacctct	agtttgaagt	300
	gagggagaaa	tgatgagcca	ggaaactggc	actttgaatg	tggggaggaa	gatattccag	360
	acaaggtttc	ctacgataaa	gcagttctct	gcttctctgt	tggcagcaal	gttagatggc	420

WO 01/46258

PCT/US00/35095

```

agagaccag aatcoasgat ggttggtggc cagatttttg tagaacagaga tgggtatttg 480
tttagttca tcttagattt tttagaact caccagcttt lallaccac: tgaattttca 540
gactactta ggttcagag agaggtcttt ttctatgac ttgtttotc agtbgatcc 600
tbaaacccat acctgctaca gccaaagact gctttgttg aggtacattt cctaagccgg 660
aacactcaag cttttttcag ggtggttggc ttctgocga aaacaattga gatgcaaca 720
gggaggatta cagtggttac agaacaacct lcgggcgcga cctggaatgg taacttttc 780
cctctcaga tgaacttaet tcaactggct ccacaagac ctcttacc: tpaacbggtt 840
tccaptgat gctctgacag cactactgat accaaaatg gactagata ttctgtac 900
tgcagtatt ctctgtata ccagtttct atgtttctc taaaacttg aagttoctca 960
ggcctgaac ttctggaaaa gatgattatt caaanatag tttggggt: acoagtgag 1020
ttggtagaa tgcocaaata attattttcc aaactggat acttttaga gtaaaaggg 1080
ctattatlag gtgggacaaa aggaataaat gaagactgc cagaaaaaac tgagactatg 1140
gacattcaaa toatgggaga aataaatttt gtgattatg tctcattgt aatgaattg 1200
acttgaaaa gaattggcct abttttaaga gattggttca gttggtca: taagggctcg 1260
ctcaclggt tctctgagt tcoltacaca ctatataag tttcttca: gtttatgat 1320
tcaactaotg tttttctct agctgacttt attttaaac aacctaaag aagatatt 1380
ctcatggca abttggtatc ctgttacagc abtgctctt aaacaactc aatatctgg 1440
ataggtgtc agtatgttaa ggaatgttc tctctagtc attttcact taectoccl 1500
gtgttcttg cctggatcct aacgtgatt tcaactcgc tgtcaaaa: atttttccc 1560
cgttaaatg tcttaatgct gctctaccat tattttacca actgtgaag ctgggttaa 1620
tttttagag gaaaagaaa gcctgcatgt gttctttatt ggtatcattt aatatatac 1680
ttttttttt ttttgtaaa ggtaggccta ttttaagata ttctttaa: ttgagcagta 1740
gcaaacagga agtatccag tgtctctctc ccttagcac acaactctg ccttgctta 1800
ccaacggag gacccaggt gaataaacg agtatgaca ttcttaabtg tttatattt 1860
ataactctc gtcttcaaa aagatttga atgtcattt uggaaaaag agccagtca 1920
gctataggc gattgtgaa gaaaactca taactatct ttatctcaa cctctgtca 1980
acttatttt cattgatgg atactttaac aaaaatgaa tttttttgg tttttaaat 2040
atgagtgatt atgacctct tggggatcat gctcaaaaa gtcagaaac tagagacaa 2100
actgtcaatt abtttaaga agaaacacac taggtcaaaa gaagatgtc tggaaatac 2160
aagtactctt taaaaacct gcatttggag naagtattg ttctctgan aaacatgat 2220
aaaaactaaa actggagct tctgtgtgt acacagtc aaatggttt cctctttat 2280
gtgtgtttt agaaaagca gaaaagtct ttcaattta aagtagaaa acctatatt 2340
tagacttca caattccaa atcagagct abtttaaaa ttagcattt ctgtatcac 2400
caatagtat tcaattgtt gaagctcaaa atttttcca tccataaal gtttgyaac 2460
ttttgacag tgcocattta aaagttaga tagcaaatc gaatacgtg aattatggg 2520
gatctctgtt gattgggatg aaaactctg ccttaaaag tcaacttta gtataaatt 2580
gcctaattag caatcattt tttttttgc tcaactcctg yctgaatc atctgtctat 2640
tcagatatt ttgttaggt ttgaaaaat gggagtgag cctaattggt gcctaactg 2700
ctgtgtatc atcacttta tcaagtgtg tcaataata agattatcc ctcaagtta 2760
acttagcag ttgctcagtt abtatctc saggtcacag tactagaaat acttggctg 2820
catcttcaag atgocattca ttttatcaag ctcaaatat agttggtca aggtattca 2880
agttttatt tgaactctc tttttgaact ggtcfaatg gaaaagtga gttgcttta 2940
aatgttaaaa atagtttaa actttatatt tccattggt tccactatt ttgctctc 3000
ttgtgtgct tgaatattt tabttttcag ttgtctca taggaaatc agtattttag 3060
ctagggatg tcttcaagt acttccact ttgtacaat clactatct tatatactat 3120
ttgacttta atcttttat gagatgtct gtaacattt tctcacttg acaaatgct 3180
ttgactgta cagtcaaat ctggctctg gggtagatg gaatgattg ctatattga 3240
gaactctgtt tatcaaacat aataaacttt ttttagatg tgaaaaaaa aaaa 3294

```

```

<210> 33
<211> 2043
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Inocyte ID No: 70064803CB1

```

```

<400> 39
gcaacatggc ggtctgccgtg gtgcagcgc cgggctgagc gacagcaagt ccagcgggt 60
cctacccggg gtgaggggtg gcctccgggt gggatcgtgc cctcttcagc ccgctccctg 120
cccgaactc acgtgtatto ogaactccc ctccggctgc tgtatctact gaggcggga 180
ggctgacag ggcctgggctc cctctccagt ggtctctgt gcttccgggc aagctcccc 240
tctccggcg cactctctc gccgtgtct ccttctctc ccttctctc gctctctcc 300
ccgcagggtg gactctctc tgggacccga ggcgggggg ggcggggccc ccggcccat 360
ggcgggtcc gacacccgc cctctccag ccaggcgtat gacccggag ccggcccat 420
gctggcacc cgggggttc nagggtccac ggggaaccc agtcccgag 480

```

WO 01/46258

PCT/US00/35095

```

cccgaccag gaggggctgc agcgcactac oggcoctgtc ccggccgctt cggctctcat 540
agtggggggg ctgtctcaaca tcaatctcct gaactacatg gaaccyctca ccgtgctggc 600
cgtccctccc gcaatcgagc agttcttcaa catcggggac agtagctctg ggctcatcca 560
gaaccctgtc atctccagtt acatgggtgt ggcacctgtg tttggctacc tgggtgacag 720
gtacaactcg aagtaactca tgtcgggggg caatgctctc tggctcctgg tgcactggg 780
gtcactcttc atcccggag agcatttctg gcctgctctc ctgcccggg gccctggtgg 840
ggctggggag gccagtattt caacactcgc gccactctc attgcccagc tcttctggc 900
cgaccaggg agcggatgc tccagctert ctaclttgoc attccgtag ccagtggtc 960
gggtacatt gcaagctcca aaglyaaaga latggctgga gactggactt gggctctgag 1020
ggtagaaccc ggtctaggag tggtyccgtt ttctgtctg tctctgtag tggggagcc 1080
gccaaaggga gccgtggag gccactcaga ttggccacc ctgacccca cctctggtg 1140
ggcagactg agggctctgg caagaaactt catcttggg ctcaacca ctctgacgg 1200
agtccctgtt gtaggctgg gtaggggat cagccggcgt cctcccaact ccaaccggc 1260
ggctgatccc ctggctctgt ccactggcct cctgggtctt gccactctc tcttctctg 1320
ccttgccctg gccctgggta gcatctggc cacltatatt ttcatctca ttggagagac 1380
ctctctgtcc atgaaactgg caactgtgg cgaactctg ctgtactgg tgatccctac 1440
ccgacctcc acggccggg ccttccagat cgtctgtcc caactgtgg gtagctctg 1500
gagccctac ctactggcc tgatctctga ccgctctgc ggaactggc cccctctct 1560
ctgtccagc ttccggctc tgcagtctc gctcaatgc tgcgctctg ttggggact 1620
gggctggcca gccctctctg gccacggcat ctcaatgag gccacggcc gggggcaaa 1680
gctgcactg caaggctctc tgcacgaagc aggtctccca gacgaccga ttctgtgcc 1740
ccagcgggag cgtccaccac ccgtgcccgt ggcagctgt ctactctgag agcctgccc 1800
tccactact gccactctgc cacagctggc ctlyggccca ccccacgag ggctggggc 1860
taaacctctg gctcggcca gcttccagag ggcactggc ccgtgtgca gctccagac 1920
actacatgg tgcctcagc aagaggtgg ggtccagga ggggatacc tctccagc 1980
ggcagcccca agggctcggc gctattgtg acggatcaa atttgtagc agacaaaaa 2040
aaa 2043

```

```

<210> 40
<211> 1915
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> Incyte ID No: 70356768ca1

```

```

<400> 40
caactcgg cgtcggcgc tgccttccc tccggccaca ggcctcggc taccgctt 60
ctaattggc cggggtctc ctgggacag caagactcc gctcagccc ctcttccaa 120
tgctccagc cctctgoga tctagatga ttcagggcag gatctctct ctgacactc 180
ggctcggc gattggtgg actttcaqt ttggctata cctctctac atcaatgcc 240
gcacttca cattcaggaa tccaccaat agacatgga ggcgctact gtagagccc 300
tgccgatca cctagcctg ctatgtgtt cctcaatgt gctctgtat cccctggag 360
gctctttgg agcactgctt gcaggtccct tggccatcc gctgggaagc aagaatccc 420
tccgtttaa taactcttt gttgtctag cagcaatcc gtttggattt agccgcaag 480
caggtcctt tgcagcgtc atgctggaa caatgctgt ggtgctaat gcaagttga 540
gcaatgaat ccaagccatg taactgggg agagccccc taaggagctc ctaggagct 600
tgccatgag ctcagccatc ttaccggctc tgggatctt gatggacag gtagctggc 660
tccagagct cctaggtggc cctcagcct ggcctctct gctggccagc tgcctgtgc 720
ccggggcct ccaagctccc tccctgcctc tgcctcctga aagcccgcc tccctctca 780
ttagctgg agacccggc gccctgctg cagagacgg tctctctgt tccagctgt 840
agttctgtt ctgttctat gcaatgccc atltlgate agcaggaag tttctctt 900
ttttgtttt gttttttgg tttttttgg gacgggtct caactctgt cccaggtgg 960
agttgttga tcttggctg ctgcagctc caactccag gcccaatgg tctctctg 1020
tcaactcct ggtggctgg caactctgg caactccag acgtctgtt aattttttt 1080
tattatttga tttttttaa agatggagtt taactcttt gccctggcag gttctcaact 1140
ctcgagatca aatgatctc cccctctggc ctcccaagt gctgggatta taggcatgag 1200
caatgtate tggtagaal gggagtttt aactgtccc ttcccaact gggccagtg 1260
atctcctct agcagcctg gtagtctgt ctctggggt ctcaactat tcatgctga 1320
cttagtgag caactgact tgcctagctt actgcaacc agactctgt gggccaggg 1380
atctctctt ctcagcctcc ttagttagctt ggaactctg caacacccc tatgctggc 1440
tctccatgt tctgggtct accctctgg atgttttcc ttctcttca cctctctga 1500
tctctctga agaggggtt gcaaatctg ctgctttga tggtlgaga aatctctag 1560
ctctctctt gctcttagag agttggggca ggtggggcc cagctcacc ctgtaatcc 1620
agggagctg agggggcag atcaagaggt caggaatca agacccgct ggcctcact 1680
gtgggaccc actctacta caatacaaa aatggctgtt gtaggtgtg cctgctgt 1740

```

WO 01/46258

PCT/US00/35095

```

gggtccggct cctgggggagc ntgaggcggg agagttgctt gggcccggga gggggaggtt 1800
gcagttggcg gagaattgct tggggcccgg gaggcggagc ttgcggtagc ccgagattgt 1860
gccagtgcaac actgcactcc agcctgggtga nagagtgaga ctccgtcttc aaaaa 1925

```

```

<210> 41
<211> 1809
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5674114CB1

```

```

<400> 41
atgggctctg ccagggcctt acgccccttc agcggcggcc tggattcggg agacagccgg 60
ggggggcttg aagggggggc cggggcccgg ctgtgcggca ccgggtgggg gcccggaccg 120
gtgcagtcac ccgtgggccc ggccccgctt cgtttctgca aaaaagacyg gcactgcaac 180
gtgctgttcc taaaacctgg tggccagggc yggcctacc tgagcgactt gttccaccac 240
tggctggagc tggctggggc ctgggatgtc ctgctctct cctgctcctt cctcgcctcc 300
tggctgctct tggcctgggc cttctggctc atlgcctcgc tgcacggcga cctggccggc 360
ccgcaccgcg ccgcgccctg cttctccacc gtggccagct tccctggcgc ctctctcttc 420
ggcctggaga cgcagacgtc catcggctac ggcgtggcga ggcctaccga gtagtgcgcc 480
ccgcctgtgg ccgcctgggt gctgcagtcg atlgccagct cctgtctcga cgcctctgct 540
gtgggtgctg tcatggccaa gatggccaaa cccaagaagc gcaacggagc gctggctctc 600
agcgagaacg ccctctgtgc gctgcggcac caccgcctct gctcctatgt gcgctggggc 660
aacctgcgcc cgaqccacct ggtcggagcc caactgcctt ccagctgctt gcagcccctg 720
gtaccccagc aggtgtgata cctcccctgt gaccaccagg atgtggatgt gggctttgat 780
ggaggcaccg atcgtatctt cctcgtgtcc cccatccaca tgcctcatga gatcgactct 840
gccagtcctc tgtatgagct agaacgtgcc gagctggcca gggctgaact tgagctggtg 900
gtcaatctcy augggalggt tgaagccaca gtcctgacca cccagctgct ctgctctac 960
ctccctgggt aactgctctg gggccatcgt ttggagccag ctctgtctca gctgggtctc 1020
cagtatgaga tggatctcc ccacttccat cgcacttatg aggtcccagg gacaccgtc 1080
tgcagtgcta aggagctgga tgaacgggca ggcagcctt cccacagcct caagcttagt 1140
ttcccggctt ctctgactgc atttgtttat gagaatgaac ttgctctgag ctgctgcag 1200
gaggaaagct agaacgatga gactgaggaa gggaaatggg tggaaacaga agatggggct 1260
gclagccccc gattctctcc accaaccttg ccctgacc ccgctccatg atgcaaac 1320
atgtcccctt cccctgttat gccccttcc ccaaggtgac agatggagg gatggggctc 1380
tctctgggaa tggggccagg tgttccctga taccgacagg cctgctgggt aatgactag 1440
gtggttaagt ttgctcctgc ctggtgaccc accatggaca tactggactt taactctct 1500
gcttctgctc tccctcctga gaaccttcta tgagctgtat tccctagctt caccgaatt 1560
ctgactcacc caaagggaaa agactggcag tctlagatto ctctatatg gggacctctg 1620
atgtttgacc aggttgagaa gccaatggtg tagactgctt ctggggaaag aagttggcag 1680
tctctgaaac gcatcagata tcaagagttt gtaggtctgg atccactaa gatccaaggg 1740
agtgttctct ccaactcag ccaactgagt aaccaatcct ttgttttaga ccactcaagg 1800
agggaaaggt 1809

```

```

<210> 42
<211> 1730
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1254635C31

```

```

<400> 42
ctttggccta ttataccatg gatgctaaaa atggttctaa ctgaaaacc ccaaccagaa 60
atagcaacaa gctctagaatt ctactactca caaaactcac ctggatccct aagggccacg 120
caaaagatga gctattacgg cagcagctat cataattaca atggggccgc aaaaaccaca 180
ggatcccgcg cagagcaaat tatagctgag aagagagagc caagaagagc attacttcc 240
aaagatggca cctgtaattg ctactcaag nacatttttg gagaatgggg aagctatgct 300
gtgacatct tcaaccctct tgtggacacc aactggcgc atctgtttg gatatttct 360
ttactctata ttctctcctg gttgatctt ggtctctgt ttggctat agccttctc 420
cgtggcgatc tattaaatga tcaagacatc acactctgtg ttgcaaacgt ccattcttct 480
ccngggcctt tttgttctc cctagagacc caaacccca ttggatctg ttakcctct 540
gtactgaagc aatgttctgt ggcctgctcc atggtgacc tccagctcat cttaagttgc 600
atcctaataa cttttatcat tggagctgcc ttggccaaaa tggcaactgc tcaagagaga 660

```

WO 01/46258

PCT/US00/35095

```

gccccaaaca ttgctttcag ctaactttgca cttataggta tgagagatgg gaagctttgc 720
ctccatgtggc gcaatgggtga ttttcggcca aaccacgtgg tagaaggaac agtttagacc 780
caactctccc gctatacaga agacagtgaa gggaggatga cgatggcatt taagacccc 840
aaatagtgca acgacaaaat catcctggtc acccggtaa ctattgtcca tgaattgaa 900
catggagccc ctctgtatgc ccltgaccgc aaagcagtag ccaagataa ctllgagatt 960
ttggtgacat ttatctatac tgggtattcc actggaacat ctaccacatc tagaagctcc 1020
tatgtcccc gagaatctct ctggggccat aggttaaatg atgtcttggg agttaaagg 1080
aagtattaca aagtgaaclg cttaacgttt caagyaatg tggagataa tgcctccttt 1140
tgcagtgaca agcaattgga ctggaaagac cagcagctcc acatgaaaa agcaccacca 1200
gttcggaaat cctgcacgtc ggaacacaag cggagacgaa ggtcaattag tgaattgoc 1260
atgttcagca gctgtgaaa cctcaggag accaccactt ccgacacaca tgaatttag 1320
gaaacacctt atcagaagc tctcctgact ttaaacgaa tctctgtaga atcccaaatg 1380
tagtctaaa ttgcaattat gaggyctacc actgaaatcat tttatctttc agccaatcaa 1440
gtctgtgaaa acgtggcttt ttgaaagtgt ttatggctat gttttatgat gatgctgggt 1500
aagtagagta agttaaactt ggtaaagat aacttaaaa tccatagtt ctcaattatt 1560
aaatttttc ttgtttgca atttgttatt aagaaatgta ttaagctaa ctgattaaa 1620
ttlatcttt ttattatctt acatgctgtt atctcagtt ggaggtgtag tattcaaaa 1680
cggggatga aggcgggaag gaggctggaa taataaaaa taatagatt 1730

```

```

<210> 43
<211> 1147
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1670595CB1

```

```

<400> 43
gcaactctct tttccggccc ccgtgcaact tccgcccag gggagcccc cggctcggg 60
ggatcgcgcc cagcgcctgc gctctgccc tgggtcact aaccatttg tggctctct 120
taactctgct cagcctatgc cagcggagag tcaactctgc tggcctacc gctcctggg 180
gaggaagggg ttgcccctcc tgccaatgg cccgggggtc ctggaggggc gctcgcggg 240
aagctctcca ccttccctgg tgtgctgtg cccactgtcc tgtccatgtt cagcatagtt 300
gtttttctga ggattgggtt cgtgtgggtt cactgtgggc taactgaggt cctggccatg 360
ctgctgtttg cctactctat cctggcaact accgtctctc ctgtctgtgc cctgcccacc 420
aatggagccc tccagggggg cggagcctac tgtatctcc aacatcgtg gactgggatg 480
ccaagggccc cagtgggctc cgggtctctc ccaagggcta cggcttggaa cctgtctgat 540
ggctccctgc tgetgggccc ttgggtggg gctctgcaact tggagccgg cctctatgac 600
cgggcctcat tctcaaatl cctgtgggic tclggctccc tggctctctg gctcctcagt 660
tttctggclg tgggcccag ggaactccgc ttgactccta ggctgcccc caatggctcc 720
tccctcgcgc cccgcttlyg ccaactcacc ggottcaaca ccagtlacct gaaggcaaac 780
ttggcgcctg gctatgctga ggaatcacc accggagccg tgaatgaatt tgcagcgtc 840
tttctgtccc tctttaaagg caggaatcat ggtggggccc aacatgtcag gggagctgaa 900
ggacccccag cgggcgatcc ctctgggccc galcgtgccc gtcgctaca ccttctctgt 960
ctatgcccctg ttttctcttc tctcagcctt ccttcactg gtccttgat gctaggggcc 1020
agggctcttc tgtgaactcg ggttaactca gtttcccact ttggcccga ctcaaggcc 1080
caccgggctg gtgatgtttt cgtctcgttt tarittttcta actctgcatg accatgaa 1140
aaagacc 1147

```

```

<210> 44
<211> 2745
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1859560CB1

```

```

<400> 44
cggcgaagcc agggacccca cgcactccga gtgaagcaac tagaacctca gggctgtgaa 60
ggccacaggl aggggctag cggagggtgg cctcaggagc gggagcccc caactctccc 120
tcagagccc cagtccaccg tagcgggtgg agcccgcctt ggtgctcagt tgyaaaacct 180
cgggcccgg ctggatctcc tggctgccc acccaacccc cggagccta cggcccacg 240
tagagatgcc tctctcgttg accggctgg gtcaggccag gtcctctggc cccgggatgg 300
ccccggagcc ctgctctgc tccctcggg cctgcagag gaggtgccc atctcagct 360
ggctgcccag ctactccctg caatggctga agatggattt cgtcggccc ctctcagttg 420

```

WO 01/46258

PCT/US00/35095

```

gcctcaclgc cattcccag ggcgtggcct atgctgaagt ggcctgactc ccgcccagt 480
atggcctcta ctctgccttc atgggctgct tctgtatatt ctctcctggc acctcccgg 540
atgtaactct gggcccacc gccattatgt cctcctggg ctctctctac accttccatg 600
agcccgccta cgtctgtctg ctggccttcc tctccggctg cctccagctg gccatggggg 660
ctctcgtttt ggggttctct ctggacttca ttctctacc cgtatataa ggcttcaact 720
ctctctctgc cgtcaccatc ggtcttggac agatcaagaa cctgctggga ctacagaaca 780
tcccaggcc gttcttctct caggttlaec acacttctct ccggatgca gggcccagg 840
taggtcagcc cgtctcgggg ctggtctgca tctgtctctc gctgtgctg aaactgatgc 900
gggaccactg cctcctcctc caccctcaga tgcctcctgg tctgctggct agccctgggc 960
tgtctctggc tgcacagaca cctctcaacg cctctgggtt ctctctcaga gccctgtgtg 1020
gctactcctt caggttgact ggtataccag ctttcatctt aacaggggag acagctgagg 1080
ggctccctcc agtcaggatc ccgcccctct cagtgaacaa agcccacggg acgatctctc 1140
tcaccagat ggtcaggac atgggagcct ggtctggcct ggtgcccctg atgggctccc 1200
tggagagcat tggctggccc aaagctctcg catctcggaa taattaccgc atcgtatgca 1260
accagagct cctggccatc ggtctcaaca acatgttggg ctccctctgc tctctctacc 1320
cgttcacagg cagcttctga cggaccggcc tgaacgtcca gttgggggtg tgcaccctgg 1380
cgggtggctt ggtgagggga tctctgtctc tctctctctc gactactcct acctcaactg 1440
tctactcaat ccccagctct gccctggctg ccgtctcatc calggccctg gccccctctg 1500
tcacaccaa gatctcagg agctctggc gtttaagag cctggacctg ctgccctctg 1560
cgttgacctt cctctctgct tctctggagg tgcagtccgg cctcctggcc ggggcccctg 1620
tctctctctc atgtctctg caactctgag ccaggcctga gaccacgtg tcaaggggc 1680
cggctctctt cctgcagccc gccagggccc tctctctccc tgcctatggg gctctctggg 1740
aggatctct aagccggccc ctggaagctg ccccgcccag ctgctggctc ctggagtcca 1800
cagctctctg cagctctgac tcaactctgg tctctggact cggcagctc ctccagact 1860
tcacagagca gggctctccc ctggcctctt tgggctctca gttccccttt ctccctctcc 1920
tctctctctg tgaactgag gggctcact actctctctc cctggaaaca cagagagagc 1980
acctgagcca ggcaccagg accagcctct acaacatcag agaagactcc atctctggac 2040
aaagctctg cctctctcag ccataatggg gccaccctg ggcctcaca gtttgcagg 2100
tctctctgaa ggtctctctc actctgattg gatctctgac gccctctgat agcatctct 2160
gctctctctg caaacctctg agcagctaac ccagggaaga gaaagagccc aggcctctgg 2220
gtccacgca ctggagctgg ggtctctctg ctctctctgg galyaclyga aaatgactc 2280
gctctctctc cctggctgaa cctctctctg aagagctgtt tggagagctc ctctctgact 2340
gacagactgt cggaggaagc aggggcaagg gtttccagcc cggctctctc gaagctctcc 2400
gggctctgca gccctctccc ggtctcaccg tgcaccctgc gggggagggc cggggcaggc 2460
aggagctctg gaggcggctc cgtctctctt gctctgggca ctctgctctc ccgagagaaa 2520
accagctgt gtaaatgac gtaactctc tatttataaa taattctgt tttctaaat 2580
ggaagagct atgctcttgg tgattctgta aaagtctaaa atgcttattg taaaaaatc 2640
agaaaccac cctcaccctt gttcaactct gtagatctc cagaccctcc cccaaactg 2700
catactgacc tctcctcag tctctggatg taatttaca gttct 2760

```

```

<210> 45
<211> 3204
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5530164Cbl

```

```

<400> 45
cgcctctctg agctactcgg cctgcaagcc cagcctctct ggcgcaggg ctgscggccc 60
agctggcctc gcccaaatcc gggggcgggac ccggggcggg caacgggccc gtcctcgggg 120
aggcctctgag ccgcaacagcc ccnccagggg tggctgctgt aaacggggct ctggatcccc 180
gaatggtctg gtttctcctg gtttgggtcc gggggaggcc caccgaagcc agcgaacccg 240
ctgacacaa ctgcccacta tgaactcctc agggcctgga agaccgcaac ccgcaacaag 300
cggccctcgc actcggcagc gactgagatc ategcctccc ggtctgtagt cggctcaacc 360
gccactctg actcagaggg aggggctcac gttgagatct ctgagagagg cgtactctg 420
cactaggaaa gccactcccc ctttccaaaa atgctccgga agtgcctctg cctlccgtaa 480
agatggccgg ggcagctggc ccaggggagg cggggtatgg cctgcccacc aagttcggc 540
gggaagatgg cggatgacaa ggtattctct cctaaactta aggaactgga atttctcaag 600
aacagctgg aaagcctgca gggcctgtta gaagaagaa tcaacagctg agtgggcccag 660
gaatgctcgc gtttctctc cccgtctctc aagggatccc tggctgggta tctgctgccc 720
aaactgagg cctcagcagt atgggcttt gctgtgggca cctgcaactg cctctatggc 780
gctcagctal atgctgtccc caactctgag aagcacttaa ggaactatt gcaattgcta 840
cgcaaggggc ccactctct ctgggtgcca tggaaagccc aggatgagca gctcagcctt 900
caggtggaga cacttctct ggatctccca cgtctctccc atctctctc tccactctc 960
ctgcccactt cactctgccc tccctctctg cagatctgag acagtagtct cctcagctgc 1020

```

WO 01/46258

PCT/US00/35095

```

acctggatt ccttctccc cttoctagct ceatgggact egccccaaqa ctgtggcttc 1080
aaggaccacc agccccttac tcttcaagcc ctgactgtgg agtltggtaga tgcctctgat 1140
ctctcagatt ctctctggca atgttcccag gcttctccct ctggggagct ggctccctaa 1200
cttgattttc cccaaacgtg ttgcaatccc tgcctccctc tagcccccga ggtctctgtg 1260
tgggtatgag tgtagaggat gggggatgfc cagggctggg ccgtcccagg cagggccgct 1320
ggccctctat gctacttcta tccactgcca tglacggltgc ceatgcccga tgcctggcac 1380
tgtgcaatgt ggaagggcga gtgcccctcc gggcccctcc agccgctctg ctgactgagc 1440
tgaaccaagt tctgttatgc gcttctccc tctctgttag ctggcaagca tggcccaggy 1500
ggccccacc ctggcccag qctgctccct tccactata agccctgctc tatggcctca 1560
acaacaacct ggtgatctat cttaagcgtt acatggaaoc cagcaactac caggtgtcga 1620
gtaactcaca gattggaagc acagctgtgc tctactgctc ctgactcggc cacocctct 1680
ctgtgctca ggggttagcg ctgctgctgc tgatggctgc gggagctgca tatgcaagc 1740
ggggcctta agtcccggg aacacccctc ccagtccccc tccagcaqct gctgcccacc 1800
ccatgcccct gctatccct ccgttaggcc tgcctgctct catctctgac tgcctcactc 1860
caggtctgtc gtcaggtac acagactgct tcatgagcgc acagggctg cctctggcac 1920
tccagacct ctctcttac actctgtgtg tctctgaa tctaggctct catgctggc 1980
goggtcttg cccagycctc ctggaggtt tctcaggalg ggcagcactc ggggtctga 2040
gccagcaact aatggactg ctctgtctg ctgtcatgaa gcatgycagc agctcacc 2100
gcctcttct ggttctctg tctgtgtgtg tcaacgcctg gctctcagca gctctgctac 2160
ggctgagct cacagccgcc tctctctgg ccacatctct catggtcctg gccatgccc 2220
tctactatgg cagccctgag tccctgacaa ctctccact gattccggac cctgtagatt 2280
ggggccacc ccagatccc cctcccagc ctctctccct ctcccactag cagccctgta 2340
acaagtgct tgtgcaaaa cctggagag cctggagcgc caggttattc tctggagatt 2400
gggtgatgaa ggggtaccoc taggagatgt gaggtgtggg ttggttaag gaagtctta 2460
ccalcccaca ccccacaca agtctctca gactaaagaa ttaaggtaac atcaatcct 2520
agggctgaga aataaccca tcttgttgg ccagctccct gcttctctc ccatgaaagc 2580
agttgatgaa agtgggtgt gggcaacaag tggcttctc tgcctacttl agtccaccag 2640
cagagccact ggagctggg agtccagccc agccatggtg catgactctt ccatnaggg 2700
tcccaacct tccacttca tgcagaagg cccagtggcc acagattata caaccattc 2760
ccaaacacct ctgacgtct cctccagtc cagcaatgcc tagagacatg ctcccctgcc 2820
tctccacgt gctgctccc acaactagcc ttgtctctg aaaccccga gaggctggg 2880
ctgactcat ctgaggaat gtgcccctg gcccctggtc taagccaca ctctgact 2940
ctctgttcc cctgaggct gctctgagc ccgtaccca ctctgagct cctaggaggt 3000
accatcttc ccatctggg gactgcccct gcatagcagt ctcccagctc ccaacagct 3060
ggggagctc tgcacagat gactgagac caggtacagc aaacctgtag ctcaatcagt 3120
gtctcttaa ctgcaaaag aataagatct taataagtc tctaggctg taggtgtgt 3180
cotacaacca cagccaaaa aaaa 3204

```

```

<210> 46
<211> 2763
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> Incyte ID No: 139115C61

```

```

<400> 46
tgcatttct atgacllga ccgttccact gacaaogcaa tatgtttatc ggagaaatg 60
ggaagaaact ggaactaca ctcttccatc tgalagcaat attctctgct gtaaaaaaa 120
caaaagcagc ccaatttttg cctccaggca ggaagtccag aaaaaagtyt cagctttcaa 180
ctctcagatg gacaaagty gattaatccc tggctctagt tctacatcaa tactttgtc 240
tattgtgat cactaaggac gaaaattccc tatgatttg tcttccgttg gtcctctgc 300
aaccagctt tggctctgtt tctctgcta ttttgccttl ccactccagc ttttctctgc 360
atctacttc atgtgtaat ttgtgggaaa ctataccaca ttttggggag ctgctctgc 420
ctataagtt gatcagtgta aagaacaca acaaaaaa attcgaatag ctatcattga 480
cttctactt ggacttgta ctggaactac agcactgca tctgctattc ttattagaa 540
gctaggtttt ggtggtgct ttctaatat tctctgtctc cttgctgta attgata 600
tattttatt tttctggag atccagtcaa agaggttca tctcagaatg tctactatgc 660
algtagtcaa ggttcaaaa acctatttta ccgaactcac atgcttttta agaatgcto 720
tgytaagaga cgalttttgc tctgtttgtt acttttaca gtaateactt abtttttgt 780
ggtaattgpc attgcccbaa ttttatctc ttaagaaatg gattccccc cctgctggaa 840
tgaatttll atzagttat gatcagcttt ggttagtccc tctttttgca ctagtctct 900
aggaatattg ctlltllcti atgtatgga agatattca atgcccctca ttggatctt 960
taccagatg aoygaatgg ctatgaccgc gtttccagc acaacactga tgaatcttt 1020
agccaggtg ccgtctctt tcaclattg gccaltctc gttctacggt ccagtctgc 1080
aaaaglggtl cttctcagc gacaaggtac cctgtttgct tgtatctct tcttagaac 1140

```

WO 01/46258

PCT/US00/35095

```

aattggaggga gtcactgag ttttacttlt taatggaatt tactcagcca ctggtgcttg 1200
gtacccctggc ttaactttcc tgetgtctgc tgggtotgta ctacttccag ccatacagct 1260
atgtgttgc eagtglacca gctggaatga gggagctat gaacttctta tacaagaaga 1320
atccagtga gatgcttccg acaggtgact gtgatttaaa caaaccaaaa aatctatga 1380
atgcaacat cacataccat gaactctgaa gactataaat gaaitecaca atcagtgct 1440
cactgagac caat tllacc latctllcl tctaaactga acagtoagag agacagctcc 1500
tggotttagc ttot tlygtt aocagocct ttgagocct ttgtgtatc atgaaatata 1560
cttgcnaac acagaacaaa ttccasatc gctcacttt tactctaga agagaacat 1620
taaaactltaa ggggtgagg agggatcnag aaactlgata aggtcaaaag caataatctc 1680
totgacat tccagctct tacactgaga ccaagagaa atctttacc cagtttctta 1740
atcagagaa tgggtthctg cctctctcna gggataaatt tgaaggcata atgaaatata 1800
tgaataaca ctcatggta gaaaataat gatataagtt tcaaatatgt atgattttac 1860
ctatacttg taatgctttg llttatagag cctgttaagc tgcattgat agtggagct 1920
tatatactg gaactctgaa gactatacat gaattccaca atcagtgctt tgttgatata 1980
aaactcttaa agggaggcc cttaaaagaa ctgtatatt tcaclttct taatgcttt 2040
catcgtgac aggtatgata atatttctat atgtaatgg taaltggga aaaaatagat 2100
ataaataaaa ttgctotaaa gaagttaaaa aactgaatga acagctata ctggtataaa 2160
gtaactaatg ttggagcca acatttgttc cttgtctag caaaggata tccatctcc 2220
atgatccctg cctgagaatt ctgctcttag tctttctac ccagctgttg tctatcctg 2280
tcaattata aatactgcta agggcaattt taataataga tctgtatgc cttaaatltg 2340
aatccctcag caagtoact catagyaaaa tgrataaaca agcaagccag tcatgattg 2400
actctctcc atctcaatctc ttaactgctt agctcaatcc tgaagttccac ctltgctctc 2460
aaaaacaca ttgtcttcca ttgctctcag tctttctcna cactyttcaa ttgtctctc 2520
ctccacatt acattgaaac ttccaagct cagtcgaaac attgctctt ctgkataga 2580
gecttttga cctccctct cactcccag tccctacagg gcttccatag ctctlttgt 2640
gactctgat cccagcattt tccatcgact lgtaaattgt tctgtaact gaaatcact 2700
gecttgagta ctgggcaaac ctttgattac tcaatatac tcaataaat atttgtgaa 2760
cta 2763

```

```

<210> 47
<211> 1639
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> Incyte ID No: 1702940CB1

```

```

<400> 47
atcgcactga ggettggatc tgacttctct cccccaccty ctgtgccctt aaactgcaga 60
gatcggggcg ggggttgggg gccaagccgc tcaagtggggt tcaaaaaact ccccaggctc 120
aacctcggtt ctgactgect gagacatggg cagctgacac agcagacett gaatcctgag 180
galtgtaggc aggtatatac tggagggccg gaggacgtgt ctggttatta cacagalca 240
cagctggagc tgggatccac acagctcaga acagttggat ctgtctcagt ctctgtcaga 300
ggaagatccc tgggcaaga ggaacctgac ttggtgtggg agtgaaggta gaggagctg 360
gaa-cgaggt taagaaaaac ctccagct ct ggaagtgac tgaagagctc caagyaayc 420
ccctcggtaa cccagccgct ggcacatga acccagagag cagtatcttt attgagatt 480
accttaagta ttccgggac caagtggca gagagaatct gctacaactg ctgactgat 540
algaactc'g gaattgattc gtgctgctg ctgaactgac cagggalgag gcagatgagc 600
tcgttaagc tetgucaag cttgcaatc acatggtcat gaaggacaaa aaccgcaag 660
ataaagacca gcaagcacag cagtgttttt tgaagagttt tctcgtgttg aaaaaggagc 720
ttgagatca cataaggaag ctccgtgccc ttgcagagga ggttgagcag gtccacagag 780
gcaccacat tgcnaatgtg gtgtccaact ctgttggcac tacctctgyc atcctgaacc 840
tcolcggcct ggtctggca ccttccacag aaggatccag tttgtgctc ttggaactg 900
gcatgggtct gggagagca gctgtctgg ctggatctac ctgaatgtg atgaaachag 960
taaacaaatt cggggcaga gcccaagccc c-aacttggc ccaaadccgc accaatgtag 1020
caaggtgat gaagggtll gtgggtggga acacacccaa tgttcttacc ttagttagaa 1080
atbtgtacca agtcaacaaa gggattggga ggaacatccg tgcacataga cgaagccagag 1140
ccaccctca gtaggagcg tatgcccac ccccgatgt catggggca atctcagctg 1200
aaggcgtga acaggttag aggtlylty aaggcccgc ccaggcaatg agcagagaa 1260
ccatgatcgt ggttcagcc actggaggca tottgetct gctgagatg gtccagcttg 1320
cctatgagtc aaagcacttg cttgagggg caaagtcaga gtoagctgag gactgaaaga 1380
agcggctca gggctgag ggaagctca actttctcac caagatccat gaaatgctc 1440
agccagcca agaccatga cccagagca gtgcagccac caggccgaa atgcccggca 1500
caggcaagga caaatgag acttttttt tttcaagtc tttgacggg aaggagcto 1560
cgctttlcc ccaagtagg gtggcgggc ccaactctg gcctgtgaa cctccgggg 1620
gggggattc gatbaacc 1639

```

WO 01/46258

PCT/US00/35095

<210> 48
<211> 1600
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1703342C31

<400> 48
caaggcggcc caggacaggc agggcctgca cggcgtgaag aancaagac gcagagggc 60
caagcccctt gccttgggto acaacagcaa aggaggcaga gccagaactc acaaccagat 120
ccagaggcaa caggacatg gccacotggg acgaaaaggc agtcaccocy agggccaaay 180
tggctccccc tgaagagatg agcaagtctc taaggcaact caggtctgty ggagacgact 240
accctgctct gaaacatcaac tacaagaatc ggggaaaga agggagggag gaggagagag 300
agcagccccc accaacacca gtctcaggcg aggagggcag agctgcaqcc ctgagctgtg 360
ccccggcccc tggccccqca cccaggggccc cccttgactt caggggcatg ttgaggaaac 420
tgttcagctc ccacaggttt caggctatca tcatctgctt ggtggttctg gatgccctcc 480
tggctgctgc tgaagctcatc ctggacctga agatcatcca gcccgacaag aataactatg 540
ctgcccaggtt atccactac atgagatca ccatcttggc cttttttatg atggagatca 600
tctttaaatt atttctcttc cgcctggagt tcttcaaca caagtttgag atccctggatg 660
ccgtctgtgt ggtggctcca tcatcctctg acattgctct cctgttccag gaggaccagt 720
ttgagctctt ggcctgctg atctctctcc agctgtggcg ggtggccagg atctcaatg 780
ggatatacat ctcaagtaag acacgttccag aaggcaact cttaaggtta aaacagatga 840
atgtcaactt ggcgcgaag attcaacacc ttgagttcag ctgctctgag aaggacaacg 900
aaattgaaag acttaacaaa ctattgcgac agcatggact tcttggtaga gtagactaga 960
cccggaccag ctccctcaa aaagaagaca ctgtctcatg ggcctgtgct gtaagagag 1020
gaacagctgc cctcctggg cgccttgggt agaggtttgg ttgataacct ctgctccct 1080
cctgcagaca tggattctgg gtggacacag ccttgtgaaa ggtccagtae ccccaagagc 1140
tggccatcca ctccaccccc acactgtalc aaatgtatca cactttctca tgttgaacac 1200
tttagcctta attgaaatg agcaacaaag ctggcaact cctagttgta tacaatatt 1260
aaattccctg aatgtacagt ttcaaatctt cagggtata ttaaggact gatgcatctg 1320
agcattctga aagaagaaga aagaagctac tttagctgcc acccattct agaaagctct 1380
cttatttcca agctgttota aatagctctg tctcagtttc ccccaagggt gtaccaggc 1440
ccctcctctg tgtgcccag ctgcatcagc cagctctcag gtgctccal tgtttctcgc 1500
caactgacaa caattttctt caattactgt caaactactg tataaataaa aacaactact 1560
gtataaataa aactctctct tttccctgga aaaaaaaaaa 1600

<210> 49
<211> 2380
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1727529CB1

<400> 49
ctgagccatg gggggaagc agcgggacga ggatgacgag gccctacggga agccagtcas 60
atacagcccc tcccttccag gccccatena gaacagaagc tgcacagatg tcatctgctg 120
cgtctctctc ctgctcttca ttctaggtta catcg'ggtg gggattgtgg cctgggtgta 180
tggagaccoc cggcaagctc tctaccocag gaactctact ggggctact gtggcatggg 240
ggagaacaaa gataagccgt atctcctgta cttcaacatc tcaagctca tctcttccag 300
caactatcct tcaagctgctg agaacggcct acagtgcccc acaccccagg tctgtgtgta 360
ctctctcccg ggggacccat ggcctgttgg aaaaaagag tctcaacaga ctgttgggga 420
agctttctat acsaaaaaca ggaacttttg tctgcccagg gtacccctga atatgacggt 480
gataoaaagc ctgcaacagg aactctgccc cagttctctc ctccctctg ctccagctct 540
ggagcctgct ttctccatgga ccaacattac tccacggcgg ctcccaggga tccacaatga 600
caccaccata cagcagggga tcaagcgtct tattgacagc ctcaatgccc gagaatcag 660
tgttaagatc tttgaagatt ttgcccagtc ctggatctgg attctgtttg ccttgggggt 720
ggctctggtc ttgagcctac tgtttatctt gcttctgccc ctggtggctg ggcctctggt 780
gctgtgctga atcttggag tcttgggggt gctggcaaac ggcactctac cctgtctgga 840
gggtacaga gtcctgggg acaaaggggc ctccatctcc cagctgggtt tcccaacca 900
cctcaagctc tccagagag tgcagggagc ctggctggcc gccctgctc tglttggcgt 960
gcttgaagcc atcctgctgc tgggtctcat ctctctgggg cagcggatcc gtattgcaa 1020
cgcctctctg aaggaggcca gcaagctgtt gggacagatg atgtctacca tgtttctacc 1080
acctgtcaac ttgtctctcc tctctatctg cattyctctac tgggccaatga ctgctctgta 1140

WO 01/46258

PCT/US00/35095

```

ctgggtaca tggggcaac cccagtatgt gctotgggca tccaaacatca gctcccocgg 1200
ctgggggaaa gtgccaataa atacatcatg caacccocag gccccoccttg tgaactcctc 1260
gtgcccaggg ctgatgtgcy tottccaggy ctactcatcc aaagocctaa tccaagctta 1320
tgtctccaat ctgcaaacct atggggctcc ggggtctctc tggaccctta actgggtaact 1380
ggcccaggcc caatggctcc tcyctggagc ctltgcccgc tctactggg ccttccacaa 1440
gcccagagac atcccacact tcccclaat ctctgncttc atccgcacac tccgttaaca 1500
caotgggtca ttggpatttg gagccctcat ctlgaccctt gtcagatag cccgggtcat 1560
cttgggtat attgaccada agctcagagg agtgcagaac ctgttagccc gctgcaatc 1620
gtgctgtttc aagtgtgccc tctggttctt ggaaaaaatt atccagttcc taaaocgcaa 1680
tgcatacacc atgatogcca tctacgggaa gaatttctgt gctccagcca aaaatggctt 1740
catgctaccc atgcaagaca ttgtcaggtt ggtcgtcctg gacaaagtca cagactgct 1800
gctgttcttt gggaaagctgc tgggtgttgg agggtgggg gtccctctct tcttttttt 1860
ctccgttccc atcccggggc tgggtaaaag ctttaagagc ccccaactca actattaact 1920
cttggccatc atgacctcca tctggggggc ctatgctatc gccagcggct tcttcaagct 1980
tttggcatg tgttgggaca cgtctctctc ctgcttctcg gaagacttgg agcggaacaa 2040
gggtctctct gaccggccct ctatcatgtc caagagcctt ctaagatct tgggcaaga 2100
gaaagagggc ccccgggaca accaagaagc gaagaagtga cagctcoggc cctgatccag 2160
gaactgaccc caccocccac gtcccgcact ccaacctcac ttcgcttacc aggtctccat 2220
tttgggttaa aaaaaggctt taggcccagg gcctgtgctc accctgttaa tccaacct 2280
tgagagctg agcggggcgg atcacctgag ccaggagctc gagaccagcc tggccaact 2340
ggtgaacct cctctctat taaaaataca aaatttagc 2380

```

```

<210> 50
<211> 3038
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Inseye ID No: 2289333CB1

```

```

<400> 50
aggggcaagg agcggggcac cagggcgggg tccctccggg caggccagat aggcctgggc 60
ctgaccggcg caagcagcgy gggggagagt gagcaatcgg gggggggcgt cctggagacc 120
cgcagagagt ggaagcggcg ggcagcggcg cggctaccgg gggggcaggg cggagagagc 180
tagatattga tgaatggagg ccttctgata atgacagcaa ttttctgggg acatcagatg 240
aagaacaatga gcaagagctt ctgcccgttc agaagcatta ccaacttgat gatcaagagg 300
gaattcact tgtacaacct ctatgccc cttctaaagg aaatattgga actggccttt 360
taggacttcc attggcaata aaaaagcag gcatagttgt tggcccaac agctctgtgt 420
ttataggaat tattctgttt cactgtatgc acatattggt acgttgcagt caotthcat 480
gtctggagtt taaaagtca acattagull atagtacac tglgaacttt gctatggaag 540
tgagtccctg gattgtctt cagaagcaag uagcctgggg ggggagtgtg gttgacttt 600
ttctggtgat aacacagctg ggtctctgta gtttttatat tgccttctta gctgaaatg 660
tgaacaagt tcatgaagga tccctggaga gtaaatgttt tatttoaat agtaccact 720
cattcaaaacc tttgtgagaga agaagtgttg acctaaggat atatatgctt tgccttcttc 780
cattataaak tctttgggtc tccatctgtg aactaaagaa tctatttcta cttctcatcc 840
ttgcccagct tccatggcti gtcagctctg tgataattta ccagtagttr gtcaggaaac 900
tgcocagatc ccccaacttt ccaatagttg ctggctggaa gaalaccda cctttttgtg 960
gactcctgtt atttctctt gaagcctagc gactgtctct tccaactgaa acccaaatga 1020
aagaatcaaa gcgtttccct caagccttga atattggcat ggggattgtt caactttgt 1080
atgtaacatt agtcaactta ggaatattgt gtllocutga tgaalcaaa ggcagcatab 1140
cttaaatct tcccagat gtatggttat atcaatcagt gaaaattcta tattcctttg 1200
gccttttgt gacatattca atteagttct atgtccagc agagatcat atccctggga 1260
tccatccaa atlctact aatgggaagc aaatctgtga attgggata agactctct 1320
tgggtatgat tacttgttcc gpagcaatc ttatctctg cttagcaatt gttattctc 1380
tgtttggagc tgtgacagc agcaactgg cccatctct gccactttg gttgaaatc 1440
ttacatttcc gaagaaacat tataatataat gtaggtctct gaaaaatatt totatgaat 1500
tcactggagt tcttggcttc ttattaggta calatatac tgltagaaga attatttat 1560
ctactcccaa agttgtagct ggcactccac agagctcttt tcaaatitg aatcaaat 1620
gcttaacatc tggttgaaa tagtaaaagc agaatcctga gttctctatt tttgtccat 1680
ttctgaaat talcaagata actagtaaaa taacttcta tatacaaaa aatggtaaca 1740
aactcgttt tcttgggac gatattaata ctttggagc aatcaaacct ccttaccagt 1800
agtggaac ctatgaaaa tcttgcctt taagtctag caatagttca acaattcag 1860
ttctgaaat tgaacaaat aatagtaaa aaatbaag aatgaaata cttctattat 1920
tctttatct agtaagaaa lacctaaac agatctctc tctttttgtc tactctttg 1980
cactcaact ggaagagaa taggatltca caataagag aataaaaata gacalgtat 2040
aacaanaagc tctctccaga tcatcctgt gaatgccaaa gtaacttcta tctacagt 2100

```

WO 01/46258

PCT/US00/35095

```

aaaaaaaa aatctcagtt atgtttttat tagcccaatt ctaatgattg gctcctggaa 2150
gtatagaana cttccattan catentataa gctcagaaa attgcaaaac ctgaatttaa 2220
ttttcaactc taatggtagt tcatctcoat agtcaagagy caclgttcaa gatcatgaac 2280
tagtgtttca atgaatttg acaggggact ltaaaactta tccagtgcaa cctcctggtt 2340
tttctcaga gaaaaggag gctagaaggt gttaagtac ttgtctgaga ccactcagcc 2400
ttgagatcaa gaaaacttaa tctctgact ccaaggccag gatgttttat tctctcaact 2450
atgtccaga aagaatata attatgttca gcttaatttt agtgttgaat ctathtgatt 2520
atatttlaat accttgaaaa tgaatgtgtg atttttaata gtatagtga cctgagcaga 2580
aaatcaggga actccagaaa gcttactctg tggccatata aacctcagca agagaagaa 2640
gcbatgttct tttaaaacag aatagagacc gcttctgctgt gaaactcctg gctagtaaga 2700
tgtgtgtcta gctatctat tttgtgcttg agotttttta attattacct tcttttcttg 2760
agtttttag gccaccakt cctgaatggc agaaaataga caoctcagaa aacggagagt 2820
tttgggact ttccagccc lgttgctttt ctatctcag ccttttattt abtatagaca 2880
gaataaaga atcagctagg tgtgtgttgc tgtgttata atccccagta cctctggaga 2940
taagtggga ggtcactctg agagccctgg agcttggac cagcctgggc agcatagta 3000
gacctgact ctataaaca taaaaaaaaa aaaaaaaaa 3038

```

```

<210> 51
<211> 2608
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 2720354CB1

```

```

<400> 51
taggctaatt ttttttacag acacgatttc gccacgttgg ccaggctggf cttgaactcc 60
tgaactcaag tgatccccc accctcagct cccaaagtgt tgggattaca ggcgtgagcc 120
actgcaactg gccaggctca tcaacttttg cgcctattgc ctggaagcca gctctctggtg 180
ggacatttag gcaagggtcc tcaagcttag tctgagaca ttggcagctca ctacagcagta 240
tgcacagcat caccgggaga accggccagk ctcaaggaggt agtctatgcc ccaatgagcc 300
tgcaactgac agccatagtg taacaagag ccttaacctg aactggtctg agatctctggg 360
aacccccac cctgtctcca gacgactgct cctttagctg tttgcatact ggcaccocat 420
cctyagaagg catagatacc cggcccaccc tgcctgggaa ttacaaaagt cttagaactgt 480
gcttgagtcg ccggcctcct tgggagaccc tcttaggcag cctaaagacc agaccgggga 540
gctgggtgct atggtctctg ctgctcctct cctcactggac cctctctctc caggttacggc 600
ttcaggtcca gaggctggag aagcctcagc accgctggac gtlgcactgc tcaagttca 660
tctcagca gaagagctg ctggcctgtt acaaaggtct ggccttctcc ctcatggggc 720
tccacttcat caacggctgt gtttctgggg tcaaggcaa cccctctcgg gccctgggc 780
acgactctcc cctcaacag tctctgctgac gtcgctggcc gggcctcctc cagtgctca 840
tctgctgcc catggagctg gccaaagcgc ggtgagctg gcaggacgc ggcaccagcc 900
gcaccacaa ggtctctgtg gactgctctg cgcagatnta ogggcagag ggtctgctg 960
gogtcaaccy gggcatggtg tccacgttgc tgcgtgagac tcccagcttc ggcgtctact 1020
tctcaacta tgaagctctc acgagggcgc tgggctgaga gccggggac cgcctctctg 1080
tgcacagct cctctggcg ggcggtcagc caggaatctg gtcctggtc tcaactatc 1140
ctgtggaag tgtcaagcty cgtctgagc ggaaggacl cggggcgcc ccggtctacc 1200
gggcatctt gacttctgt caccagact accgctcaga ggtctgccc gctctcaacc 1260
gggggtctg gttcaagctg ctgctgctct tcccctcaa cctgctcacc ttggccacc 1320
tcaaggtggt gctcaactac gctgctggcg agyagctgg gcccgagggc gaggctgtgc 1380
cgcgccccc tgggggctct gccctggcgc agccctccag cctgtgacc tccaccgcc 1440
ctctctccc agggctcctt ctcaaaaacc tgggacataa attggccctt gactcagtt 1500
cctctgctcc tgcctggatg ctgctgagctg tggagctctc cagatgtggg ctgaattttg 1560
ctgtctagct ggttagtctt ggcctgagac tgcacttgcc tcaagtgtct catctatgaa 1620
ataagacc ctatggcacc actgttagag caggaagctc agagattatt ccaagagca 1680
gcaagacc gctctgctg aggcattgc accgttatcc tggaaactga ggcagacc 1740
ccagccctt tctggatcc tggccagctc atgtgtctcc tgcctctcag gctgctccc 1800
gggggtctct gatggcacc caaggggcca ccaagggacc tetaactca caatctctc 1860
accgggggg gtagggggcc accctctctg tctgtttag ggcagagga aaacttggty 1920
tgcctctgg tctcaagaa ctggatctct tgcatacccc agotctctca catgcaactg 1980
ctaggggtac ccaagctctt gccactctct ctggagggtg aactggggac cctgcaacct 2040
ccggagacc atgggtctg ctggagacc catatcagcc tccgggacta ggggtgggag 2100
caaacagcc agcgtggag gtcctggcag ttcaagctgt atccagctgt ggcagggaga 2160
aatctctcc cctctggcc ttaggtctgc tgtccmha atggggat ggccctga 2220
cggacaactg agtctccgc ccaactacca ccgcagcca ggatcccca aagtgctcag 2280
aggtctcag agagaacgt atgggcccc claccagcc clggacacc tccagccaca 2340
aagaaacca agtctagctc ctctgtccc cagcaaggy tgcctccag gattctcag 2400

```

WO 01/46258

PCT/US00/35095

tgcacaggct	tactccctgt	gaaggcacag	ggcctgtctg	tggyccacag	gggtggctag:	2460
lggggcctgg	ggcagaggag	ggctgcacca	ggcctctctg	ggatgttctc	cagtgaagac	2520
gaacactggc	ttgcacagc	ctgggtctgc	tgtaacagaa	ctgtcaaggg	aataaagtg:	2580
tctttgtttt	ttaaaaaaa	aaaaaaaa				2608

<210> 51
 <211> 3804
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Inocyte ID No: 3038193CB1

<400> 52
 ccccttccctg tcaactggcta ctaccactcc caaccctcct caaagccgcc ggagcaacc 60
 ccaggctcttt actttacaat cggcaatttg acttgcctctg ctgcaatgtc ggagggaaca 120
 aggaagatgt ggagacgctc caaggatag gtgatcggag ctgaaaaga aaaaagcca 180
 acaaaataaa caaaaaccac ccaccctaac aaatatggag ctgctggaga gaatgagga 240
 agactggctc atgtctggaa tagtctctgc gatctctga cctaaactgg agccgtccat 300
 aggggtgcaat ggggaccac tgaagccaga aataactgta tccacacttg ctgttgcac 360
 aatallcctt aacagtgga tactatgaa aacagaggag ctgaccagtg ctttggtgca 420
 tctaaactgt tctctttta ttoagttctt tctctctgca ttcttccag caactatag 480
 gcttttctct cagcttttat caatcacacc cctcaacgaa tggcttttaa aaggtttgca 540
 gacagtgggt tgcctgcctc cgcctgtgic ttctgcagtg allttaaaca aggcagttg 600
 tggaaatgag ggcactgcta taacaccctt gctctctctg ctttttctg gttcactctc 660
 tctctgctct tccacatcta ttttttctca gcttttatg actgtgttg tctctctcat 720
 catggacag atgtctcgaa gatacatcaa ggattggctt gagagaana agcctccttt 780
 tgggtctatc agcagcagtg tactctctat gatcactac caaacattct gtgacagtt 840
 ctcaaccoca aatatgacc tggataaatt cagcctgttt ctcaactgt tcaataatt 900
 tctctccag ctgagttta tggctttaa tctctcttt tcaacaaga acaattggg 960
 tttcaacca ccagacacag tggctctcat ttctgttct acacacaat cctttactt 1020
 gggaaattcc agctgaaga tctgttttc agctatgag catctctct taatatctg 1080
 accctgtctc atctaccacc cagctcagat ccttctggga agtgtgttg tggcaacaat 1140
 caagtcttgg atggtatcaa ggcagagaa actactcaca accagggggc cactggctaa 1200
 cttgaataat ccagaaggct tggaaatatc atccatcaaa ttggggcatt aaaaataca 1260
 ccaaggtcc atctctcagg gactgaact gacaagccc acagtatac aaagggagt 1320
 gacttctgt agcaatgat atatgacag gctgtctcat actagcaat ctggagcct 1380
 gtaacttgga atggtgctc atgcaatatt tttatctttt ccacaaaat atgcaatct 1440
 gtttaagtc cttcaaatgt attgaccag agcgttattt ccacaatg ctttgttat 1500
 tactgcccgg ggtggtacaa tatttggggg ttaattttgc ttctaatg cagggaatcag 1560
 tcaatglaag tgcacaaaag caaacatgct tctctctcag caccttctg taatacaacc 1620
 ctatngtaag tactgtaag ttgaaatga ggtcacacca tcaggaaaat gccctctga 1680
 tgaactgaa aatttcaaaa gtcttatca tgaactcttt gatttactgt gtattcttt 1740
 ttcttcaaga ctgtgacatg cctctctct atcaactcag caggggtcat agatcgaata 1800
 gatctgaaa agcttaaat atatgcaatc ctgtgacata ttcttaaga ctttctctca 1860
 aatagttcc ccaagaaat tctctctctc cttatgaga gattgtggtt atatgtotta 1920
 aatttattat aagctctctc aagaagaggg hetgaattt tgaattatga gtraaatcat 1980
 gtgaaatttt gagttaaact ctgtgatttg atttccagg ttcttaaaal atactttaa 2040
 atcttctctc tctttatca ataatctctg tcttgcactt acacactctc aacagcaaaa 2100
 tatgaggcac aaaaatgta caatcagttt gaaagcagca tcaattaatg gtatattota 2160
 tcaacattcc acaaccaga ccaaattttt tctctatcac gcagatgtgc tggacccttt 2220
 ccagattgccc cctgttggcc aaaagcagcc tgttatcccc tggaaatlaag cacacttaag 2280
 gtaattgaga caatttaata atgaaattct ccttggcaga ttggcaaat gtggcaata 2340
 tttttttaa agttaaata tattgtctct atgaaataa gaaatataa agtcaatga 2400
 tgcacaacaa tttlacatal acacattctg tctctccaga tpaasaagac atgcaaaaco 2460
 atttaaatac caaaatatac agtaaaatla gttcccaagc gggcagcagc tttcaatga 2520
 gtgtccaata tttgtctctg ctatagctgc aagaactgta actggaccca agtagagat 2580
 gaagcagct atagaaactac gagaacactt tctgtgtttt ccccaatgcc gtctgtctac 2640
 atctcttac agtctctctc ttgatltgat agcaaatatt ggcactctgg gtctcactga 2700
 ggcctgtcta tgtctctcagc agctgttttt gtgtttctgt tattatgccc acaaaaaaa 2760
 atactctctt agaaactcac caagtctatc tactgtgtaa atttatatta ctgttactac 2820
 caggtctcat ctttctgcaa tgtctatgaa taaattctat aagagttctt ctgagtgga 2880
 attttaagc taatgcaaga tctgcacaaa atctatgcta accagcclgt agtcaactc 2940
 gttataaga attttaactg tgtctaaaac taocagtaat ttgottaggt aattgtgott 3000
 acotatggag cacaggaag cctctagttt ttgtctctac aagtctcttt gaatttggga 3060
 gtaaatggaa gtgtctgtct gtcgtctctc taltctgccc tcttctctcc 3120

WO 01/46258

PCT/US00/35095

```

taacattaaa atactgatcc cgcgccccaa cttatctacc tctattgtct aacacctata 3199
gtagggtgga tcatgggata naattcaact gaanaatgcta tgateacatt ttatcgtttg 3240
ctttaaaaat gtgctttgtt ttcaaatat cttacatag tgaacttgg tggcgttagt 3300
gatattgtta tgcctatttc ttttttttac acaaatccct tggcatattt ttctaatag 3360
aacaataaat aaaaacaaaa ttttatttta attaatgctt attgggattt aattattcag 3420
agcttaaaat attttgttat gttataacac tgaagagota tctgttttat gcatttgttt 3480
tgttcaaatg tatttatgaa agnaaatat tagat'abat t'atgtttac tca'lttt'cc 3540
acctggattt ttttcaatgg tigttaaaaa attagatttt ttaaggggta ataattgttg 3600
tattttcatg tttttcttta gatttaaaat ttttgggtgt tttttaaant ttttccclal 3660
tctgttaaaa attaacacac ccttagctaa tgttcagtgt ttgtgtleaa taccaaattt 3720
tttcaaaagg attgtt'aa'g tcatanagt'g gattttttat gatgact'gga agat'gaaaat 3780
aattatatga ttaaaaaaag aatg                                     3804

```

```

<214> 53
<211> 1894
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3460979CB1

<220>
<221> Unsure
<222> 1651
<223> a, t, c, g, or other

```

```

<400> 53
acgatacaact agtatggcgc gcaatgtgct ggaaggggaa caaacatggc cgotctggcg 60
ccccctgggt ccccgcttc cggcgttct agctggcg cggcctccg cctgctccca 120
atgctgggtt tgcctcagtt gctgcccag cctggcctgg gcggcgtcca tcaactggca 180
ctcaaggatg atgtgagcca taagtctcat ctgaacacct ttggctcttt cangyatggg 240
tacaatgttg tgaatgctag tagcctctca ctgaaatgag ctgaagacaa ggaatgact 300
attgatttia gcttagaccg tacaagaat gatgcttct ctcttacct ggaatgact 360
gtgaattact gtattttaa gaaacagtct gtctctgcca cccittaat cctagacato 420
tccagaagtg agttagagt aaagtctcca ccagaagctg gtacccagtt accaaagatc 480
atcttcagca ggaatgaga agtcttgggt cagagccagg agcctaagg taacctgct 540
tccagcagca accagccca gaagacaaa gatggggaa agtcaaaa'g agtcaagt'g 600
gattcaaa'g ccatgggaga gaattctt'ct tctgtt'cata ataatgg't'g ggcagt'g'ca 660
tttcaagt'tt ttttacaat cagcactgat gacaaaga'g gcc'tt'acag tctt't'at'tt 720
cataaatt'cc ttggaaaaga attgccaagt gacaagt'tta cattoagcct tgattatt'gag 780
atcaacag'ga agaatccl'ga cagctacctc tcaycagg'ag aat'tcct'ct cccaaat'ta 840
cacatctcaa tggcct'tt'tt ctctt'tt'tt tctgg'g'acca tctg'g'at'cca tctcct't'ga 900
aaa'c'g'ag'ga atg'at'g'at't taaa't'cc'ac tgg'ot'g'at'gg cgg'cc'ct'cc t'tc'c'caa'g 960
tctctt'ct'ct t'gt'g't'cca t'g'a'a't'g'ac t'ac'c'a'c'aa t'ct'c'c'cca g'g'ct't'cc'ct 1020
atcgaag'cct g'g'ct'g't't'gt g'c'a't'c'a't'a act'c'ac't'tt t'g'a'a'g'g'g'c g'c'a'c't'ct'to 1080
atcacc'att'g cact'c't'bgg oact'g'ct'gg gctt'c'a't'a agc'ac't'ct t'ct'g'e't'a'a 1140
gacaaaa'ga tctt'c'at'gat t'g't'a'at't'ca ct'c'ag't'cc t'g'c'c'a'at'gt agc't'c'at'o 1200
atcatag'agt ccacc'g'ag'ga g'g'g'c'ag'act gaat'at'g'ct t'g't'g'a'ag'ga ct'ct'c'at'tt 1260
ctg'gt'oy'ac c't't'g't'g't'g t'g'g't'g'c'at'c ct'ct't'cc'ag t'g'g't'g'g't'o aat'c'ag'act 1320
ttacaagaag catcagcaac agatgganaa gctgctatta acttgcaaa gctgaaactt 1380
ttcagcaact attagctct gattgtgtgt t'c'c'at'act t'c'ac't'agg'at cat't'g'c'at't 1440
ctctc'aa'ac t'c'ct'g't'cc at'c'c'ag't'g'g aag'g'g'ct'ct ac'c'ag'ct'ct g'g'at'g'a'a'c'g 1500
gc'c'ac'ct'g'g tcttctt'gt tctaacgggg tataaattcc gtc'c'g'ct'tc agataa'ccc 1560
tacctaccac tttctcagga agaaagagac ttggaaa'gg agt'c'c'g't'g'ta agaaat'ct'tt 1620
cttccct'clt ccttagccct gazccct'tg nctaacacaa agc'ag'c'ag'c'g t'g't'g'a'at'c'g'a 1680
gc'c'g'c't'g'gt ct'c'ag'c'at't ct'g'g'ot'g'oa g'g'g't'g'g't'o ct'c'at'at'tt ag'c'g'a'ag'g'g 1740
acc'g'c'act'g gag'c'c'a'gg g'g't'c'g't'c'g g't'g'a'a'g'g'ca ag'at't'g'g'ca acc'c'at'c't'g 1800
gctgt'cc'g'g aaagg'ana'g g'g'g'g'g'c'ca a'a'c'c'a'at't'g g'g'g'c'g'g'c't c'aaaa'ac'c'g 1860
g'g'c'g'a'c'a'ag agaaaaagcg g'g'c'c'ag'g'ag aa'ag                                     1894

```

```

<210> 54
<211> 1668
<212> DNA
<213> Homo sapiens

<220>

```

WO 01/46258

PCT/US00/35095

<221> misc_feature

<223> IncyTe ID No: 7472200CB1

<400> 54

```

atgacactgg tttactttcc tcttccaag cttcagcagc agcagcagcc atcgagatcc 60
agtgcctctg cccaacagtt ggcccacatc tcttggcagc tggccctgag ctttggcaaa 120
cggaccacta tccacggcct ggacaggctg cttagtcca aggcacatcg atggagcga 180
ttctgtctgg tgtgcacett tctgcaagcc ttcttgggg cgggtacgt ttgcttgatt 240
ctctccggcc gctacaacgc cgtccacctc cagacggtgg tggatagcgc ggggtttcgg 300
gthttaccga taccatttc ggtcataacg atttgcaacc ggaatgcct caactggcaa 360
cgcctggcgg aggggaagtc aagattctct gccaacggca gcaactccgc ccagcggagg 420
ctcttcgagc tgattgtggg caactacgac gctgcttact tcygtcaact tcagtccttc 480
gagcgaattc gcaaccagcc aacggagctc ctcaactatg tcaatttcag ccaggtggtg 540
gatlttatga cctggcgtcg caacggagctg ctccgggaat gctgtgtggc caaccatgcc 600
taagactgct cggagatccg ctgcaagcgg cgcagcaaga accgcttctg ctgggctttc 660
aactcctctg agaccgaaga gggcaggctg atgcagctgc tggatcccat gggccctggg 720
cgtactgggt cggcgggtcc caagagcgc ctctcctgac gttctctcat ccagccctcg 780
aagcactggc cgggcacccg ggaacgaat gccatgaagc gcatcgatgt catggttac 840
gagcattttg tglggcaca caatccttct ttctgtggcc cgaacacgga gacgacctg 900
gagatgagac ccttcatcta cttctatgac aacgacccc ggggagttcg ctccgaccag 960
cctccagtgc tottogatga tgagcacaac agcaaggatt tcaagtcct gcaaggatc 1020
gtttacatga tggaaaactg tcagtccyag tgcactcagg agtacttggc ggcctattgc 1080
aactgcacaa tggactact gthttccacc gacctgctca tctactccca caatcccg 1140
gagaaggagt tegtctgcaa ccaatttcag ggaatgctc gcaagtgcct ccgcaactgc 1200
tactccctca actacatcag cnatgtccgg cccgccttcc tgcaccgga tctgtaccga 1260
aacaacccct abtggaccc ggtgtgcaac ttctccttcc agaccattat ggtctatcgc 1320
accagctctg tottccgctg ggtggactta atggttagct ttggaggaat tgcctgtctt 1380
ttctctggct gctccclaat tagtggcatg gaactggcct atttccctgt cattgaggtg 1440
cggccttttg ggtcggatgg actgctcga aggtggaagg ctccagggca gatggatctg 1500
ggcgtaacgg tgcccacgcc cactltgaac ttccaacaa ccacgcccag tcaagtatg 1560
gagaactaca ttatgcaact gaagcctgag aagggcaac agcagaaggc gaactttcaa 1620
aactggcacc gcataacatt tgcctcaaaag catgttatg gcaagtga 1668

```

【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
28 June 2004 (28.06.2004)

PCT

(10) International Publication Number
WO 01/46258 A3

(51) International Patent Classification: C07K 14/705,
A61K 38/17, C07K 16/18, C12N 5/00, 15/63, C12Q 1/68

(21) International Application Number: PCT/US00/25095

(22) International Filing Date:
22 December 2000 (22.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
69/172,000 24 December 1999 (23.12.1999) US
69/176,083 14 January 2000 (14.01.2000) US
69/177,332 21 January 2000 (21.01.2000) US
69/178,572 28 January 2000 (28.01.2000) US
69/179,758 2 February 2000 (02.02.2000) US
69/181,625 10 February 2000 (10.02.2000) US

(US) LU, Dyoung, Anna, M. [US/US]: 233 Coy Drive, San Jose, CA 95124 (US); YANG, Junming [CN/US]: 7125 Bark Lane, San Jose, CA 95129 (US); REDDY, Roopa [IN/US]: 1233 W. McKinley Avenue #3, Sunnyvale, CA 94086 (US); LAL, Preeti [IN/US]: 2382 Lass Drive, Santa Clara, CA 95054 (US); HILLMAN, Jennifer, L. [US/US]: 250 Marner Drive #17, Mountain View, CA 94040 (US); AZEMZAI, Yelda [US/US]: 5518 Bonita Canyon Drive, Castro Valley, CA 94552 (US); YUE, Henry [US/US]: 826 Lois Avenue, Sunnyvale, CA 94087 (US); NGUYEN, Damiel, B. [US/US]: 1405 Ridgewood Drive, San Jose, CA 95118 (US); YAO, Monique, G. [US/US]: 111 Frederick Court, Mountain View, CA 94043 (US); GANDHI, Ameena, R. [US/US]: 837 Roble Avenue #1, Menlo Park, CA 94025 (US); TANG, Y., Tom [CN/US]: 4230 Ramwick Court, San Jose, CA 95118 (US); KHAN, Farrah, A. [IN/US]: 333 Escuela Avenue #221, Mountain View, CA 94040 (US).

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

(65) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:
US 69/172,000 (CIP)
Filed on 23 December 1999 (23.12.1999)
US 69/176,083 (CIP)
Filed on 14 January 2000 (14.01.2000)
US 69/177,332 (CIP)
Filed on 21 January 2000 (21.01.2000)
US 69/178,572 (CIP)
Filed on 28 January 2000 (28.01.2000)
US 69/179,758 (CIP)
Filed on 2 February 2000 (02.02.2000)
US 69/181,625 (CIP)
Filed on 10 February 2000 (10.02.2000)

(81) Designated States (national): AF, 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, GR, HU, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LF, LK, LR, LS, LU, LV, MA, MD, MG, MK, MN, MW, MX, MY, NI, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

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

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

Published:
with international search report

(72) Inventors; and
(73) Inventors/Applicants (for US only): BAUGHN, Mariah, R. [US/US]: 11211 Santiago Road, San Leandro, CA 94577 (US); BURROLD, Neil [GB/US]: 103 Whitehead Circle, Durham, CT 06832 (US); AL-YOUNG, Janice [US/US]: 233 Golden Eagle Lane, Brisbane, CA 94005

(85) Date of publication of the international search report:
24 January 2002

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



WO 01/46258 A3

(54) Title: TRANSPORTERS AND ION CHANNELS

(57) Abstract: The invention provides human transporters and ion channels (TRICH) and polynucleotides which identify and encode TRICH. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of TRICH.

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PC 00/25295
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C07K14/705 A61K38/17 C07K16/18 C12N5/00 C12H15/63 C12Q1/68		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C07K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) BIOSIS, EPO-Internal, NPI Data, PAJ, EMBL		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] EBI; Accession No: A1094911, 18 August 1998 (1998-08-18) NCI/NINDS-CGAP: "Integral Membrane Protein" XP002162957 abstract	1-28
A	EP 0 834 561 A (UNIV LEIDEN) 8 April 1998 (1998-04-08) the whole document	1-28
	--- -/-	
<input checked="" type="checkbox"/> Further documents are listed in the continuation of C.		
<input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents:		
"A" document defining the general state of the art which is not considered to be of particular relevance		"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier document but published on or after the international filing date		"X" documents 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
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another claim or other special reason (as specified)		"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" documents referring to an oral disclosure, use, or exhibition or other means		"Z" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 16 March 2001		Date of making of the international search report 18.06.01
Name and mailing address of the ISA European Patent Office, P.O. Box 5819, Patenkamp 2 NL - 5200 PH Rijswijk Tel: (+31-70) 340-2040, Tx: 31 881 000 nl Fax: (+31-70) 340-3018		Authorized officer Schwachtgen, J-L

Form PCT/ISA/10 (second sheet) July 1992)

INTERNATIONAL SEARCH REPORT

International Application No.
PCT / 00/35095

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category	Relevant to claim No.
P,X	1-28
IBBERSON MARK ET AL: "GLUTX1, a novel mammalian glucose transporter expressed in the central nervous system and insulin-sensitive tissues." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 275, no. 7, 18 February 2000 (2000-02-18), pages 4607-4612, XP002162956 ISSN: 0021-9258 cited in the application the whole document -----	

Form PCT/ISA210 (Continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Publication No.
PCT/US 00/35095

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Although claims 18 and 24 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the composition.
2. Claims Nos.: 20, 23
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this International application, as follows:

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-26 all partially

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 00/35095

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-28 all partially; Invention 1

An isolated polypeptide comprising at least a fragment of a transporter or receptor protein comprising the amino acid sequence SEQ ID NO:1 and subject-matter referring to said polypeptide.

2. Claims: 1-28 all partially; Inventions 2-28

Idem as subject 1 but limited to each of the polypeptides comprising the amino acid sequences SEQ ID NO:2-27.

International Application No. PCT/US 00/35095

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 219

Continuation of Box 1.2

Claims Nos.: 20, 23

Claims 20 and 23 are directed to agonists and antagonists of the polypeptide of SEQ ID NO:1. As no such compounds are defined in the application the subject-matter of said claims is considered to be purely speculative and unclear in the sense of Article 6 PCT to a degree that no meaningful search can be carried out.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.
PCT/JP 00/35002

Patent document cited in search report	Publication date	Patent family members	Publication date
EP 0834561 A	08-04-1998	AU 4402197 A	17-04-1998
		EP 0932677 A	04-08-1999
		JP 2001500743 T	23-01-2001
		WO 9813490 A	02-04-1998

フロントページの続き

(51) Int.Cl. ⁷	F I	テーマコード(参考)
A 6 1 P 3/00	A 6 1 P 3/06	4 C 0 8 4
A 6 1 P 3/06	A 6 1 P 3/10	4 H 0 4 5
A 6 1 P 3/10	A 6 1 P 7/00	
A 6 1 P 7/00	A 6 1 P 7/06	
A 6 1 P 7/06	A 6 1 P 7/08	
A 6 1 P 7/08	A 6 1 P 9/00	
A 6 1 P 9/00	A 6 1 P 9/10	
A 6 1 P 9/10	A 6 1 P 9/12	
A 6 1 P 9/12	A 6 1 P 11/06	
A 6 1 P 11/06	A 6 1 P 13/02	
A 6 1 P 13/02	A 6 1 P 15/00	
A 6 1 P 15/00	A 6 1 P 17/00	
A 6 1 P 17/00	A 6 1 P 17/16	
A 6 1 P 17/16	A 6 1 P 19/02	
A 6 1 P 19/02	A 6 1 P 19/08	
A 6 1 P 19/08	A 6 1 P 19/10	
A 6 1 P 19/10	A 6 1 P 21/00	
A 6 1 P 21/00	A 6 1 P 25/00	
A 6 1 P 25/00	A 6 1 P 25/02	1 0 3
A 6 1 P 25/02	A 6 1 P 25/06	
A 6 1 P 25/06	A 6 1 P 25/08	
A 6 1 P 25/08	A 6 1 P 25/14	
A 6 1 P 25/14	A 6 1 P 25/16	
A 6 1 P 25/16	A 6 1 P 25/18	
A 6 1 P 25/18	A 6 1 P 25/24	
A 6 1 P 25/24	A 6 1 P 25/28	
A 6 1 P 25/28	A 6 1 P 27/02	
A 6 1 P 27/02	A 6 1 P 27/12	
A 6 1 P 27/12	A 6 1 P 29/00	
A 6 1 P 29/00	A 6 1 P 29/00	1 0 1
A 6 1 P 31/04	A 6 1 P 31/04	
A 6 1 P 31/10	A 6 1 P 31/10	
A 6 1 P 31/12	A 6 1 P 31/12	
A 6 1 P 31/18	A 6 1 P 31/18	
A 6 1 P 33/00	A 6 1 P 31/18	
A 6 1 P 35/00	A 6 1 P 33/00	
A 6 1 P 37/00	A 6 1 P 35/00	
A 6 1 P 37/08	A 6 1 P 37/00	
A 6 1 P 43/00	A 6 1 P 37/08	
C 0 7 K 14/705	A 6 1 P 43/00	1 1 1
C 0 7 K 16/28	C 0 7 K 14/705	
C 1 2 N 1/15	C 0 7 K 16/28	
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 P 21/02	C 1 2 P 21/02	C
C 1 2 Q 1/02	C 1 2 Q 1/02	

C 1 2 Q 1/68	C 1 2 Q 1/68	A
G 0 1 N 33/15	G 0 1 N 33/15	Z
G 0 1 N 33/50	G 0 1 N 33/50	Z
G 0 1 N 33/53	G 0 1 N 33/53	M
G 0 1 N 33/566	G 0 1 N 33/566	
	C 1 2 N 5/00	A
	A 6 1 K 37/02	

- (31)優先権主張番号 60/178,572
 (32)優先日 平成12年1月28日(2000.1.28)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/179,758
 (32)優先日 平成12年2月2日(2000.2.2)
 (33)優先権主張国 米国(US)
 (31)優先権主張番号 60/181,625
 (32)優先日 平成12年2月10日(2000.2.10)
 (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 バーフォード、ニール
 アメリカ合衆国コネチカット州06422・ダラム・ワイルドウッドサークル 105
 (72)発明者 オウ・ヤング、ジャニス
 アメリカ合衆国カリフォルニア州94005・ブリスベン・ゴールドデンイーグルレーン 233
 (72)発明者 リュ、デュング・アイナ・エム
 アメリカ合衆国カリフォルニア州95123・サンノゼ・コイドライブ 233
 (72)発明者 ヤング、ジュンミン
 アメリカ合衆国カリフォルニア州95129・サンノゼ・パークレーン 7125
 (72)発明者 レディ、ルーパ
 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
 (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95054・サンタクララ・ラスドライブ 2382
 (72)発明者 ヒルマン、ジェニファー・エル
 アメリカ合衆国カリフォルニア州94040・マウンテンビュー・#17・モンロードライブ 230
 (72)発明者 アジムザイ、ヤルダ
 アメリカ合衆国カリフォルニア州94552・カストロパレー・ボールダーキャニオンドライブ 5518
 (72)発明者 ユエ、ヘンリー
 アメリカ合衆国カリフォルニア州94087・サニーベイル・ルイスアベニュー 826
 (72)発明者 ニュエン、ダニエル・ビー
 アメリカ合衆国カリフォルニア州95118・サンノゼ・リッジウッドドライブ 1403
 (72)発明者 ヤオ、モニック・ジー
 アメリカ合衆国カリフォルニア州94043・マウンテンビュー・フレデリックコート 111
 (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 タング、ワイ・トム
 アメリカ合衆国カリフォルニア州 9 5 1 1 8 ・サンノゼ ・ランウィックコート 4 2 3 0
 (72)発明者 カーン、ファラ・エイ
 アメリカ合衆国カリフォルニア州 9 4 0 4 0 ・マウンテンビュー ・ # 2 2 1 ・エスケラアベニュー
 3 3 3

F ターム(参考) 2G045 AA40 DA12 DA13 DA14 DA36 FB02
 4B024 AA01 AA11 BA44 BA63 CA01 CA04 CA09 CA11 DA01 DA02
 DA05 DA11 EA01 EA02 EA03 EA04 FA02 GA01 GA11 HA01
 HA03 HA11
 4B063 QA01 QA18 QA19 QQ01 QQ13 QQ42 QQ52 QR08 QR33 QR42
 QR55 QR59 QR62 QR74 QR80 QS05 QS25 QS34 QS36 QX02
 4B064 AG20 AG26 CA01 CA19 CA20 CC24 DA01 DA13
 4B065 AA01X AA57X AA87X AA93Y AB01 AB02 BA01 BA08 CA24 CA25
 CA44 CA46
 4C084 AA02 AA17 BA22 CA18 CA53 DC50 NA14 ZA01 ZA06 ZA07
 ZA08 ZA12 ZA15 ZA16 ZA18 ZA20 ZA22 ZA24 ZA33 ZA36
 ZA42 ZA43 ZA45 ZA51 ZA52 ZA54 ZA55 ZA59 ZA66 ZA81
 ZA89 ZA94 ZA96 ZA97 ZB07 ZB08 ZB11 ZB13 ZB15 ZB26
 ZB31 ZB32 ZB35 ZB37 ZB38 ZB39 ZC35 ZC42 ZC55
 4H045 AA10 AA11 AA20 AA30 BA10 CA40 DA50 DA75 EA20 EA50
 FA72 FA74

专利名称(译)	转运蛋白和离子通道		
公开(公告)号	JP2004500814A	公开(公告)日	2004-01-15
申请号	JP2001547167	申请日	2000-12-20
[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	ボーグンマライアアール バーフォードニール オウヤングジャンニス リュデュングアイナエム ヤングジュンミング レディルーパ ラルプリーティ ヒルマンジェニファーエル アジムザイヤルダ ユエヘンリー ニュエンダニエルビー ヤオモニークジー ガンディーアミーナアール タングワイトム カーンファラエイ		
发明人	ボーグン、マライア・アール バーフォード、ニール オウ・ヤング、ジャンニス リュ、デュング・アイナ・エム ヤング、ジュンミング レディ、ルーパ ラル、プリーティ ヒルマン、ジェニファー・エル アジムザイ、ヤルダ ユエ、ヘンリー ニュエン、ダニエル・ビー ヤオ、モニーク・ジー ガンディー、アミーナ・アール タング、ワイトム カーン、ファラ・エイ		
IPC分类号	A01K67/027 A61K38/00 A61K39/00 A61K45/00 A61P1/00 A61P3/00 A61P3/06 A61P3/10 A61P7/00 A61P7/06 A61P7/08 A61P9/00 A61P9/10 A61P9/12 A61P11/06 A61P13/02 A61P15/00 A61P17/00 A61P17/16 A61P19/02 A61P19/08 A61P19/10 A61P21/00 A61P25/00 A61P25/02 A61P25/06 A61P25 /08 A61P25/14 A61P25/16 A61P25/18 A61P25/24 A61P25/28 A61P27/02 A61P27/12 A61P29/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P37/00 A61P37/08 A61P43 /00 C07K14/47 C07K14/705 C07K16/28 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12P21 /02 C12Q1/02 C12Q1/68 G01N33/15 G01N33/50 G01N33/53 G01N33/566		
CPC分类号	A61K39/00 A61P1/00 A61P3/00 A61P11/06 A61P13/02 A61P15/00 A61P17/00 A61P17/16 A61P19/02 A61P19/08 A61P19/10 A61P21/00 A61P25/00 A61P25/02 A61P25/06 A61P25/08 A61P25/14 A61P25 /16 A61P25/18 A61P25/24 A61P25/28 A61P27/02 A61P27/12 A61P29/00 C07K14/47		
FI分类号	C12N15/00.ZNA.A A01K67/027 A61K45/00 A61P1/00 A61P3/00 A61P3/06 A61P3/10 A61P7/00 A61P7 /06 A61P7/08 A61P9/00 A61P9/10 A61P9/12 A61P11/06 A61P13/02 A61P15/00 A61P17/00 A61P17		

/16 A61P19/02 A61P19/08 A61P19/10 A61P21/00 A61P25/00 A61P25/02 A61P25/02.103 A61P25/06 A61P25/08 A61P25/14 A61P25/16 A61P25/18 A61P25/24 A61P25/28 A61P27/02 A61P27/12 A61P29/00 A61P29/00.101 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P37/00 A61P37/08 A61P43/00.111 C07K14/705 C07K16/28 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12Q1/02 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.M G01N33/566 C12N5/00.A A61K37/02

F-TERM分类号

2G045/AA40 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/FB02 4B024/AA01 4B024/AA11 4B024/BA44 4B024/BA63 4B024/CA01 4B024/CA04 4B024/CA09 4B024/CA11 4B024/DA01 4B024/DA02 4B024/DA05 4B024/DA11 4B024/EA01 4B024/EA02 4B024/EA03 4B024/EA04 4B024/FA02 4B024/GA01 4B024/GA11 4B024/HA01 4B024/HA03 4B024/HA11 4B063/QA01 4B063/QA18 4B063/QA19 4B063/QQ01 4B063/QQ13 4B063/QQ42 4B063/QQ52 4B063/QR08 4B063/QR33 4B063/QR42 4B063/QR55 4B063/QR59 4B063/QR62 4B063/QR74 4B063/QR80 4B063/QS05 4B063/QS25 4B063/QS34 4B063/QS36 4B063/QX02 4B064/AG20 4B064/AG26 4B064/CA01 4B064/CA19 4B064/CA20 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA01X 4B065/AA57X 4B065/AA87X 4B065/AA93Y 4B065/AB01 4B065/AB02 4B065/BA01 4B065/BA08 4B065/CA24 4B065/CA25 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA17 4C084/BA22 4C084/CA18 4C084/CA53 4C084/DC50 4C084/NA14 4C084/ZA01 4C084/ZA06 4C084/ZA07 4C084/ZA08 4C084/ZA12 4C084/ZA15 4C084/ZA16 4C084/ZA18 4C084/ZA20 4C084/ZA22 4C084/ZA24 4C084/ZA33 4C084/ZA36 4C084/ZA42 4C084/ZA43 4C084/ZA45 4C084/ZA51 4C084/ZA52 4C084/ZA54 4C084/ZA55 4C084/ZA59 4C084/ZA66 4C084/ZA81 4C084/ZA89 4C084/ZA94 4C084/ZA96 4C084/ZA97 4C084/ZB07 4C084/ZB08 4C084/ZB11 4C084/ZB13 4C084/ZB15 4C084/ZB26 4C084/ZB31 4C084/ZB32 4C084/ZB35 4C084/ZB37 4C084/ZB38 4C084/ZB39 4C084/ZC35 4C084/ZC42 4C084/ZC55 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/CA40 4H045/DA50 4H045/DA75 4H045/EA20 4H045/EA50 4H045/FA72 4H045/FA74

优先权

60/172000 1999-12-23 US
 60/176083 2000-01-14 US
 60/177332 2000-01-21 US
 60/178572 2000-01-28 US
 60/179758 2000-02-02 US
 60/181625 2000-02-10 US

外部链接

Espacenet

摘要(译)

本发明中，人的转运和离子通道 (TRICH) ，并确定多核苷酸编码 TRICH。本发明还提供了表达载体和宿主细胞，抗体，激动剂，拮抗剂。此外，本发明提供了诊断，治疗和预防与TRICH异常表达相关的疾病的方法。

核苷酸 编号 ID	核苷酸 编号 ID	核苷酸 编号 ID	核苷酸 编号 ID	核苷酸 编号 ID
146107	146108	146109	146110	146111
146112	146113	146114	146115	146116
146117	146118	146119	146120	146121
146122	146123	146124	146125	146126
146127	146128	146129	146130	146131
146132	146133	146134	146135	146136
146137	146138	146139	146140	146141
146142	146143	146144	146145	146146
146147	146148	146149	146150	146151
146152	146153	146154	146155	146156
146157	146158	146159	146160	146161
146162	146163	146164	146165	146166
146167	146168	146169	146170	146171
146172	146173	146174	146175	146176
146177	146178	146179	146180	146181
146182	146183	146184	146185	146186
146187	146188	146189	146190	146191
146192	146193	146194	146195	146196
146197	146198	146199	146200	146201
146202	146203	146204	146205	146206
146207	146208	146209	146210	146211
146212	146213	146214	146215	146216
146217	146218	146219	146220	146221
146222	146223	146224	146225	146226
146227	146228	146229	146230	146231
146232	146233	146234	146235	146236
146237	146238	146239	146240	146241
146242	146243	146244	146245	146246
146247	146248	146249	146250	146251
146252	146253	146254	146255	146256
146257	146258	146259	146260	146261
146262	146263	146264	146265	146266
146267	146268	146269	146270	146271
146272	146273	146274	146275	146276
146277	146278	146279	146280	146281
146282	146283	146284	146285	146286
146287	146288	146289	146290	146291
146292	146293	146294	146295	146296
146297	146298	146299	146300	146301
146302	146303	146304	146305	146306
146307	146308	146309	146310	146311
146312	146313	146314	146315	146316
146317	146318	146319	146320	146321
146322	146323	146324	146325	146326
146327	146328	146329	146330	146331
146332	146333	146334	146335	146336
146337	146338	146339	146340	146341
146342	146343	146344	146345	146346
146347	146348	146349	146350	146351
146352	146353	146354	146355	146356
146357	146358	146359	146360	146361
146362	146363	146364	146365	146366
146367	146368	146369	146370	146371
146372	146373	146374	146375	146376
146377	146378	146379	146380	146381
146382	146383	146384	146385	146386
146387	146388	146389	146390	146391
146392	146393	146394	146395	146396
146397	146398	146399	146400	146401
146402	146403	146404	146405	146406
146407	146408	146409	146410	146411
146412	146413	146414	146415	146416
146417	146418	146419	146420	146421
146422	146423	146424	146425	146426
146427	146428	146429	146430	146431
146432	146433	146434	146435	146436
146437	146438	146439	146440	146441
146442	146443	146444	146445	146446
146447	146448	146449	146450	146451
146452	146453	146454	146455	146456
146457	146458	146459	146460	146461
146462	146463	146464	146465	146466
146467	146468	146469	146470	146471
146472	146473	146474	146475	146476
146477	146478	146479	146480	146481
146482	146483	146484	146485	146486
146487	146488	146489	146490	146491
146492	146493	146494	146495	146496
146497	146498	146499	146500	146501
146502	146503	146504	146505	146506
146507	146508	146509	146510	146511
146512	146513	146514	146515	146516
146517	146518	146519	146520	146521
146522	146523	146524	146525	146526
146527	146528	146529	146530	146531
146532	146533	146534	146535	146536
146537	146538	146539	146540	146541
146542	146543	146544	146545	146546
146547	146548	146549	146550	146551
146552	146553	146554	146555	146556
146557	146558	146559	146560	146561
146562	146563	146564	146565	146566
146567	146568	146569	146570	146571
146572	146573	146574	146575	146576
146577	146578	146579	146580	146581
146582	146583	146584	146585	146586
146587	146588	146589	146590	146591
146592	146593	146594	146595	146596
146597	146598	146599	146600	146601
146602	146603	146604	146605	146606
146607	146608	146609	146610	146611
146612	146613	146614	146615	146616
146617	146618	146619	146620	146621
146622	146623	146624	146625	146626
146627	146628	146629	146630	146631
146632	146633	146634	146635	146636
146637	146638	146639	146640	146641
146642	146643	146644	146645	146646
146647	146648	146649	146650	146651
146652	146653	146654	146655	146656
146657	146658	146659	146660	146661
146662	146663	146664	146665	146666
146667	146668	146669	146670	146671
146672	146673	146674	146675	146676
146677	146678	146679	146680	146681
146682	146683	146684	146685	146686
146687	146688	146689	146690	146691
146692	146693	146694	146695	146696
146697	146698	146699	146700	146701
146702	146703	146704	146705	146706
146707	146708	146709	146710	146711
146712	146713	146714	146715	146716
146717	146718	146719	146720	146721
146722	146723	146724	146725	146726
146727	146728	146729	146730	146731
146732	146733	146734	146735	146736
146737	146738	146739	146740	146741
146742	146743	146744	146745	146746
146747	146748	146749	146750	146751
146752	146753	146754	146755	146756
146757	146758	146759	146760	146761
146762	146763	146764	146765	146766
146767	146768	146769	146770	146771
146772	146773	146774	146775	146776
146777	146778	146779	146780	146781
146782	146783	146784	146785	146786
146787	146788	146789	146790	146791
146792	146793	146794	146795	146796
146797	146798	146799	146800	146801
146802	146803	146804	146805	146806
146807	146808	146809	146810	146811
146812	146813	146814	146815	146816
146817	146818	146819	146820	146821
146822	146823	146824	146825	146826
146827	146828	146829	146830	146831
146832	146833	146834	146835	146836
146837	146838	146839	146840	146841
146842	146843	146844	146845	146846
146847	146848	146849	146850	146851
146852	146853	146854	146855	146856
146857	146858	146859	146860	146861
146862	146863	146864	146865	146866
146867	146868	146869	146870	146871
146872	146873	146874	146875	146876
146877	146878	146879	146880	146881
146882	146883	146884	146885	146886
146887	146888	146889	146890	146891
146892	146893	146894	146895	146896
146897	146898	146899	146900	146901
146902	146903	146904	146905	146906
146907	146908	146909	146910	146911
146912	146913	146914	146915	146916
146917	146918	146919	146920	146921
146922	146923	146924	146925	146926
146927	146928	146929	146930	146931
146932	146933	146934	146935	146936
146937	146938	146939	146940	146941
14				