(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2004-537254
(P2004-537254A)
(43)公表日　平成16年12月16日(2004.12.16)

(51)Int.Cl.[7]	FI		テーマコード（参考）
C12N 15/09	C12N 15/00	ZNAA	2G045
A61K 38/00	A61K 45/00		4B024
A61K 45/00	A61P 3/00		4B063
A61P 3/00	A61P 21/00		4B064
A61P 21/00	A61P 25/00		4B065

審査請求 未請求　予備審査請求 有　（全 332 頁）　最終頁に続く

(21)出願番号　特願2002-526934(P2002-526934)
(86)(22)出願日　平成13年9月14日(2001.9.14)
(85)翻訳文提出日　平成15年3月14日(2003.3.14)
(86)国際出願番号　PCT/US2001/028938
(87)国際公開番号　WO2002/022684
(87)国際公開日　平成14年3月21日(2002.3.21)
(31)優先権主張番号　60/232,685
(32)優先日　平成12年9月15日(2000.9.15)
(33)優先権主張国　米国(US)
(31)優先権主張番号　60/234,842
(32)優先日　平成12年9月22日(2000.9.22)
(33)優先権主張国　米国(US)
(31)優先権主張番号　60/236,882
(32)優先日　平成12年9月29日(2000.9.29)
(33)優先権主張国　米国(US)

(71)出願人　301005050
インサイト・ゲノミックス・インコーポレイテッド
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
(74)代理人　100089266
弁理士　大島　陽一
(72)発明者　リー、アーンスティーン・エイ
アメリカ合衆国カリフォルニア州９４７０６・アルバニー・ケインズストリート　６２４

最終頁に続く

(54)【発明の名称】輸送体及びイオンチャネル

(57)【要約】
本発明はヒト輸送体およびイオンチャネル（ＴＲＩＣＨ）、並びにＴＲＩＣＨを同定し、コードするポリヌクレオチドを提供する。本発明はまた、発現ベクター、宿主細胞、抗体、アゴニストおよびアンタゴニストをも提供する。本発明はまた、ＴＲＩＣＨの異常発現に関連する疾患を診断、治療または予防する方法をも提供する。

【特許請求の範囲】
【請求項１】
以下の（ａ）乃至（ｄ）からなる群から選択した単離されたポリペプチド。
（ａ）ＳＥＱ ＩＤ ＮＯ：１－２６（配列番号１乃至２６）からなる群から選択したアミノ酸配列を含むポリペプチド
（ｂ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列と少なくとも９０％が同一であるような天然のアミノ酸配列を含むポリペプチド
（ｃ） ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片
（ｄ） ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片
【請求項２】
ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択した請求項１に記載の単離されたポリペプチド。
【請求項３】
請求項１のポリペプチドをコードする単離されたポリヌクレオチド。
【請求項４】
請求項２のポリペプチドをコードする単離されたポリヌクレオチド。
【請求項５】
ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択した請求項４に記載の単離されたポリヌクレオチド。
【請求項６】
請求項３に記載のポリヌクレオチドに機能的に連結したプロモーター配列を含む組換えポリヌクレオチド。
【請求項７】
請求項６に記載の組換えポリヌクレオチドを用いて形質転換した細胞。
【請求項８】
請求項６に記載の組換えポリヌクレオチドを含む遺伝形質転換体。
【請求項９】
請求項１のポリペプチドを生産する方法であって、以下の過程を含む方法。
（ａ） 前記ポリペプチドの発現に好適な条件下で、請求項１のポリペプチドをコードするポリヌクレオチドに機能的に連結されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養する過程と、
（ｂ） そのように発現した前記ポリペプチドを回収する過程。
【請求項１０】
前記ポリペプチドが、ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有することを特徴とする請求項９に記載の方法。
【請求項１１】
請求項１に記載のポリペプチドと特異的に結合する単離された抗体。
【請求項１２】
以下の（ａ）乃至（ｄ）からなる群から選択した単離されたポリヌクレオチド。
（ａ） ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド
（ｂ） ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択したポリヌクレオチド配列と少なくとも９０％が同一であるような天然のポリヌクレオチド配列を含むポリヌクレオチド
（ｃ） （ａ）のポリヌクレオチドに相補的なポリヌクレオチド
（ｄ）（ｂ）のポリヌクレオチドに相補的なポリヌクレオチド
（ｅ）（ａ）～（ｄ）のＲＮＡ等価物
【請求項１３】

請求項１２に記載のポリヌクレオチドの少なくとも６０の連続したヌクレオチドを含む単離されたポリヌクレオチド。
【請求項１４】
請求項１２に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中から検出する方法であって、
（ａ）前記サンプル中の前記標的ポリヌクレオチドに相補的な配列を含む少なくとも２０の連続したヌクレオチドを含むプローブを用いて前記サンプルをハイブリダイズする過程（ただし、前記プローブと前記標的ポリヌクレオチドまたは断片の間でハイブリダイゼーション複合体が形成されるような条件下で、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）と、
（ｂ）前記ハイブリダイゼーション複合体の存在・不存在を検出し、該複合体が存在する場合にはオプションでその量を検出する過程を含むことを特徴とする方法。
【請求項１５】
前記プローブが少なくとも６０の連続したヌクレオチドを含むことを特徴とする請求項１４に記載の方法。
【請求項１６】
請求項１２に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中から検出する方法であって、
（ａ）ポリメラーゼ連鎖反応増幅を用いて前記標的ポリヌクレオチドまたはその断片を増幅する過程と、
（ｂ）前記標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含むことを特徴とする方法。
【請求項１７】
請求項１のポリペプチドと、薬剤として許容できる賦形剤とを含む組成物。
【請求項１８】
前記ポリペプチドが、ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有することを特徴とする請求項１７に記載の成分。
【請求項１９】
機能的なＴＲＩＣＨの発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項１７の組成物を投与することを含むことを特徴とする治療方法。
【請求項２０】
請求項１に記載のポリペプチドのアゴニストとして有効性を確認するために化合物をスクリーニングする方法であって、
（ａ）請求項１のポリペプチドを含むサンプルを化合物に曝露する過程と、
（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。
【請求項２１】
請求項２０に記載の方法によって同定したアゴニスト化合物と薬剤として許容できる賦形剤とを含む組成物。
【請求項２２】
機能的なＴＲＩＣＨの発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項２１の組成物を投与することを含むことを特徴とする治療方法。
【請求項２３】
請求項１に記載のポリペプチドのアンタゴニストとして有効性を確認するために化合物をスクリーニングする方法であって、
（ａ）請求項１のポリペプチドを含むサンプルを化合物に曝露する過程と、
（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。
【請求項２４】

請求項２３に記載の方法によって同定したアンタゴニスト化合物と、薬剤として許容できる賦形剤とを含む組成物。
【請求項２５】
機能的なＴＲＩＣＨの過剰な発現に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項２４の組成物を投与することを含むことを特徴とする治療方法。
【請求項２６】
請求項１に記載のポリペプチドに特異結合する化合物をスクリーニングする方法であって、
（ａ）適切な条件下で請求項１に記載のポリペプチドを少なくとも１つの試験化合物に混合させるステップと、
（ｂ）請求項１に記載のポリペプチドの試験化合物との結合を検出し、それによって請求項１に記載のポリペプチドに特異結合する化合物を同定する過程とを含むことを特徴とする方法。
【請求項２７】
請求項１に記載のポリペプチドの活性を調節する化合物をスクリーニングする方法であって、
（ａ）請求項１に記載のポリペプチドの活性が許容された条件下で、請求項１に記載のポリペプチドを少なくとも１つの試験化合物に結合させる過程と、
（ｂ）請求項１に記載のポリペプチドの活性を試験化合物の存在下で算定する過程と、
（ｃ）試験化合物の存在下での請求項１に記載のポリペプチドの活性を、試験化合物の不存在下での請求項１に記載のポリペプチドの活性と比較する過程とを含み、試験化合物の存在下での請求項１に記載のポリペプチドの活性の変化が、請求項１に記載のポリペプチドの活性を調節する化合物を標示することを特徴とする方法。
【請求項２８】
請求項５の配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、
（ａ）前記標的ポリヌクレオチドの発現に好適な条件下で、前記標的ポリヌクレオチドを含むサンプルを化合物に曝露する過程と、
（ｂ）前記標的ポリヌクレオチドの変異発現を検出する過程と、
（ｃ）可変量の前記化合物の存在下と前記化合物の不存在下で、前記標的ポリヌクレオチドの発現を比較する過程とを含むことを特徴とする方法。
【請求項２９】
試験化合物の毒性を算定する方法であって、
（ａ）核酸を含む生物学的サンプルを前記試験化合物で処理する過程と、
（ｂ）処理した前記生体サンプルの核酸と、請求項１２のポリヌクレオチドの少なくとも２０の連続するヌクレオチドを含むプローブをハイブリダイズさせるステップであって、このハイブリダイゼーションゼーションが、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われ、前記標的ポリヌクレオチドが、請求項１２のポリヌクレオチドのポリヌクレオチド配列またはその断片を含むポリヌクレオチドである、前記ステップと、
（ｃ）ハイブリダイゼーション複合体の収量を定量するステップと、
（ｄ）前記処理された生物学的サンプル中の前記ハイブリタイゼーション複合体の量を、処理されていない生物学的サンプル中の前記ハイブリタイゼーション複合体の量と比較する過程とを含み、前記処理された生物学的サンプル中の前記ハイブリタイゼーション複合体の量の差が、前記試験化合物の毒性を標示することを特徴とする方法。
【請求項３０】
生物学的サンプル中のＴＲＩＣＨの発現に関連する症状または疾患に対する診断試験法であって、
（ａ）前記抗体が前記ポリペプチドに結合し、抗体とポリペプチドとの複合体が形成され

るのに適した条件下で、前記生物学的サンプルを請求項１１に記載の抗体と結合する過程と、
（ｂ）前記複合体を検出する過程とを含み、前記複合体の存在が、前記生物学的サンプル中の前記ポリペプチドの存在と相関することを特徴とする方法。
【請求項３１】
前記抗体が、
（ａ）キメラ抗体
（ｂ）単鎖抗体
（ｃ）Ｆａｂ断片
（ｄ）$F(ab')_2$ 断片
（ｅ）ヒト化抗体 のいずれかであることを特徴とする請求項１１に記載の抗体。
【請求項３２】
請求項１１に記載の抗体と、許容できる賦形剤とを含む化合物。
【請求項３３】
被検者のＴＲＩＣＨ の発現に関連する病状又は疾患の診断方法であって、請求項３２に記載の化合物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。
【請求項３４】
前記抗体が標識されることを特徴とする請求項３２に記載の化合物。
【請求項３５】
被検者のＴＲＩＣＨ の発現に関連する病状又は疾患の診断方法であって、請求項３４に記載の化合物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。
【請求項３６】
請求項１１に記載の抗体の特異性を有するポリクローナル抗体を調製する方法であって、
（ａ）抗体反応を誘発する条件下で、ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドまたはその免疫原性断片を用いて動物を免疫化する過程と、
（ｂ）前記動物から抗体を単離する過程と、
（ｃ）前記単離された抗体をポリペプチドでスクリーニングし、それによって、ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するようなポリクローナル抗体を同定する過程とを含むことを特徴とする方法。
【請求項３７】
請求項３６に記載の方法で産出したポリクローナル抗体。
【請求項３８】
請求項３７に記載のポリクローナル抗体及び適切なキャリアを含む組成物。
【請求項３９】
請求項１１に記載の抗体の特異性を有するモノクローナル抗体を製造する方法であって、
（ａ）抗体反応を誘発する条件下で、ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドまたはその免疫原性断片を用いて動物を免疫化する過程と、
（ｂ）前記動物から抗体産出細胞を単離する過程と、
（ｃ）前記抗体産出細胞を不死化の細胞と融合して、モノクローナル抗体を産出するハイブリドーマ細胞を形成する過程と、
（ｄ）前記ハイブリドーマ細胞を培養する過程と、
（ｅ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するような前記培養モノクローナル抗体から単離する過程とを含むことを特徴とする方法。
【請求項４０】
請求項３９に記載の方法で産出したモノクローナル抗体。
【請求項４１】
請求項４０に記載のモノクローナル抗体及び適切なキャリアを含む組成物。

【請求項４２】
Ｆａｂ発現ライブラリのスクリーニングにより前記抗体を産出することを特徴とする請求項１１に記載の抗体。
【請求項４３】
組換え免疫グロブリンライブラリのスクリーニングにより前記抗体を産出することを特徴とする請求項１１に記載の抗体。
【請求項４４】
ＳＥＱ ＩＤ ＮＯ：１－２６ からなる群から選択したアミノ酸配列を有するポリペプチドをサンプル中から検出する方法であって、
（ａ）前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項１１に記載の抗体をインキュベートする過程と、
（ｂ）特異結合を検出する過程とを含み、 該特異結合が、ＳＥＱ ＩＤ ＮＯ：１－２６ からなる群から選択したアミノ酸配列を有するポリペプチドがサンプル中に存在することを標示することを特徴とする方法。
【請求項４５】
ＳＥＱ ＩＤ ＮＯ：１－２６ からなる群から選択したアミノ酸配列を有するポリペプチドを精製する方法であって、
（ａ）前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項１１に記載の抗体をインキュベートする過程と、
（ｂ）前記サンプルから前記抗体を分離し、ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有する精製ポリペプチドを得る過程とを含むことを特徴とする方法。
４６． 少なくとも一つのエレメントが請求項１３のポリヌクレオチドであるようなマイクロアレイ。
【請求項４６】
ポリヌクレオチドを含むサンプルの転写イメージを作製する方法であって、
（ａ）サンプルのポリヌクレオチドを標識する過程と、
（ｂ）ハイブリダイゼーション複合体の形成に適した条件下で請求項４６のマイクロアレイのエレメントを標識されたポリヌクレオチドに接触させる過程と、
（ｃ）サンプル内のポリヌクレオチドの発現を定量化する過程からなる方法。
【請求項４７】
固体基板上の異なる物理的位置に固定された異なるヌクレオチド分子を含むアレイであって、該ヌクレオチド分子が標的ポリヌクレオチドの少なくとも３０の連続したヌクレオチドと特異的にハイブリダイズする最初のオリゴヌクレオチドまたはポリヌクレオチドを含み、該標的ポリヌクレオチドが請求項１２のポリヌクレオチドであることを特徴とするアレイ。
【請求項４８】
前記の最初のオリゴヌクレオチドまたはポリヌクレオチド配列が該標的ポリヌクレオチドの少なくとも３０の連続するヌクレオチドと完全に相補的であることを特徴とする、請求項４８に記載のアレイ。
【請求項４９】
前記の最初のオリゴヌクレオチドまたはポリヌクレオチド配列が該標的ポリヌクレオチドの少なくとも６０の連続するヌクレオチドと完全に相補的であることを特徴とする、請求項４８に記載のアレイ。
【請求項５０】
前記の最初のオリゴヌクレオチドまたはポリヌクレオチド配列が該標的ポリヌクレオチドと完全に相補的であることを特徴とする、請求項４８に記載のアレイ。
【請求項５１】
マイクロアレイであることを特徴とする、請求項４８に記載のアレイ。
【請求項５２】

前記の最初のオリゴヌクレオチドまたはポリヌクレオチド配列を含むヌクレオチド分子にハイブリダイズされた該標的ポリヌクレオチドをさらに含むことを特徴とする請求項４８に記載のアレイ。
【請求項５３】
リンカーが該ヌクレオチド分子のうちの少なくとも一つを該固体基板に結合することを特徴とする請求項４８に記載のアレイ。
【請求項５４】
基板上の異なる各物理的位置が複数のヌクレオチド分子を含み、任意の物理的位置にあるその複数のヌクレオチド分子が同じ配列を含み、基板上の異なる各物理的位置が基板上の他の物理的位置におけるヌクレオチド分子の配列と異なる配列を有することを特徴とする、請求項４８に記載のアレイ。
【請求項５５】
ＳＥＱ ＩＤ ＮＯ：１のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項５６】
ＳＥＱ ＩＤ ＮＯ：２のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項５７】
ＳＥＱ ＩＤ ＮＯ：３のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項５８】
ＳＥＱ ＩＤ ＮＯ：４のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項５９】
ＳＥＱ ＩＤ ＮＯ：５のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項６０】
ＳＥＱ ＩＤ ＮＯ：６のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項６１】
ＳＥＱ ＩＤ ＮＯ：７のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項６２】
ＳＥＱ ＩＤ ＮＯ：８のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項６３】
ＳＥＱ ＩＤ ＮＯ：９のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項６４】
ＳＥＱ ＩＤ ＮＯ：１０のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項６５】
ＳＥＱ ＩＤ ＮＯ：１１のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項６６】
ＳＥＱ ＩＤ ＮＯ：１２のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項６７】
ＳＥＱ ＩＤ ＮＯ：１３のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項６８】
ＳＥＱ ＩＤ ＮＯ：１４のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項６９】
ＳＥＱ ＩＤ ＮＯ：１５のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項７０】
ＳＥＱ ＩＤ ＮＯ：１６のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項７１】
ＳＥＱ ＩＤ ＮＯ：１７のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項７２】
ＳＥＱ ＩＤ ＮＯ：１８のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項７３】
ＳＥＱ ＩＤ ＮＯ：１９のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項７４】

ＳＥＱ ＩＤ ＮＯ：２０のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項７５】
ＳＥＱ ＩＤ ＮＯ：２１のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項７６】
ＳＥＱ ＩＤ ＮＯ：２２のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項７７】
ＳＥＱ ＩＤ ＮＯ：２３のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項７８】
ＳＥＱ ＩＤ ＮＯ：２４のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項７９】
ＳＥＱ ＩＤ ＮＯ：２５のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項８０】
ＳＥＱ ＩＤ ＮＯ：２６のアミノ酸配列を含む請求項１に記載のポリペプチド。
【請求項８１】
ＳＥＱ ＩＤ ＮＯ：２７のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項８２】
ＳＥＱ ＩＤ ＮＯ：２８のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項８３】
ＳＥＱ ＩＤ ＮＯ：２９のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項８４】
ＳＥＱ ＩＤ ＮＯ：３０のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項８５】
ＳＥＱ ＩＤ ＮＯ：３１のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項８６】
ＳＥＱ ＩＤ ＮＯ：３２のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項８７】
ＳＥＱ ＩＤ ＮＯ：３３のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項８８】
ＳＥＱ ＩＤ ＮＯ：３４のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項８９】
ＳＥＱ ＩＤ ＮＯ：３５のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項９０】
ＳＥＱ ＩＤ ＮＯ：３６のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項９１】
ＳＥＱ ＩＤ ＮＯ：３７のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項９２】
ＳＥＱ ＩＤ ＮＯ：３８のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項９３】

ＳＥＱ ＩＤ ＮＯ：３９のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項９４】
ＳＥＱ ＩＤ ＮＯ：４０のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項９５】
ＳＥＱ ＩＤ ＮＯ：４１のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項９６】
ＳＥＱ ＩＤ ＮＯ：４２のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項９７】
ＳＥＱ ＩＤ ＮＯ：４３のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項９８】
ＳＥＱ ＩＤ ＮＯ：４４のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項９９】
ＳＥＱ ＩＤ ＮＯ：４５のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項１００】
ＳＥＱ ＩＤ ＮＯ：４６のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項１０１】
ＳＥＱ ＩＤ ＮＯ：４７のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項１０２】
ＳＥＱ ＩＤ ＮＯ：４８のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項１０３】
ＳＥＱ ＩＤ ＮＯ：４９のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項１０４】
ＳＥＱ ＩＤ ＮＯ：５０のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項１０５】
ＳＥＱ ＩＤ ＮＯ：５１のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【請求項１０６】
ＳＥＱ ＩＤ ＮＯ：５２のポリヌクレオチド配列を含む請求項１２に記載のポリヌクレオチド。
【発明の詳細な説明】
【０００１】
（技術分野）
本発明は、輸送体及びイオンチャネルの核酸配列及びアミノ酸配列に関する。 本発明はまた、これらの配列を利用した輸送障害、神経の疾患、筋疾患、及び免疫疾患および細胞の異常増殖の診断・治療・予防に関する。 本発明はさらに、輸送体及びイオンチャネルの核酸配列及びアミノ酸配列の発現における外来性化合物の効果についての評価に関する。
【０００２】

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

【０００３】
膜の一方の側から他方の側へ単一溶質を輸送する担体タンパク質をユニポーターと呼ぶ。対照的に、共役輸送体は、１つの溶質の移動を第２溶質の移動に連結させ、その移動は同時または順次に起こるかのいずれかであり、、移動の方向は同一方向（シンポート）または逆方向（アンチポート）のいずれかである。例えば、腸管及び腎臓の上皮は、原形質膜に存在するナトリウム勾配によって駆動される種々のシンポーター系を含む。ナトリウムは、電気化学勾配に従って細胞内に移動し、その移動と共に溶質を細胞内に運ぶ。溶質を取り込むための駆動力を供給するナトリウム勾配は、遍在性のNa^+／K^+ ＡＴＰアーゼ系によって維持される。ナトリウム共役輸送体には、哺乳動物グルコース輸送体（ＳＧＬＴ１）、ヨウ化物輸送体（ＮＩＳ）及び多ビタミン輸送体（ＳＭＶＴ）がある。３つの輸送体はすべて、１２個の推定上の膜貫通セグメント、細胞外グリコシル化部位、並びに細胞質内側にあるＮ末端及びＣ末端を有する。ＮＩＳは、放射性ヨウ素甲状腺イメージング技術及び甲状腺への放射性同位元素の特定のターゲッ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．ＵＳＡ９４：５５６８－５５７３）。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などの水溶性ビタミンの輸送に関与していると考えられる（Ｐｒａｓａｄ，Ｐ．Ｄ．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、硝酸、ヌクレオチド、モノカルボン酸及び薬剤のための輸送体を含む。真核生物に見られるＭＦＳ輸送体はすべて１２個の膜貫通セグメントを含む構成となっている（Ｐａｏ，Ｓ．Ｓ．ら（１９９８）Ｍｉｃｒｏｂｉｏｌ．Ｍｏｌｅｃ．Ｂｉｏｌ．Ｒｅｖ．６２：１３４）。ＭＦＳ輸送体の最大のファミリーは糖輸送体ファミリーであり、ヒトに見られるような、グルコース及びその他の六炭糖の輸送に必要な７つのグルコース輸送体（ＧＬＵＴ１～ＧＬＵＴ７）を含む。これらのグルコース輸送タンパク質は、独自の組織分布及び生理学的機能を有する。ＧＬＵＴ１は、基底板でのグルコース要求を有する多くの細胞型に提供して、上皮及び内皮の障壁を通ってグルコースを輸送し、ＧＬＵＴ２は、グルコースの取り込みまたは肝臓からの流出を促進し、ＧＬＵＴ３は、ニューロンへのグルコースの供給を調整し、ＧＬＵＴ４は、インシュリンによって調整されるグルコース分解に関与し、ＧＬＵＴ５は、骨格筋へのフルクトースの取り込みを調節する。グルコ

ース輸送体の欠陥は、最近同定された神経症候群であって、糖原貯蔵症、ファンコニービッケル症候群、及び非インシュリン依存糖尿病のみならず、乳児てんかん及び発達障害の原因となるような神経症候群に関与している（Ｍｕｅｃｋｌｅｒ，Ｍ．（１９９４）Ｅｕｒ．Ｊ．Ｂｉｏｃｈｅｍ．２１９：７１３－７２５、Ｌｏｎｇｏ，Ｎ．およびＬ．Ｊ．Ｅｌｓａｓ（１９９８）Ａｄｖ．Ｐｅｄｉａｔｒ．４５：２９３－３１３）。

【０００５】

モノカルボン酸アニオン輸送体は、Ｌ－乳酸、ピルビン酸及びケトン体アセテート、アセトアセテート及びβヒドロキシ酪酸を含む広範な基質特異性を有するプロトン共役シンポーターである。これまでに少なくとも７つのアイソフォームが同定されている。アイソフォームは、ＴＭ６とＴＭ７との間に大きな細胞内ループを有する１２個の膜貫通（ＴＭ）らせん状ドメインを有し、解糖中に乳酸と共に化学量論的に生成されるプロトンを除去することにより、細胞内ｐＨを維持する際に重要な役割を果たすと推定されている。最も良く特徴付けられたH^+ モノカルボン酸輸送体は、Ｌ－乳酸及び広範囲の他の脂肪族モノカルボン酸を輸送する赤血球膜の輸送体である。その他の細胞は、異なる基質及び抑制因子選択性を有するH^+ 共役モノカルボン酸輸送体を含む。特に心筋及び腫瘍細胞は、或る基質に対するＫｍ値（Ｄ－乳酸よりＬ－乳酸を選択する立体化学的選択性を含む）及び抑制因子に対する感受性が異なる輸送体を有する。腸管及び肝臓上皮の管腔表面にNa^+ モノカルボン酸共輸送体があり、それによってこれらの組織における乳酸、ピルビン酸及びケトン体の取り込みが可能になる。更に、肝臓、腸及び肺を含む器官に、有機カチオン及び有機アニオンのための特異的且つ選択的な輸送体がある。有機アニオン輸送体は、電子誘引側基を有する疎水性、荷電分子に選択的である。アンモニウム輸送体などの有機カチオン輸送体は、種々の薬剤及び内因性代謝の分泌を媒介し、細胞内ｐＨの維持に寄与する（Ｐｏｏｌｅ，Ｒ．Ｃ．およびＡ．Ｐ．Ｈａｌｅｓｔｒａｐ（１９９３）Ａｍ．Ｊ．Ｐｈｙｓｉｏｌ．２６４：Ｃ７６１－Ｃ７８２、Ｐｒｉｃｅ，Ｎ．Ｔ．ら（１９９８）Ｂｉｏｃｈｅｍ．Ｊ．３２９：３２１－３２８、Ｍａｒｔｉｎｅｌｌｅ，Ｋ．および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、嚢胞性繊維症膜通過伝導制御因子（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１つのＭＳＤが含まれる。これらの「半分子」は、小胞体での主要組織適合複合体（ＭＨＣ）ペプチド輸送系であるＴａｐ１及びＴａｐ２などのホモ及びヘテロ二量体を形成する。幾つかの遺伝病は、ＡＢＣ輸送体の欠陥に起因する。疾患及びそれに対応するタンパク質の例としては、嚢胞性繊維症（ＣＦＴＲ、イオンチャネル）、副腎白質ジストロフィー（副腎白質ジストロフィータンパク質、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．および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補助因子として重要である。例えば、銅は、スーパーオキシドジスムターゼ、フェロキシダーゼ（セルロプラスミン）及びリシルオキシダーゼなどのオキシドレダクターゼにおいて補助因子として作用することにより、ヘモグロビン合成、結合組織代謝及び骨格の発達に関与する。 銅及びその他の金属イオンは、食餌中に与えなければならないものであり、消化管で輸送体に吸収される。血漿タンパク質は、金属イオンを肝臓及びその他の標的器官へ輸送し、ここで特異輸送体が必要に応じてイオンを細胞及び細胞のオルガネラに移動させる。金属イオンでの代謝の不均衡は、多数の疾患状態に関連してきた（Ｄａｎｋｓ，Ｄ．Ｍ．（１９８６）Ｊ．Ｍｅｄ．Ｇｅｎｅｔ．２３：９９－１０６）。

【０００８】

原形質膜を通過する脂肪酸の輸送は、高容量で低親和性プロセスである拡散により起こる。しかしながら、通常の生理学的条件下では、脂肪酸輸送の内の有意の割合が高親和性、低容量のタンパク質媒介輸送プロセスによって起こっているように見える。脂肪酸輸送タンパク質（ＦＡＴＰ）は、４つの膜貫通セグメントを有する膜内在性タンパク質であり、筋肉、心臓及び脂肪などの、高レベルの原形質膜脂肪酸流入を示す組織において発現される。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、コハク酸、フマル酸及びリン酸を輸送する）、トリカルボン酸担体（クエン酸及びリンゴ酸を輸送する）、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）Ｂｉｏｃｈｅｍｉｓｔｒｖ，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、またミトコンドリア脱共役タンパク質も含まれる。この脱共役タンパク質はミトコンドリア内膜でプロトンリークを起こさせ、酸化的リン酸化をＡＴＰ合成から脱共役させる。結果として、熱の形態でのエネルギー散逸が起こる。ミトコンドリア脱共役タンパク質は、温度調節及び代謝率のモジュレーターとして意味づけられ、肥満症などの代謝病に対する薬剤の潜在的標的として提案されてきた（Ｒｉｃｑｕｉｅｒ，Ｄ．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起こり、かつ維持される。イオンの動きは、膜内にイオン選択性ポアを形成するようなイオンチャネルを必要とする。イオンチャネルには２つの基本型があり、それはイオン輸送体とイオンチャネルがある。イオン輸送体は、ＡＴＰ加水分解から得られるエネルギーを利用して、イオンの濃度勾配に逆らってイオンを能動的に輸送する。イオンチャネルは、制限された条件下でイオンの電気化学勾配に従ってイオンの受動的流出を起こさせる。同時に、これらのタイプのイオンチャネルは、１）神経細胞の軸索に沿った電気インパルス伝導、２）濃度勾配に逆らった細胞内への分子の輸送、３）筋収縮の開始、及び４）内分泌細胞の分泌に用

いられるような電気化学勾配を発生させ、維持及び利用する。

【００１２】

イオン輸送体

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

【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きな細胞質内領域を有する１００ｋＤ 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加水分解に関与する周辺複合体であり、V_2 ドメインは、プロトンの膜通過移動に関与する、膜内在性の複合体である。Ｆ－ＡＴＰ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、１２個のコピーのｃサブユニットを含んでおり、各サブユニットは高度に疎水性のタンパク質で、２つの膜貫通ドメインから成り、プロトン輸送に不可欠なＴＭ２に１つの埋没されたカルボキシル基を含む。Ｖ－ＡＴＰ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サブユニットも含む（Ｆｏｒｇａｃ，Ｍ．（１９９９）Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２７４：１２９５１－１２９５４）。

【００１４】

細胞の静止電位は、担体タンパク質及びイオンチャネルに関連する多くのプロセスに利用される。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（シンポート）にカップルされ、Na^+ の電気化学勾配に従った動きが他の分子を濃度勾配に逆らって輸送させる。同様にして、心筋は、細胞からのCa^{2+} の移動を細胞内へのNa^+ の輸送にカップルさせる（アンチポート）。

【００１５】

イオンチャネル

イオンチャネルは、ポアの開閉を調節することによってイオンの流れを制御する。種々のゲーティング機構によってイオン流速を制御する能力は、ニューロン及び内分泌シグナル伝達、筋収縮、受精並びに、イオン及びｐＨバランスの調整などの多様なシグナル伝達及び恒常性機能をイオンチャネルに媒介させ得る。イオンチャネルは、開口機能の制御方法に応じて区分される。機械依存性チャネルは、機械的応力に応じてポアを開け、電位依存性チャネル（例えばNa^+、K^+、Ca^{2+} 及びCl^- チャネル）は膜電位に応じてポアを開け、リガンド依存性チャネル（例えばアセチルコリン依存性、セロトニン依存性、及びグルタミン依存性カチオンチャネル、並びにＧＡＢＡ依存性及びグリシン依存性クロライドチャネル）は、特定のイオン、ヌクレオチドまたは神経伝達物質の存在下でポアを開ける。特定のイオンチャネルの開口特性（即ち開閉に対するその閾値及び持続時間）は、補助チャネルタンパク質及び／またはリン酸化などの翻訳後修飾と関連して時折調節さ

れる。

【００１６】

機械依存性または機械刺激イオンチャネルは、触覚、聴覚及び平衡感覚のトランスデューサとして作用し、細胞容積調整、平滑筋収縮及び心リズム生成においても重要な役割を果たす。機械受容不活性化チャネル（ＳＩＣ）は、最近ラットの腎臓からクローニングされた。ＳＩＣチャネルは、細胞膜への圧力またはストレスによって活性化され、Ca^{2+}及びNa^{+}を共に伝導するようなあるグループのチャネルに属している（Ｓｕｚｕｋｉ，Ｍ．ら（１９９９）Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２７４：６３３０－６３３５）。

【００１７】

電位依存性カチオンチャネルのポア形成サブユニットは、イオンチャネルタンパク質のスーパーファミリーを形成する。チャネルタンパク質の特徴的なドメインには、６つの膜貫通ドメイン（Ｓ１～Ｓ６）、Ｓ５とＳ６との間に配置されたポア形成領域（Ｐ）及び、細胞内にあるアミノ末端及びカルボキシ末端がある。Na^{+}及びCa^{2+}サブファミリーでは、このドメインは４回繰り返され、K^{+}チャネルサブファミリーでは、各チャネルは、同一または非類似のいずれかであるサブユニットの四量体から形成される。Ｐ領域は、チャネルに対するイオン選択性を特定する情報を含む。K^{+}チャネルの場合、ＧＹＧトリペプチドがこの選択性に関与する（Ｉｓｈｉｉ，Ｔ．Ｍ．ら（１９９７）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ９４：１１６５１－１１６５６）。

【００１８】

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

【００１９】

電位依存性Na^{+}チャネルは、２つの更に小さな補助サブユニットβ１及びβ２に関連する２６０ｋＤａのポア形成αサブユニットから構成されるようなヘテロ三量体の複合体である。β２サブユニットは、細胞外Ｉｇドメインを含む複合膜糖タンパク質であり、そのα及びβ１サブユニットとの関連は、チャネルの機能的発現の増加、そのゲーティング特性の変化、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）Ｃｅｌｌ８３：４３３－４４２）。

【００２０】

非電位依存性Na^{+}チャネルには、アミロライド感受性Na^{+}チャネル／ｄｅｇｅｎｅｒｉｎ（ＮａＣ／ＤＥＧ）ファミリーのメンバーが含まれる。このファミリーのチャネルサブユニットは、アミノ末端及びカルボキシル末端が細胞内にあり、長い細胞外ループに隣接する２つの膜貫通ドメインを含むと考えられる。ＮａＣ／ＤＥＧファミリーには、気道、遠位結腸、腎臓の皮質集合管及び外分泌腺導管を含む上皮におけるNa^{+}再吸収に関与する上皮性Na^{+}チャネル（ＥｎａＣ）がある。ＥｎａＣにおける突然変異は、Ｉ型偽性低アルドステロン症及びリドル症候群（偽性高アルドステロン症）を引き起こす。ＮａＣ／ＤＥＧファミリーには、最近特徴付けられたH^{+}依存性カチオンチャネルまたは酸感受性イオンチャネル（ＡＳＩＣ）もある。ＡＳＩＣサブユニットは、脳で発現し、ヘテロ多量体のNa^{+}浸透性チャネルを形成する。これらのチャネルは、活性化のため

に酸性ｐＨ変動を必要とする。ＡＳＩＣサブユニットは、元々は線虫から単離された機械依存性チャネルのファミリーであるｄｅｇｅｎｅｒｉｎとの相同性を示す。Ｄｅｇｅｎｅｒｉｎでの突然変異は、神経変性を引き起こす。組織アシドーシスが痛みを発生させるものであるので、ＡＳＩＣサブユニットは、神経機能または痛みの知覚においても役割を有し得る（Ｗａｌｄｍａｎｎ，Ｒ．およびＭ．Ｌａｚｄｕｎｓｋｉ（１９９８）Ｃｕｒｒ．Ｏｐｉｎ．Ｎｅｕｒｏｂｉｏｌ．８：４１８－４２４；Ｅｇｌｅｎ，Ｒ．Ｍ．ら（１９９９）Ｔｒｅｎｄｓ Ｐｈａｒｍａｃｏｌ．Ｓｃｉ．２０：３３７－３４２）。

【００２１】

K^+ チャネルはすべての細胞種に存在し、電圧、ＡＴＰ濃度、またはCa^{2+}及びｃＡＭＰなどのセカンドメッセンジャーによってK^+ チャネルを制御し得る。非興奮性組織では、K^+ チャネルが、タンパク質合成、内分泌性分泌の制御及び膜内外の浸透平衡の維持に関与している。ニューロン及びその他の興奮性細胞では、活動電位の調整及び膜の再分極に加えて、K^+ チャネルが静止膜電位の設定に関与する。サイトゾルは、非拡散性アニオンを含み、この負の電荷総ての平衡を保持するために、細胞にはNa^+、K^+及びCl^- の再分布を提供するNa^+／K^+ ポンプ及びイオンチャネルが含まれる。ポンプは、能動的にNa^+ を細胞へ輸送し、K^+ を細胞内に輸送する（Na^+：K^+ の割合は３：２である）。原形質膜のイオンチャネルは、K^+ 及びCl^- が受動的な拡散により流出入することを許容する。サイトゾル内の電荷が高度に負であるので、Cl^- は細胞外に流出する。K^+ の流れは、K^+ を細胞内に引き込む起電力及びK^+ を細胞外に押し出す濃度勾配によって均衡が保たれる。従って、静止膜電位は、主としてK^+ の流れによって調整される（Ｓａｌｋｏｆｆ，Ｌ．およびＴ．Ｊｅｇｌａ（１９９５）Ｎｅｕｒｏｎ １５：４８９－４９２）。

【００２２】

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

【００２３】

K^+ チャネルの第２のスーパーファミリーは、内向き整流型チャネル（Ｋｉｒ）で構成される。Ｋｉｒチャネルは、優先的にK^+ の流れを内向きに導く特性を有する。これらのタンパク質は、１つのカリウム選択性ポアドメイン及び２つの膜貫通ドメインから構成され、これらは電位依存性Ｋチャネルの第５及び第６膜貫通ドメインに対応する。Ｋｉｒサブユニットは、四量体としても会合している。Ｋｉｒファミリーには、腎尿細管疾患のバ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。α_1 サブユニットは、膜ポア及び電圧センサを形成し、$\alpha_2\delta$ 及びβサブユニットは、チャネルの電圧依存、ゲーティング特性及び電流振幅を調節する。これらのサブユニットは、少なくとも６つのα_1、１つの$\alpha_2\delta$、４つのβ遺伝子によってコードされる。第４サブユニットγは、骨格筋において同定されてきた（Ｗａｌｋｅｒ，Ｄ．ら（１９９８）Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２７３：２３６１－２３６７、ＭｃＣｌｅｓｋｅｙ，Ｅ．Ｗ．（１９９４）Ｃｕｒｒ．Ｏｐｉｎ．Ｎｅｕｒｏｂｉｏｌ．４：３０４－３１２）。

【００２６】
カルシウムイオンチャネルの過渡的な受容体ファミリ（Ｔｒｐ）は容量性カルシウム流入（ＣＣＥ）を媒介すると考えられている。ＣＣＥは、イノシトール三リン酸（ＩＰ３）や多くのホルモンや成長因子に応答する他の物質の作用によって消耗されるＣａ$^{2+}$ 貯蔵を再供給するためのＣａ$^{2+}$ の細胞内への流入である。ＴｒｐおよびＴｒｐ様物質は、最初にショウジョウバエからクローンされ、Ｓ３からＳ６領域においてＣａ２＋チャネルにゲート開閉をする電位に類似性を持つ。これはＴｒｐまたはその関連タンパク質が、哺乳動物のＣＣＣ流入チャネルを形成することを示唆する（Ｚｈｕ，Ｘ．ら（１９９６）Ｃｅｌｌ８５：６６１－６７１；Ｂｏｕｌａｙ，Ｇ．ら（１９９７）Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２７２：２９６７２－２９６８０）．。メラスタチン（ｍｅｌａｓｔａｔｉｎ）はマウスおよびヒトの両方において単離される遺伝子であり、黒色腫細胞におけるその発現は、生体内における黒色腫の侵襲性と逆比例している。このヒトｃＤＮＡ転写物はＴｒｐファミリのメンバーに相同体を有する１５３３基のアミノ酸をもつタンパク質に相当する。メラスタチンｍＲＮＡ発現状態と腫瘍の厚さを組み合わせて利用すると、転移性疾患の発症に対するリスクの低いグループと高いグループに患者のグループ分けを決定できる可能性が提示されている（Ｄｕｎｃａｎ，Ｌ．Ｍ．ら（２００１）Ｊ．Ｃｌｉｎ．Ｏｎｃｏｌ．１９：５６８－５７６）．。

【００２７】
クロライドチャネルは、内分泌性分泌及びサイトゾルオルガネラのｐＨ調整に必要である。分泌性上皮細胞では、Ｃｌ$^{-}$ はＮａ$^{+}$、Ｋ$^{+}$／Ｃｌ$^{-}$ 共輸送体によって側底膜を通過して細胞に入り、電気化学的平衡濃度以上で細胞内に蓄積する。ホルモンの刺激に応じて、尖端表面からのＣｌ$^{-}$ を分泌することにより、Ｎａ$^{+}$ 及び水が分泌性ルーメンに流入する。嚢胞性線維症膜通過伝導制御因子（ＣＦＴＲ）は、ヒトによく見られる致命的な遺伝病である嚢胞性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、６つの膜貫通ドメインを各々有するような２つのドメインとその後に続くヌクレオチド結合部位から構成される。ＣＦＴＲ機能の損失は、経上皮水分泌を減少させ、結果的に、呼吸樹、膵管及び腸を被膜する粘液の層は、脱水されて、取り除くのが困難になる。結果的にこれらの部位が閉塞することにより、膵不全、「胎便性イレウス」及び悲惨な「慢性閉塞性肺疾患」を生じさせる（Ａｌ－Ａｗｑａｔｉ，Ｑ．ら（１９９２）Ｊ．Ｅｘｐ．Ｂｉｏｌ．１７２：２４５－２６６）。

【００２８】
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。骨格筋で主に発現されるよ

うなＣＬＣ－１の突然変異は、常染色体劣性全身性ミオトニー及び常染色体優性先天性ミオトニーの原因であり、腎臓チャネルＣＬＣ－５における突然変異は、腎臓結石を引き起こす（Ｊｅｎｔｓｃｈ，Ｔ．Ｊ．（１９９６）Ｃｕｒｒ．Ｏｐｉｎ．Ｎｅｕｒｏｂｉｏｌ．６：３０３－３１０）。

【００２９】

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、γアミノ酪酸（ＧＡＢＡ）及びグリシンなどの抑制神経伝達物質に応じて開き、膜の過分極及びそれに続く活動電位の生成を引き起こす。神経伝達物質依存性イオンチャネルは、４つの膜貫通ドメインを有し、おそらく五量体として機能する（前出のＪｅｎｔｓｃｈ）。第２膜貫通ドメインのアミノ酸は、チャネル浸透及び選択性の決定において重要であるように見える（Ｓａｔｈｅｒ，Ｗ．Ａ．ら（１９９４）Ｃｕｒｒ．Ｏｐｉｎ．Ｎｅｕｒｏｂｉｏｌ．４：３１３－３２３）。

【００３０】

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

【００３１】

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

【００３２】

別のタイプのイオンチャネルタンパク質の活性は、種々の細胞内シグナル伝達タンパク質によっても調節し得る。多くのチャネルは、１個若しくは数個のタンパク質キナーゼによるリン酸化のための部位を有する。 タンパク質キナーゼには、タンパク質キナーゼＡ、

タンパク質キナーゼＣ、チロシンキナーゼ及びカゼインキナーゼＩＩがあり、これらはすべて細胞におけるイオンチャネル活性を調節する。Ｋｉｒチャネルは、ヘテロ三量体のＧタンパク質のＧｐｙサブユニットの結合により活性化される（Ｒｅｉｍａｎｎ，Ｆ．およびＦ．Ｍ．Ａｓｈｃｒｏｆｔ（１９９９）Ｃｕｒｒ．Ｏｐｉｎ．Ｃｅｌｌ．Ｂｉｏｌ．１１：５０３－５０８）。別のタンパク質は、細胞膜の特定部位へのイオンチャネルの局在化に関与している。このようなタンパク質にはＰＤＺドメインタンパク質があり、これはニューロンのシナプスでイオンチャネルのクラスター形成を調整するＭＡＧＵＫ（膜関連グアニル酸キナーゼ）として知られている（Ｃｒａｖｅｎ，Ｓ．Ｅ．およびＤ．Ｓ．Ｂｒｅｄｔ（１９９８）Ｃｅｌｌ９３：４９５－４９８）。

疾患との相関関係

無数のヒトの疾患及び障害の原因は、膜を通過する分子の輸送における欠陥に起因し得る。膜結合輸送体及びイオンチャネルの輸送における欠陥は、嚢胞性繊維症、グルコース－ガラクトース吸収不良症候群、高コレステロール血症、フォンギルケ病及び或る種の糖尿病など幾つかの疾患に関連している。膜を通過して小分子を輸送できないことに起因する単一の遺伝子欠陥疾患には、例えばシスチン尿症、イミノグリシン尿症、Ｈａｒｔｕｐ病及びファンコーニ病がある（ｖａｎ’ｔ Ｈｏｆｆ，Ｗ．Ｇ．（１９９６）Ｅｘｐ．Ｎｅｐｈｒｏｌ．４：２５３－２６２、Ｔａｌｅｎｔｅ，Ｇ．Ｍ．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、筋形質カルシウム放出チャネル、Ｔ細管カルシウムチャネル及び筋肉ナトリウムチャネルにおける突然変異は、悪性高熱症を引き起こす。ＱＴ延長症候群及び特発性心室細動などの心不整脈疾患は、カリウム及びナトリウムチャネルにおける突然変異に起因する（Ｃｏｏｐｅｒ，Ｅ．Ｃ．およびＬ．Ｙ．Ｊａｎ（１９９８）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ９６：４７５９－４７６６）。既知のヒト真性てんかん遺伝子は、４つがすべてイオンチャネルタンパク質をコードする（Ｂｅｒｋｏｖｉｃ，Ｓ．Ｆ．およびＩ．Ｅ．Ｓｃｈｅｆｆｅｒ（１９９９）Ｃｕｒｒ．Ｏｐｉｎ．Ｎｅｕｒｏｌｏｇｙ１２：１７７１８２）。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）。

【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．およびＬ．Ｓ．Ｎａｒａｓｉｍｈａｎ（１９９７）Ａｄｖ．Ｐｈａｒｍａｃｏｌ．３９：４７－９８）。種々のクラスのイオンチャネルは、痛みの知覚においても重要な役割を果たすので、新たな鎮痛薬の潜在的標的である。このようなイオンチャネルにはバニロイド（ｖａｎｉｌｌｏｉｄ）依存性イオンチャネルがあり、これは有害熱によってのみならずバニロイドカプサイシンによっても活性化される。電位依存性Na^+チャネルを遮断するリドカインやメキシレチンなどの局部麻酔薬は、神経障害性の痛みの治療において有用であった（前出のＥｇｌｅｎ）。

【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Ｋｖ１．３のペプチドアンタゴニストは、ブタにおける遅延型過敏症及び異質遺伝子型反応を抑制し、安全で効果的な免疫抑制薬としてのチャネル遮断薬の使用のアイデアを実証している（Ｃａｌａｈａｎ，Ｍ．Ｄ．ａｎｄ Ｋ．Ｇ．Ｃｈａｎｄｙ（１９９７）Ｃｕｒｒ．Ｏｐｉｎ．Ｂｉｏｔｅｃｈｎｏｌ．８：７４９－７５６）。

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

【００３７】

（発明の概要）

本発明は、総称して「ＴＲＩＣＨ」、個別にはそれぞれ「ＴＲＩＣＨ－１」、「ＴＲＩＣＨ－２」、「ＴＲＩＣＨ－３」、「ＴＲＩＣＨ－４」、「ＴＲＩＣＨ－５」、「ＴＲＩＣＨ－６」、「ＴＲＩＣＨ－７」、「ＴＲＩＣＨ－８」、「ＴＲＩＣＨ－９」、「ＴＲＩＣＨ－１０」、「ＴＲＩＣＨ－１１」、「ＴＲＩＣＨ－１２」、「ＴＲＩＣＨ－１３」、「ＴＲＩＣＨ－１４」、「ＴＲＩＣＨ－１５」、「ＴＲＩＣＨ－１６」、「ＴＲＩＣＨ－１７」、「ＴＲＩＣＨ－１８」、「ＴＲＩＣＨ－１９」、「ＴＲＩＣＨ－２０」、「ＴＲＩＣＨ－２１」、「ＴＲＩＣＨ－２２」、「ＴＲＩＣＨ－２３」、「ＴＲＩＣＨ－２４」、「ＴＲＩＣＨ－２５」、および「ＴＲＩＣＨ－２６」と呼ぶ輸送体及びイオンチャネルである精製されたポリペプチドを提供する。或る実施態様において本発明は、（ａ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列からなるポリペプチド、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列と少なくとも９０％が同一である天然のアミノ酸配列を含むポリペプチド、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を含む群から選択した単離されたポリペプチドを提供する。 一実施態様では、ＳＥＱ ＩＤ ＮＯ：１－２６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。

【００３８】

また、本発明は（ａ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列と少なくとも９０％の相同性を有する天然のアミノ酸配列を含むポリペプチド、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列の生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列の免疫原性断片を含む群から選択されたポリペプチドをコードするような単離されたポリヌクレオチドを提供する。 一実施態様では、ポリヌクレオチドはＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したポリペプチドをコードする。 別の実施態様では、ポリヌクレオチドはＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択される。

【００３９】

本発明は更に、（ａ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を含むポリペプチド、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列と少なくとも９０％の相同性のある天然のアミノ酸配列を含むポリペプチド、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－２６を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択し

たアミノ酸配列を有するポリペプチドの免疫原性断片を有する群から選択したポリペプチドをコードするようなポリヌクレオチドと機能的に連結したプロモーター配列を有する組換えポリヌクレオチドを提供する。 一実施態様では、本発明は組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様では、本発明はこの組換えポリヌクレオチドを含む遺伝形質転換体を提供する。

【００４０】

また、本発明は、（ａ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチド、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列と少なくとも９０％が同一である天然のアミノ酸配列を含むポリペプチド、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および（ｄ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群から選択されたポリペプチドを製造する方法を提供する。 製造方法は、（ａ）組換えポリヌクレオチドを用いて形質転換した細胞をポリペプチドの発現に適した条件下で培養する過程と、（ｂ）そのように発現したポリペプチドを回収する過程とを有し、組換えポリヌクレオチドはポリペプチドをコードするポリヌクレオチドに機能的に連結したプロモーター配列を有する。

【００４１】

本発明は更に、（ａ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチド、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列と少なくとも９０％同一である天然のアミノ酸配列を含むポリペプチド、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および（ｄ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片から構成される群から選択されたポリペプチドに特異結合するような単離された抗体を提供する。

【００４２】

本発明は更に、（ａ）ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択したポリヌクレオチド配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択したポリヌクレオチド配列と少なくとも９０％の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、（ｃ）（ａ）に相補的なポリヌクレオチド配列、（ｄ）（ｂ）に相補的なポリヌクレオチド配列、および（ｅ）（ａ）～（ｄ）のＲＮＡ等価物からなる群から選択された単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドは少なくとも６０の連続したヌクレオチドを含む。

【００４３】

本発明は更に、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。 ここで、標的ポリヌクレオチドは（ａ）ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、（ｂ）ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択したポリヌクレオチド配列と少なくとも９０％の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、（ｃ）（ａ）に相補的なポリヌクレオチド、（ｄ）（ｂ）に相補的なポリヌクレオチド配列、および（ｅ）（ａ）～（ｄ）のＲＮＡ等価物からなる群から選択されたポリヌクレオチド配列を含む。検出方法は、（ａ）サンプル中の標的ポリヌクレオチドに相補的な配列を含む少なくとも２０の連続したヌクレオチドを含むプローブを用いて該サンプルをハイブリダイズする過程と、（ｂ）ハイブリダイゼーション複合体の存在・不存在を検出し、複合体が存在する場合にはオプションでその量を検出する過程からなり、プローブと標的ポリヌクレオチドあるいはその断片の間でハイブリダイゼーション複合体が形成されるような条件下で、プローブは標的ポリヌクレオチドに特異的にハイブリダイズする。一実施態様では、プローブは少なくとも６０の連続したヌクレオチドを含む。

【００４４】

本発明はまた、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。 ここで、標的ポリヌクレオチドは、（ａ）ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択し

たポリヌクレオチド配列を含むポリヌクレオチド、（ｂ）ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択したポリヌクレオチド配列と少なくとも９０％の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、（ｃ）（ａ）のポリヌクレオチドに相補的なポリヌクレオチド、（ｄ）（ｂ）のポリヌクレオチドに相補的なポリヌクレオチド、および（ｅ）（ａ）～（ｄ）のＲＮＡ等価物からなる群から選択された配列のポリヌクレオチドを有する。検出方法は、（ａ）ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、（ｂ）標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含む。

【００４５】

本発明は更に、有効量のポリペプチドと薬剤として許容できる賦形剤とを含む成分を提供する。 有効量のポリペプチドは、（ａ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列と少なくとも９０％の同一性を有する天然のアミノ酸配列を含むポリペプチド、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を含むポリペプチドの生物学的活性断片、および（ｄ）ＳＥＱ ＩＤ ＮＯ：１－２６を有する群から選択したアミノ酸配列の免疫原性断片からなる群から選択される。一実施例では、ＳＥＱ ＩＤ ＮＯ：１－２６からなる一群から選択されたアミノ酸配列を含む組成物を提供する。 更に、本発明は、患者にこの組成物を投与することを含む、機能的ＴＲＩＣＨの発現の低下に関連した疾患やその症状の治療方法を提供する。

【００４６】

本発明はまた、（ａ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列と少なくとも９０％の同一性を有する天然のアミノ酸配列を含むポリペプチド、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列の生物学的活性断片、および（ｄ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列の免疫原性断片からなる群から選択されたポリペプチドのアゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。 スクリーニング方法は、（ａ）ポリペプチドを有するサンプルを化合物に曝す過程と、（ｂ）サンプル中のアゴニスト活性を検出する過程とを含む。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的ＴＲＩＣＨの発現の低下に関連した疾患やその症状の治療方法を提供する。

【００４７】

本発明は更に、（ａ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を含むポリペプチド、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列と少なくとも９０％の同一性を有する天然のアミノ酸配列を含むポリペプチド、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列の生物学的活性断片、および（ｄ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列のポリペプチドの免疫原性断片からなる群から選択されたポリペプチドのアンタゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。 スクリーニング方法は、（ａ）ポリペプチドを含むサンプルを化合物に曝す過程と、（ｂ）サンプル中のアンタゴニスト活性を検出する過程とを含む。一実施態様で本発明は、この方法によって同定したアンタゴニスト化合物と薬剤として許容できる賦形剤とを含む成分を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的ＴＲＩＣＨの過剰な発現に関連した疾患やその症状の治療方法を提供する。

【００４８】

本発明は更に、（ａ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を含むポリペプチド、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列と少なくとも９０％の同一性を有する天然のアミノ酸配列を含むポリペプチド、（

ｃ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列の生物学的活性断片、または（ｄ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列の免疫原性断片を含むポリペプチドに特異結合する化合物をスクリーニングする方法を提供する。 スクリーニング方法は、（ａ）ポリペプチドを適切な条件下で少なくとも１つの試験化合物に混合させる過程と、（ｂ）試験化合物とのポリペプチドの結合を検出し、それによってポリペプチドに特異結合する化合物を同定する過程とを含む。

【００４９】

本発明は更に、（ａ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列、（ｂ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列と少なくとも９０％の相同性を有する天然のアミノ酸配列を含むポリペプチド、（ｃ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を含むポリペプチドの生物学的活性断片、および（ｄ）ＳＥＱ ＩＤ ＮＯ：１－２６からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片からなる群から選択されたポリペプチドの活性を調節する化合物をスクリーニングする方法を提供する。 スクリーニング方法は、（ａ）ポリペプチドの活性が許容された条件下で、ポリペプチドを少なくとも１つの試験化合物と混合させる過程と、（ｂ）ポリペプチドの活性を試験化合物の存在下で算定する過程と、（ｃ）試験化合物の存在下でのポリペプチドの活性を試験化合物の不存在下でのポリペプチドの活性と比較する過程とを含み、試験化合物の存在下でのポリペプチドの活性の変化は、ポリペプチドの活性を調節する化合物であることを意味する。

【００５０】

本発明は更に、標的ポリヌクレオチドの変異発現の有効性を確認するために化合物をスクリーニングする方法を提供する。 標的ポリヌクレオチドは、ＳＥＱ ＩＤ ＮＯ：２７－５２を有する群から選択した配列を含む。 スクリーニング方法は、（ａ）標的ポリヌクレオチドを含むサンプルを化合物に曝す過程と、（ｂ）標的ポリヌクレオチドの変異発現を検出する過程とを含む。

【００５１】

本発明は更に、（ａ）核酸を含む生物学的サンプルを試験化合物で処理する過程と、（ｂ）（ｉ）ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、（ｉｉ）ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択したポリヌクレオチド配列と少なくとも９０％の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、（ｉｉｉ）（ｉ）に相補的な配列を有するポリヌクレオチド、（ｉｖ）（ｉｉ）のポリヌクレオチドに相補的なポリヌクレオチド、（ｖ）（ｉ）～（ｉｖ）のＲＮＡ等価物からなる群から選択したポリヌクレオチドの少なくとも２０の連続したヌクレオチドを含むプローブを用いて、処理した生物学的サンプルの核酸をハイブリダイズする過程とを含む試験化合物の毒性の算定方法を提供する。ハイブリダイゼーションは、上記プローブと生物学的サンプル中の標的ポリヌクレオチドの間に特定のハイブリダイゼーション複合体が形成されるような条件下で発生し、上記標的ポリヌクレオチドは、（ｉ）ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、（ｉｉ）ＳＥＱ ＩＤ ＮＯ：２７－５２からなる群から選択したポリヌクレオチド配列と少なくとも９０％の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、（ｉｉｉ）（ｉ）のポリヌクレオチドに相補的な配列を有するポリヌクレオチド、（ｉｖ）（ｉｉ）のポリヌクレオチドに相補的なポリヌクレオチド、および（ｖ）（ｉ）～（ｉｖ）のＲＮＡ等価物からなる群から選択する。或いは、標的ポリヌクレオチドは、上記（ｉ）～（ｖ）からなる群から選択したポリヌクレオチド配列の断片と、（ｃ）ハイブリダイゼーション複合体の量を定量する過程と、（ｄ）処理した生物学的サンプルのハイブリダイゼーション複合体の量を、非処理の生物学的サンプルのハイブリダイゼーション複合体の量と比較する過程を含み、処理した生物学的サンプルのハイブリダイゼーション複合体の量の差は、試験化合物の毒性を示す。

【００５２】

（発明を実施するための形態）

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

【００５３】

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

【００５４】

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

【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。

【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「対立遺伝子変異配列」は、ＴＲＩＣＨをコードする遺伝子の別の形を指す。対立遺伝子変異体は、核酸配列における少なくとも１つの突然変異から作製し得る。 また、変異ｍＲＮＡまたはポリペプチドを作製し得る。その構造または機能は、変異することもしないこともある。遺伝子は、天然の対立遺伝子変異体を全く有しないか、１個若しくは数個の天然の対立遺伝子変異体を有し得る。一般に対立遺伝子変異体を生じさせる通常の突然変異性変化は、ヌクレオチドの自然欠失、付加または置換に帰するものである。これら各変化は、単独或いは他の変化と共に、所定の配列内で１回若しくは数回生じ得る。

【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対立遺伝子変異配列との不適当或いは予期しないハイブリダイゼーション、並びに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型性を含む。コードされたタンパク質も変異され得り、サイレント変化を生じ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、負に帯電したアミノ酸にはアスパラギン酸及びグルタミン酸があり

、正に帯電したアミノ酸にはリジン及びアルギニンがある。親水性値が近似している非荷電極性側鎖を有するアミノ酸には、アスパラギンとグルタミン、セリンとスレオニンがある。親水性値が近似している非荷電側鎖を有するアミノ酸には、ロイシンとイソロイシンとバリン、グリシンとアラニン、フェニルアラニンとチロシンがある。

【００５９】

「アミノ酸」または「アミノ酸配列」の語は、オリゴペプチド、ペプチド、ポリペプチド若しくはタンパク質の配列またはその断片を指し、天然または合成分子を指す。ここで、「アミノ酸配列」は天然のタンパク質分子のアミノ酸配列を指すものであり、「アミノ酸配列」及び類似の語は、アミノ酸配列を、列挙したタンパク質分子に関連する完全な本来のアミノ酸配列に限定しようとするものではない。

【００６０】

「増幅」は、核酸配列の追加複製に関連する。増幅は通常、当業者によく知られたポリメラーゼ連鎖反応（ＰＣＲ）技術を用いて行う。

【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】

「抗体」の語は、エピトープの決定基と結合することができる、そのままの免疫グロブリンやその断片、例えばＦａ、Ｆ（ａｂ’）２ 及びＦｖ断片を指す。ＴＲＩＣＨポリペプチドと結合する抗体は、免疫抗原として、そのままのポリペプチド、または関心のある小ペプチドを含む断片を用いて作製可能である。動物（マウス、ラット、ウサギ等）を免疫化するために用いるポリペプチドまたはオリゴペプチドは、ＲＮＡの翻訳、または化学合成によって得られるポリペプチドまたはオリゴペプチドに由来し得るもので、好みに応じて担体タンパク質に抱合することも可能である。通常用いられる担体であってペプチドと化学結合するものは、ウシ血清アルブミン、サイログロブリン及びスカシガイのヘモシアニン（ＫＬＨ）等がある。結合ペプチドは、動物を免疫化するために用いる。

【００６３】

「抗原決定基」の語は、特定の抗体と接触している分子の領域（即ちエピトープ）を指す。タンパク質またはタンパク質断片を用いて宿主動物を免疫化する場合、タンパク質の多数の領域が、抗原決定基（タンパク質の特定の領域または３次元構造）に特異結合する抗体の産生を誘導し得る。抗原決定基は、抗体への結合において無損傷抗原（即ち免疫応答を誘導するために用いられる免疫原）と競合し得る。

【００６４】

「アプタマー」の語は特定の分子標的に結合する核酸またはオリゴヌクレオチド分子を指す。アプタマーはｉｎ ｖｉｔｒｏの進化プロセスから誘導される（たとえば米国特許番号第５，２７０，１６３号に記載のＳＥＬＥＸ（Ｓｙｓｔｅｍａｔｉｃ Ｅｖｏｌｕｔｉｏｎ ｏｆ Ｌｉｇａｎｄｓ ｂｙ ＥＸｐｏｎｅｎｔｉａｌ Ｅｎｒｉｃｈｍｅｎｔ））。このプロセスでは大きな組み合わせライブラリから標的に特異的なアプタマーの配列が選択される。アプタマーの組成は二本鎖また一本鎖で、デオキシリボヌクレオチド、リボヌクレオチド、ヌクレオチド誘導体またはその他のヌクレオチド様分子を含むことがある。アプタマーのヌクレオチド成分の糖基を修飾することにより（たとえばリボヌクレオチドの２’－ＯＨ 基を２’－Ｆ または ２’－NH_2 で置換）、核酸分解酵素への耐性を改善したり、血液中における寿命を延ばすなど、望ましい改良を加えることができる。循環系からアプタマーが除去される速度を遅くするために、アプタマーを高分子量担体等の分子に抱合させることができる。アプタマーは、たとえば光活性化または架橋剤によって各々のリガンドと特異的に架橋させることができる（Ｂｒｏｄｙ， Ｅ．Ｎ． および Ｌ．Ｇｏｌｄ （２０００） Ｊ． Ｂｉｏｔｅｃｈｎｏｌ． ７４：５－１３等を参照）。

【００６５】

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

【００６６】

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

【００６７】

「アンチセンス」の語は、特定の核酸配列の「センス」（コーディング）鎖と塩基対を形成することが可能な任意の成分を指す。アンチセンス成分には、ＤＮＡや、ＲＮＡや、ペプチド核酸（ＰＮＡ）や、ホスホロチオ酸、メチルホスホン酸またはベンジルホスホン酸等の修飾されたバックボーン連鎖を有するオリゴヌクレオチドや、２’－メトキシエチル糖または２’－メトキシエトキシ糖等の修飾された糖類を有するオリゴヌクレオチドや、或いは５－メチルシトシン、２－デオキシウラシルまたは７－デアザ－２’－デオキシグアノシン等の修飾された塩基を有するオリゴヌクレオチドが含まれうる。アンチセンス分子は、化学合成または転写を含む任意の方法で製造することができる。相補的アンチセンス分子は、ひとたび細胞に導入されたら、細胞が形成した天然の核酸配列と塩基対を形成し、転写または翻訳を妨害する二重鎖を形成する。「負」若しくは「マイナス（－）」の語が参照ＤＮＡ分子のアンチセンス鎖を、「正」若しくは「プラス（＋）」がセンス鎖を指すことがある。

【００６８】

「生物学的に活性」の語は、天然分子の構造的機能、調節機能または生化学的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。

【００６９】

「相補（的）」または「相補性」の語は、塩基対形成によってアニーリングする２つの一本鎖核酸の間の関係を指す。例えば、配列「５’Ａ－Ｇ－Ｔ３’」は、相補配列「３’Ｔ－Ｃ－Ａ５’」に結合する。

【００７０】

「所定のポリヌクレオチド配列を含む成分」及び「所定のアミノ酸配列を含む成分」は、所定のポリヌクレオチド配列またはアミノ酸配列を含む広範囲の任意の成分を指す。この成分には、乾燥製剤または水溶液が含まれ得る。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は、塩（例えばＮａＣｌ）、界面活性剤（例えばドデシル硫酸ナトリウム；ＳＤＳ）及びその他の構成エレメント（例えばデンハート液、脱脂粉乳、サケの精子のＤＮＡ等）を含む水溶液中にプローブを分散させることができる。

【００７１】

「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、再度シークエンシングされた核酸配列、またはＧＥＬＶＩＥＷ 断片構築システム（ＧＣＧ，Ｍａｄｉｓｏｎ，ＷＩ）またはＰｈｒａｐ（Ｕｎｉｖｅｒｓｉｔｙ ｏｆ Ｗａｓｈｉｎｇｔｏｎ，Ｓｅａｔｔｌｅ ＷＡ）等の断片構築用のコンピュータプログラムを用いて１つ或いはそれ以上の重複するｃＤＮＡやＥＳＴ、またはゲノムＤＮＡ断片から構築された核酸配列を指す。伸長及び構築の両方を行ってコンセンサス配列を作製する配列もある。

【００７２】
「保存的なアミノ酸置換」は、置換がなされた時に元のタンパク質の特性を殆ど損なわないと予測されるような置換、即ちタンパク質の構造と特に機能が保存され、そのような置換による大きな変化がない置換を指す。下表は、タンパク質中で元のアミノ酸と置換可能で、保存アミノ酸置換と認められるアミノ酸を示している。

元の残基	保存的な置換
Ala	Gly, Set
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

【００７３】
保存アミノ酸置換では通常、（ａ）置換領域におけるポリペプチドのバックボーン構造、例えばβシートやαヘリックス構造、（ｂ）置換部位における分子の電荷または疎水性、及び／または（ｃ）側鎖の大部分を保持する。
【００７４】
「欠失」は、結果的に１個若しくは数個のアミノ酸残基またはヌクレオチドが失われてなくなるようなアミノ酸またはヌクレオチド配列における変化を指す。
【００７５】
「誘導体」の語は、ポリペプチドまたはポリヌクレオチドの化学修飾を指す。例えば、アルキル基、アシル基、ヒドロキシル基またはアミノ基による水素の置換は、ポリヌクレオチド配列の化学修飾に含まれ得る。ポリヌクレオチド誘導体は、天然分子の生物学的または免疫学的機能を少なくとも１つは保持しているポリペプチドをコードする。ポリペプチド誘導体は、グリコシル化、ポリエチレングリコール化（ｐｅｇｙｌａｔｉｏｎ）、或いは任意の同様なプロセスであって誘導起源のポリペプチドから少なくとも１つの生物学的若しくは免疫学的機能を保持しているプロセスによって、修飾されたポリペプチドである。
【００７６】
「検出可能な標識」は、測定可能な信号を生成することができ、ポリヌクレオチドまたはポリペプチドに共有結合または非共有結合するようなレポーター分子または酵素を指す。
【００７７】
「示差発現」は少なくとも２つの異なったサンプルを比較することによって決められる、

増加、または非調節、あるいは減少、下方調節、または欠損遺伝子またはタンパク発現を指す。このような比較は例えば、治療後サンプルと未治療のサンプルまたは病態のサンプルと正常サンプルの間で行われ得る。

【００７８】

「エキソンシャフリング」は、異なるコード領域の組換えを指す（エキソン）。エキソンはコードされたタンパク質の構造的或いは機能的ドメインを表しうるので、新たなタンパク質は安定的な下部構造の新規な再組合せにより構築される。それによって新たなタンパク質機能の進化の加速が可能となる。

【００７９】

用語「断片」は、ＴＲＩＣＨまたはＴＲＩＣＨをコードするポリヌクレオチドの固有の部分であって、その親配列（ｐａｒｅｎｔ　ｓｅｑｕｅｎｃｅ）と同一であるがその配列より長さが短いものを指す。断片は、画定された配列の全長から１ヌクレオチド／アミノ酸残基を差し引いた長さ以下の長さを有し得る。例えば或る断片は、５～１０００の連続したヌクレオチドまたはアミノ酸残基を有し得る。プローブ、プライマー、抗原、治療用分子として、或いはその他の目的のために用いられる断片は、少なくとも５、１０、１５、１６、２０、２５、３０、４０、５０、６０、７５、１００、１５０、２５０若しくは５００の連続したヌクレオチド或いはアミノ酸残基長さであり得る。断片は、分子の特定領域から優先的に選択し得る。例えば、ポリペプチド断片は、所定の配列に示すような最初の２５０または５００アミノ酸（またはポリペプチドの最初の２５％または５０％）から選択された或る長さの連続したアミノ酸を有し得る。これらの長さは明らかに例として挙げているものであり、本発明の実施例では、配列表、表及び図面を含む明細書に裏付けされた任意の長さであってよい。

【００８０】

ＳＥＱ　ＩＤ　ＮＯ：２７－５２の断片には、固有のポリヌクレオチド配列領域が含まれる。この領域は、ＳＥＱ　ＩＤ　ＮＯ：２７－５２を特異的に同定するものであり、例えばこの断片を得たゲノム中のその他の配列とは異なるものである。ＳＥＱ　ＩＤ　ＮＯ：２７－５２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：２７－５２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。ＳＥＱ　ＩＤ　ＮＯ：２７－５２　の断片の正確な長さ及び断片に対応するＳＥＱ　ＩＤ　ＮＯ：２７－５２　の領域は、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

【００８１】

ＳＥＱ　ＩＤ　ＮＯ：１－２６　の断片は、ＳＥＱ　ＩＤ　ＮＯ：２７－５２の断片によってコードされる。ＳＥＱ　ＩＤ　ＮＯ：１－２６　の断片には、ＳＥＱ　ＩＤ　ＮＯ：１－２６を特異的に同定する固有のアミノ酸配列領域が含まれている。例えば、ＳＥＱ　ＩＤ　ＮＯ：１－２６　の断片は、ＳＥＱ　ＩＤ　ＮＯ：１－２６を特異認識する抗体を産出するための免疫原性ペプチドとして有用である。ＳＥＱ　ＩＤ　ＮＯ：１－２６　の断片及び断片に対応するＳＥＱ　ＩＤ　ＮＯ：１－２６　の領域の正確な長さは、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

【００８２】

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

【００８３】

「相同性」の語は、配列類似性即ち２つ以上のポリヌクレオチド配列または２つ以上のポリペプチド配列の配列間で互換可能な配列同一性である。

【００８４】

ポリヌクレオチド配列に適用される「一致率」または「一致％」の語は、標準化されたアルゴリズムを用いてアラインメントされた少なくとも２つ以上のポリヌクレオチド配列間で一致する残基の割合を意味する。このようなアルゴリズムは、２配列間のアラインメン

トを最適化するために比較する配列において、標準化された再現性のある方法でギャップを挿入するので、２つの配列をより有意に比較できる。

【００８５】

ポリヌクレオチド配列間の一致率は、ＭＥＧＡＬＩＧＮ ｖｅｒｓｉｏｎ ３．１２ｅ配列アラインメントプログラムに組込まれているようなＣＬＵＳＴＡＬ Ｖアルゴリズムのデフォルトのパラメータを用いて決定できる。このプログラムは、ＬＡＳＥＲＧＥＮＥ ソフトウエアパッケージ（一組の分子生物学的分析プログラム）（ＤＮＡＳＴＡＲ，Ｍａｄｉｓｏｎ ＷＩ）の一部である。このＣＬＵＳＴＡＬ Ｖは、Ｈｉｇｇｉｎｓ，Ｄ．Ｇ．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を２つ１組でアラインメントする際のデフォルトパラメータは、Ｋｔｕｐｌｅ＝２、ｇａｐ ｐｅｎａｌｔｙ＝５、ｗｉｎｄｏｗ＝４、「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。ＣＬＵＳＴＡＬ Ｖは、アラインメントされたポリヌクレオチド配列対間の「類似率」として一致率を報告する。

【００８６】

或いは、米国国立バイオテクノロジー情報センター（ＮＣＢＩ）のＢａｓｉｃ Ｌｏｃａｌ Ａｌｉｇｎｍｅｎｔ Ｓｅａｒｃｈ Ｔｏｏｌ（ＢＬＡＳＴ）が一般的に用いられ、且つ、無料で入手可能な配列比較アルゴリズム一式を提供している（Ａｌｔｓｃｈｕｌ，Ｓ．Ｆ．ら（１９９０）Ｊ．Ｍｏｌ．Ｂｉｏｌ．２１５：４０３－４１０）。このアルゴリズムは、幾つかの情報源から入手可能であり、メリーランド州ベセスダにあるＮＣＢＩ及びインターネット（ｈｔｔｐ：／／ｗｗｗ．ｎｃｂｉ．ｎｌｍ．ｎｉｈ．ｇｏｖ／ＢＬＡＳＴ／）からも入手可能である。ＢＬＡＳＴソフトウェア一式には様々な配列分析プログラムが含まれており、既知のポリヌクレオチド配列を種々のデータベースから得た別のポリヌクレオチド配列とアラインメントする「ｂｌａｓｔｎ」もその１つである。その他にも、２つのヌクレオチド配列を対で直接比較するために用いる「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」と称されるツールも利用可能である。「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」は、ｈｔｔｐ：／／ｗｗｗ．ｎｃｂｉ．ｎｌｍ．ｎｉｈ．ｇｏｖ／ｇｏｒｆ／ｂｌ２．ｈｔｍｌにアクセスして対話形式で利用することが可能である。「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」ツールは、ｂｌａｓｔｎ と ｂｌａｓｔｐ（後述）の両方に用いることができる。ＢＬＡＳＴプログラムは、一般的には、ギャップ及びデフォルト設定に設定された他のパラメータと共に用いる。例えば、２つのヌクレオチド配列を比較するために、デフォルトパラメータとして設定された「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」ツールＶｅｒｓｉｏｎ ２．０．１２（２０００年４月２１日）を用いてｂｌａｓｔｎを実行してもよい。デフォルトパラメータの設定例を以下に示す。

【００８７】

Matrix:BLOSUM62
Reward for match: 1 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

【００８８】

一致率は、完全に画定された（例えば特定の配列番号で画定された）配列長さと比較して測定し得る。 或いは、より短い長さ、例えばより大きな画定された配列から得られた断片（例えば少なくとも２０、３０、４０、５０、７０、１００または２００の連続したヌ

クレオチドの断片）の長さと比較して一致率を測定してもよい。ここに挙げた長さは単なる例示的なものに過ぎず、表、図及び配列リストを含めた本明細書に記載された配列に裏付けられた任意の配列長さの断片を用いて、一致率を測定し得る長さを説明し得ることを理解されたい。

【００８９】

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

【００９０】

ポリペプチド配列に適用される「一致率」または「一致％」の語は、標準化されたアルゴリズムを用いてアラインメントされた少なくとも２以上のポリペプチド配列間で一致する残基の割合を意味する。ポリペプチド配列アラインメントの方法は公知である。保存的アミノ酸置換を考慮するアラインメント方法もある。既に詳述したこのような保存的置換は通常、置換部位の電荷及び疎水性を保存するので、ポリペプチドの構造を（従って機能も）保存する。

【００９１】

ポリペプチド配列間の一致率は、ＭＥＧＡＬＩＧＮ ｖｅｒｓｉｏｎ ３．１２ｅ配列アラインメントプログラムに組込まれているようなＣＬＵＳＴＡＬ Ｖアルゴリズムのデフォルトのパラメータを用いて決定できる（既に説明したのでそれを参照されたい）。ＣＬＵＳＴＡＬ Ｖを用いて、ポリペプチド配列を２つ１組でアラインメントする際のデフォルトパラメータは、Ｋｔｕｐｌｅ＝１、ｇａｐ ｐｅｎａｌｔｙ＝３、ｗｉｎｄｏｗ＝５、「ｄｉａｇｏｎａｌｓ ｓａｖｅｄ」＝５と設定する。デフォルトの残基重み付け表としてＰＡＭ２５０マトリクスを選択する。ポリヌクレオチドアラインメントと同様に、ＣＬＵＳＴＡＬ Ｖは、アラインメントされたポリペプチド配列対間の「類似率」として一致率を報告する。

【００９２】

或いは、ＮＣＢＩ ＢＬＡＳＴソフトウェア一式を用いてもよい。例えば、２つのポリペプチド配列を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。デフォルトパラメータの設定例を以下に示す。

【００９３】

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

一致率は、完全に画定された（例えば特定の配列番号で画定された）ポリペプチド配列の長さと比較して測定し得る。 一致率は、配列或いは、より短い長さ、例えばより大きな画定されたポリペプチド配列から得られた断片（例えば少なくとも１５、２０、３０、４０、５０、７０または１５０の連続した残基の断片）の長さと比較して一致率を測定してもよい。ここに挙げた長さは単なる例示的なものに過ぎず、表、図及び配列リストを含めた本明細書に記載された配列に裏付けられた任意の配列長さの断片を用いて、その長さに対して一致率を測定し得る長さを記述し得ることを理解されたい。

【００９４】

「ヒト人工染色体（ＨＡＣ）」は、約６ｋｂ（キロベース）～１０Ｍｂのサイズの ＤＮＡ配列を含み得る、染色体の複製、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。

【００９５】

「ヒト化抗体」の語は、非抗体結合領域におけるアミノ酸配列はヒト抗体により近づくように変異させた抗体分子であって、本来の結合能力はそのまま保持しているような抗体分子を指す。

【００９６】

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

【００９７】

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

【００９８】

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０．２ｘ 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濃度は、約０．１％のＳＤＳ存在下で、約０．１～２×ＳＳＣの範囲で変化し得る。通常は、ブロッキング剤を用いて非特異ハイブリダイゼーションを阻止する。このような遮断剤には、例えば、約１００～２００μｇ／ｍｌの変性サケ精子ＤＮＡがある。例えばＲＮＡとＤＮＡのハイブリダイゼーションのような特定条件下では、有機溶剤、例えば約３５～５０％ｖ／ｖの濃度のホルムアミドを用いることもできる。洗浄条件の有用なバリエーションは、当業者には自明であろう。 ハイブリダイゼーションは、特に高ストリンジェント条件下では、ヌクレオチド間の進化的な類似性を示唆し得る。このような類似性は、ヌクレオチド及びヌクレオチドにコードされるポリペプチドに対する類似の役割を強く示唆している。

【００９９】

「ハイブリダイゼーション複合体」の語は、相補的塩基対間の水素結合の形成力によって２つの核酸配列間に形成された複合体を指す。ハイブリダイゼーション複合体は、溶解状態で形成し得る（C_0tまたはR_0t解析等）。 或いは、一方の核酸配列が溶解状態で存在し、もう一方の核酸配列が固体支持体（例えば紙、膜、フィルタ、チップ、ピンまたはガラススライド、或いは他の適切な基板であって細胞若しくはその核酸が固定される基板）に固定されているような２つの核酸配列間に形成され得る。

【０１００】

「挿入」及び「付加」の語は、１個若しくは数個のアミノ酸残基またはヌクレオチドを各

々付加するようなアミノ酸またはヌクレオチド配列における変化を指す。
【０１０１】
「免疫応答」は、炎症、外傷、免疫異常症、伝染性疾患または遺伝性疾患に関連する症状を指し得る。 これらの症状は、細胞及び全身の防御系に作用し得る種々の因子、例えばサイトカイン、ケモカイン、その他のシグナル伝達分子の発現によって特徴づけることができる。
【０１０２】
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。
【０１０３】
「マイクロアレイ」の語は、基板上の複数のポリヌクレオチド、ポリペプチドまたはその他の化合物の構成を指す。
【０１０４】
「エレメント」または「アレイエレメント」の語は、マイクロアレイ上に定義された固有の位置にあるハイブリダイズ可能なポリヌクレオチド、ポリペプチドその他の化合物を指す。
【０１０５】
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。
【０１０６】
「核酸」及び「核酸配列」の語は、ヌクレオチド、オリゴヌクレオチド、ポリヌクレオチドまたはこれらの断片を指す。「核酸」及び「核酸配列」の語は、ゲノム起源または合成起源のＤＮＡまたはＲＮＡであって一本鎖または二本鎖であるか或いはセンス鎖またはアンチセンス鎖を表し得るようなＤＮＡまたはＲＮＡや、ペプチド核酸（ＰＮＡ）や、任意のＤＮＡ様またはＲＮＡ様物質を指すこともある。
【０１０７】
「機能的に連結した」は、第１核酸配列が第２核酸配列と機能的な関係があるように配置された状態を指す。例えば、プロモーターがコード配列の転写または発現に影響を及ぼす場合には、そのプロモーターはそのコード配列に機能的に連結している。機能的に連結したＤＮＡ配列は非常に近接するか、或いは連続し得る。そして、２つのタンパク質コード領域をつなぐ必要がある場合は、同一のリーディングフレーム内にある。
【０１０８】
「ペプチド核酸」（ＰＮＡ）は、アンチセンス分子または抗遺伝子物質であって、リジンを末端とするアミノ酸残基のペプチドバックボーンに結合した、少なくとも約５ヌクレオチドの長さのオリゴヌクレオチドからなるものを指す。末端のリジンは、成分に溶解性を与える。ＰＮＡは、相補的一本鎖ＤＮＡまたはＲＮＡに優先的に結合して転写の拡張を停止するものであり、ポリエチレングリコール化して細胞におけるＰＮＡの寿命を延長し得る。
【０１０９】
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。
【０１１０】
「プローブ」とは、同一配列或いは対立遺伝子核酸配列、関連する核酸配列の検出に用いる、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、単離されたオリゴヌクレオチドまたはポリヌクレオチドであって、検

出可能な標識またはレポーター分子に結合したものである。典型的な標識には、放射性アイソトープ、リガンド、化学発光試薬及び酵素がある。「プライマー」は、短い核酸、通常はＤＮＡオリゴヌクレオチドであり、相補的塩基対を形成することで標的ポリヌクレオチドにアニーリングされ得る。プライマーは次に、ＤＮＡポリメラーゼ酵素によって標的ＤＮＡ鎖に延在し得る。プライマー対は、例えばポリメラーゼ連鎖反応（ＰＣＲ）による核酸配列の増幅（及び同定）に用い得る。

【０１１１】

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

【０１１２】

プローブ及びプライマーの調製及び使用方法については、Ｓａｍｂｒｏｏｋ，Ｊ．ら（１９８９）Ｍｏｌｅｃｕｌａｒ Ｃｌｏｎｉｎｇ：Ａ Ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ，第２版，１－３巻，Ｃｏｌｄ Ｓｐｒｉｎｇ Ｈａｒｂｏｒ Ｐｒｅｓｓ，Ｐｌａｉｎｖｉｅｗ ＮＹ、Ａｕｓｕｂｅｌ，Ｆ．Ｍ．ら，（１９８７）Ｃｕｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｍｏｌｅｃｕｌａｒ Ｂｉｏｌｏｇｙ，Ｇｒｅｅｎｅ Ｐｕｂｉ．Ａｓｓｏｃ．＆Ｗｉｌｅｙ－Ｉｎｔｅｒｓｃｉｅｎｃｅｓ，Ｎｅｗ Ｙｏｒｋ ＮＹ、Ｉｎｎｉｓら（１９９０）ＰＣＲ Ｐｒｏｔｏｃｏｌｓ，Ａ Ｇｕｉｄｅ ｔｏ Ｍｅｔｈｏｄｓ ａｎｄ Ａｐｐｌｉｃａｔｉｏｎｓ Ａｃａｄｅｍｉｃ Ｐｒｅｓｓ，Ｓａｎ Ｄｉｅｇｏ ＣＡ等を参照されたい。ＰＣＲプライマー対は、その目的のためのコンピュータプログラム、例えばＰｒｉｍｅｒ（Ｖｅｒｓｉｏｎ ０．５，１９９１，Ｗｈｉｔｅｈｅａｄ Ｉｎｓｔｉｔｕｔｅ ｆｏｒ Ｂｉｏｍｅｄｉｃａｌ Ｒｅｓｅａｒｃｈ，Ｃａｍｂｒｉｄｇｅ ＭＡ）を用いるなどして既知の配列から得ることができる。

【０１１３】

プライマーとして用いるオリゴヌクレオチドの選択は、そのような目的のために本技術分野でよく知られているソフトウェアを用いて行う。例えばＯＬＩＧＯ ４．０６ソフトウェアは、各１００ヌクレオチドまでのＰＣＲプライマー対の選択に有用であり、オリゴヌクレオチド及び最大５，０００ヌクレオチドまでの大きめのポリヌクレオチドであって３２キロベースまでのインプットポリヌクレオチド配列から得たものを分析するのにも有用である。類似のプライマー選択プログラムには、拡張能力のための追加機能が組込まれている。例えば、ＰｒｉｍＯＵプライマー選択プログラム（テキサス州ダラスにあるテキサス大学南西部医療センターのゲノムセンターから一般向けに入手可能）は、メガベース配列から特定のプライマーを選択することが可能であり、従ってゲノム全体の範囲でプライマーを設計するのに有用である。Ｐｒｉｍｅｒ３プライマー選択プログラム（マサチューセッツ州ケンブリッジのＷｈｉｔｅｈｅａｄ Ｉｎｓｔｉｔｕｔｅ／ＭＩＴゲノム研究センターから一般向けに入手可能）ではユーザーが「ミスプライミング・ライブラリ」をインプットすることができ、ここでプライマー結合部位として避けたい配列はユーザーが指定する。Ｐｒｉｍｅｒ３は特に、マイクロアレイのためのオリゴヌクレオチドの選択に有用である。（後二者のプライマー選択プログラムのソースコードは、各自のソースから得てユーザー固有のニーズを満たすように変更してもよい。）ＰｒｉｍｅｒＧｅｎプログラム（英国ケンブリッジ市の英国ヒトゲノムマッピングプロジェクト－リソースセンターから一般向けに入手可能）は、多数の配列アラインメントに基づいてプライマーを設計し、それによって、アラインメントされた核酸配列の最大保存領域または最小保存領域の何れかとハイブリダイズするようなプライマーの選択を可能にする。従って、このプログラムは、固有な、および保存されたオリゴヌクレオチド及びポリヌクレオチドの断片の同定に有用である。上記選択方法のいずれかによって同定したオリゴヌクレオチド及びポリヌク

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

【０１１４】

「組換え核酸」は天然配列ではない配列であるか或いは人為的に組み合わせなければ離隔しているような配列の２以上のセグメントを人為的に組み合わせて産出した配列を有する配列である。この人為的組合せはしばしば化学合成によって達成するが、より一般的には核酸の単離セグメントの人為的操作によって、例えばのＳａｍｂｒｏｏｋらの文献（前出）に記載されているような遺伝子工学的手法によって達成する。組換え核酸の語は、単に核酸の一部を付加、置換または欠失した変異核酸も含む。しばしば組換え核酸には、プロモーター配列に機能的に連結した核酸配列が含まれる。このような組換え核酸は、例えばある細胞を形質転換するために使用されるベクターの一部とすることが可能である。

【０１１５】

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

【０１１６】

「調節エレメント」は、通常は遺伝子の未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’の未翻訳領域（ＵＴＲ）を含む。調節エレメントは、転写、翻訳またはＲＮＡ安定性を調節する宿主タンパク質またはウイルスタンパク質と相互作用する。

【０１１７】

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

【０１１８】

ＤＮＡ配列に関する「ＲＮＡ等価物」は、窒素塩基チミンが全てウラシルに置換されていることと、糖のバックボーンがデオキシリボースではなくリボースから構成されていることを除いて、参照ＤＮＡ配列と同一のヌクレオチド線形配列から構成されている。

【０１１９】

「サンプル」の語は、その最も広い意味で用いられる。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。

【０１２０】

「特異結合」または「特異的に結合する」の語は、タンパク質またはペプチドと、アゴニスト、抗体、アンタゴニスト、小分子、任意の天然成分または合成結合成分との間の相互作用を指す。この相互作用は、タンパク質の特定の構造（例えば抗原決定基即ちエピトープ）であって結合分子が認識するものが存在するか否かに依存している。例えば、抗体がエピトープ「Ａ」に対して特異的である場合、標識された遊離したＡ及びその抗体を含む反応において、エピトープＡ（つまり遊離し、標識されていないＡ）を含むポリペプチドの存在が、抗体に結合している標識されたＡの量を低減させる。

【０１２１】

「実質上精製された」の語は、自然環境から取り除かれ、或いは単離または分離された核酸またはアミノ酸配列であって、自然に会合するその他の構成エレメントの少なくとも約６０％、好ましくは少なくとも約７５％、最も好ましいのは少なくとも約９０％が遊離しているものを指す。

【０１２２】
「置換」は、１個若しくは数個のアミノ酸残基またはヌクレオチドを各々別のアミノ酸残基またはヌクレオチドに置換することを意味する。
【０１２３】
「基板」は、任意の好適な固体または半固体の支持体を指すものであって、膜、フィルタ、チップ、スライド、ウエハ、ファイバー、磁性または非磁性ビーズ、ゲル、管、プレート、ポリマー、微細粒子、毛管が含まれる。基板は、凹み、溝、ピン、チャネル、孔等、様々な表面形態を有することができ、基板表面にはポリヌクレオチドやポリペプチドが結合する。
【０１２４】
「転写イメージ」は、所与の時間、条件での固有の細胞タイプまたは組織による遺伝子発現の集合的パターンを指す。
【０１２５】
「形質転換」は、外来性のＤＮＡが受入細胞に入り込むプロセスを表す。形質転換は、本技術分野で知られている種々の方法に従って自然条件または人工条件下で生じ得るものであり、外来性の核酸配列を原核または真核宿主細胞に挿入する任意の既知の方法を基にし得る。形質転換の方法は、形質転換する宿主細胞の種類によって選択する。 限定するものではないが形質転換方法には、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）、熱ショック、リポフェクション及び微粒子銃を用いる方法がある。「形質転換された」細胞には、導入されたＤＮＡが自律的に複製するプラスミドとして或いは宿主染色体の一部として複製可能である安定的に形質転換された細胞が含まれる。 さらに、限られた時間に一過的に導入ＤＮＡ若しくは導入ＲＮＡを発現する細胞も含まれる。
【０１２６】
ここで用いる「遺伝形質転換体」とは任意の生物体であり、限定するものではないが動植物を含み、生物体の１個若しくは数個の細胞が、ヒトの関与によって、例えば本技術分野でよく知られている形質転換技術によって導入された異種核酸を有する。 核酸の細胞への導入は、直接または間接的に、細胞の前駆物質に導入することによって、計画的な遺伝子操作によって、例えば微量注射法によって或いは組換えウイルスの導入によって行う。遺伝子操作の語は、古典的な交雑育種或いはｉｎ ｖｉｔｒｏ受精を指すものではなく、組換えＤＮＡ分子の導入を指すものである。本発明に基づいて予期される遺伝形質転換体には、バクテリア、シアノバクテリア、真菌及び動植物がある。本発明の単離されたＤＮＡは、本技術分野で知られている方法、例えば感染、形質移入、形質転換またはトランス接合によって宿主に導入することができる。本発明のＤＮＡをこのような有機体に移入する技術はよく知られており、前出のＳａｍｂｒｏｏｋ ら （１９８９） 等の参考文献に与えられている。
【０１２７】
特定の核酸配列の「変異体」は、核酸配列１本全部の長さに対して特定の核酸配列と少なくとも４０％の相同性を有する核酸配列であると定義する。 その際、デフォルトパラメータに設定した「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」ツールＶｅｒｓｉｏｎ ２．０．９（１９９９年５月７日）を用いてｂｌａｓｔｎを実行する。このような核酸対は、所定の長さに対して、例えば少なくとも５０％、６０％、７０％、８０％、８５％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％またはそれ以上の相同性を示し得る。或る変異体は、例えば「対立遺伝子」変異体（前述）、「スプライス」変異体、「種」変異体または「多型性」変異体として説明し得る。スプライス変異体は参照分子とかなりの相同性を有し得るが、ｍＲＮＡプロセッシング中のエキソンの交互スプライシングによって通常、より多数またはより少数のポリヌクレオチドを有することになる。対応するポリペプチドは、追加機能ドメインを有するか或いは参照分子に存在するドメインが欠落していることがある。種変異体は、種相互に異なるポリヌクレオチド配列である。結果的に生じるポリペプチドは通常、相互にかなりのアミノ酸相同性を有

する。多型性変異体は、与えられた種の個体間で特定の遺伝子のポリヌクレオチド配列が異なる。また、多型性変異体は、１つのヌクレオチド塩基によってポリヌクレオチド配列が変化する「１塩基多型性」（ＳＮＰ）を含み得る。ＳＮＰの存在は、例えば特定の個体群、病状または病状性向を示し得る。

【０１２８】

特定のポリペプチド配列の「変異体」は、ポリペプチド配列の１本の長さ全体で特定のポリペプチド配列に対して少なくとも４０％の相同性を有するポリペプチド配列として画定される。 ここで、デフォルトパラメータに設定した「ＢＬＡＳＴ ２ Ｓｅｑｕｅｎｃｅｓ」ツールＶｅｒｓｉｏｎ ２．０．９（１９９９年5月7日）を用いてｂｌａｓｔｐを実行する。このようなポリペプチド対は、所定の長さに対して、例えば少なくとも５０％、６０％、７０％、８０％、９０％、９１％、９２％、９３％、９４％、９５％、９６％、９７％、９８％、９９％またはそれ以上の配列同一性を示し得る。

【０１２９】

（発明）

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

【０１３０】

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

【０１３１】

表２は、ＧｅｎＢａｎｋタンパク質（ｇｅｎｐｅｐｔ）データベースに対するＢＬＡＳＴ分析によって同定されたような、本発明のポリペプチドとの相同性を有する配列を示している。列１および列２はそれぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号（Ｐｏｌｙｐｅｐｔｉｄｅ ＳＥＱ ＩＤ ＮＯ ：）およびそれに対応するＩｎｃｙｔｅ ポリペプチド配列番号（Ｉｎｃｙｔｅ Ｐｏｌｙｐｅｐｔｉｄｅ ＩＤ）を示す。と対応するＩｎｃｙｔｅポリペプチド配列番号（ＩｎｃｙｔｅポリペプチドＩＤ）によって表示した。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。（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、ＧｅｎＢａｎｋ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。

【０１３２】

表３は、本発明のポリペプチドの様々な構造的特徴を示す。列１および列２はそれぞれ、本発明の各ポリペプチドのポリペプチド配列識別番号（ＳＥＱ ＩＤ ＮＯ ：）およびそれに対応するＩｎｃｙｔｅ ポリペプチド配列番号（Ｉｎｃｙｔｅ Ｐｏｌｙｐｅｐｔｉｄｅ ＩＤ）を示す。と対応するＩｎｃｙｔｅポリペプチド配列番号（ＩｎｃｙｔｅポリペプチドＩＤ）を示す。列３は、各ポリペプチドのアミノ酸残基数を示す。列４および列５はそれぞれ、ＧＣＧ配列分析ソフトウェアパッケージのＭＯＴＩＦＳプログラム（Ｇｅｎｅｔｉｃｓ Ｃｏｍｐｕｔｅｒ Ｇｒｏｕｐ，Ｍａｄｉｓｏｎ ＷＩ）によって決定された、リン酸化およびグリコシル化の可能性のある部位を示す。列６は、シグネチャ（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。

【０１３３】

表２及び３は共に、本発明の各々のポリペプチドの特性を要約しており、それら特性が請求の範囲に記載されたポリペプチドが輸送体及びイオンチャネルであることを確立している。例えば、ＳＥＱ ＩＤ ＮＯ：２は、Ｂａｓｉｃ Ｌｏｃａｌ Ａｌｉｇｎｍｅｎｔ Ｓｅａｒｃｈ Ｔｏｏｌ （ＢＬＡＳＴ）で決定されるようにマウスのａｂｃ２輸送体（ＧｅｎＢａｎｋ ＩＤ ｇ４９５２５９）とアミノ酸９６５から２４３６にかけて９４％同一である（表２参照）。ＢＬＡＳＴ確率スコアは０．０であるが、これは観測されたポリペプチド配列が偶然に得られる確率を示す。ＳＥＱ ＩＤ ＮＯ：２はまた、ＡＢＣ輸送体ドメインを有するが、 これは、隠れマルコフモデル（ＨＭＭ）を基にした保存されたタンパク質ファミリードメインのＰＦＡＭデータベースにおいて、統計的に有意な一致を検索して決定された（表３参照）。ＭＯＴＩＦＳおよび ＰＲＯＦＩＬＥＳＣＡＮ 解析からのデータは、ＳＥＱ ＩＤ ＮＯ：３ がＡＢＣ輸送体である、さらに実証的な証拠を提供する。別の例において、ＳＥＱ ＩＤ ＮＯ：１３ はＢａｓｉｃ Ｌｏｃａｌ Ａｌｉｇｎｍｅｎｔ Ｓｅａｒｃｈ Ｔｏｏｌ （ＢＬＡＳＴ）によって同定されるヒト筋肉アセチルコリン受容体のガンマサブユニット前駆体（ＧｅｎＢａｎｋ ＩＤ ｇ８２５６１８）に９７％の同一性を有する（表２参照）。ＢＬＡＳＴ確率スコアは３．０ｅ－２７３であり、それは、偶然に観測されたポリペプチド配列を得る確率を示す。ＳＥＱ ＩＤ ＮＯ：１３はまた、１つの神経伝達物質依存性イオンチャネルドメインを有するが、 これは、隠れマルコフモデル（ＨＭＭ）を基にした保存されたタンパク質ファミリードメインのＰＦＡＭデータベースにおいて、統計的に有意な一致を検索して決定された（表３参照）。ＢＬＩＭＰＳ、ＭＯＴＩＦＳ、及びＰＲＯＦＩＬＥＳＣＡＮ解析よりのデータは、ＳＥＱ ＩＤ ＮＯ：１３ が神経伝達物質依存性イオンチャネルである、さらに実証的な証拠を提供する。別の例において、ＳＥＱ ＩＤ ＮＯ：１９ はＢａｓｉｃ Ｌｏｃａｌ Ａｌｉｇｎｍｅｎｔ Ｓｅａｒｃｈ Ｔｏｏｌ （ＢＬＡＳＴ）によって同定されるヒトの液泡プロトン輸送体ＡＴＰ分解酵素 （ＧｅｎＢａｎｋ ＩＤ ｇ３７６４３）に６２％の同一性を有する（表２参照）。ＢＬＡＳＴ確率スコアは３．２ｅ－１２９であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。ＢＬＡＳＴ解析よりのデータは、配列識別番号１９がＡＴＰシンターゼである、さらに実証的な証拠を提供する。別の例において、ＳＥＱ ＩＤ ＮＯ：２２ はＢａｓｉｃ Ｌｏｃａｌ Ａｌｉｇｎｍｅｎｔ Ｓｅａｒｃｈ Ｔｏｏｌ （ＢＬＡＳＴ）によって同定されるラットのＧＡＢＡ（Ａ）受容体ガンマ１サブユニット （ＧｅｎＢａｎｋ ＩＤ ｇ５６１７６）に９４％の同一性を有する（表２参照）。ＢＬＡＳＴ確率スコアは４．４ｅ－２４４であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。ＳＥＱ ＩＤ ＮＯ：２２はまた、１つの神経伝達物質依存性イオンチャネルドメインを有するが、これは、隠れマルコフモデル（ＨＭＭ）を基にした保存されたタンパク質ファミリードメインのＰＦＡＭデータベースにおいて、統計的に有意な一致を検索して決定された。（表３参照）ＢＬＩＭＰＳ、ＭＯＴＩＦＳ、及びＰＲＯＦＩＬＥＳＣＡＮ解析よりのデータは、ＳＥＱ ＩＤ ＮＯ：２２ が神経伝達物質依存性イオンチャネルである、さらに実証的な証拠を提供する。別の例として、ＳＥＱ ＩＤ ＮＯ：２６はウサギのペルオキシソームのＣａ依存性溶質担体（ＧｅｎＢａｎｋ ＩＤ ｇ２３５２４２７） と６１％の同一性を有することがＢａｓｉｃ Ｌｏｃａｌ Ａｌｉｇｎｍｅｎｔ Ｓｅａｒｃｈ Ｔｏｏｌ （ＢＬＡＳＴ）によって示された（表２参照）。ＢＬＡＳＴ確率スコアは６．４ｅ－１５６であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。ＳＥＱ ＩＤ ＮＯ：２６はまた、３つのミトコンドリア担体タンパク質ドメインおよび３つＥＦハンドドメインのを有するが、 これは、隠れマルコフモデル（ＨＭＭ）を基にした保存されたタンパク質ファミリードメインのＰＦＡＭデータベースにおいて、統計的に有意な一致を検索して確認された（表３参照）。ＢＬＩＭＰＳ、ＭＯＴＩＦＳ、及びＰＲＯＦＩＬＥＳＣＡＮ解析よりのデータは、ＳＥＱ ＩＤ ＮＯ：２６ がカルシウム

依存性担体タンパク質である、さらに実証的な証拠を提供する。別の例において、ＳＥＱ ＩＤ ＮＯ：１７ はＢａｓｉｃ Ｌｏｃａｌ Ａｌｉｇｎｍｅｎｔ Ｓｅａｒｃｈ Ｔｏｏｌ （ＢＬＡＳＴ）によって同定されるＡｍｂｙｓｔｏｍａ ｔｉｇｒｉｎｕｍ の起電性$NaHCO_3$ 共輸送体 （ＧｅｎＢａｎｋ ＩＤ ｇ２１９８８１５）に６９％の同一性を有する（表２参照）。ＢＬＡＳＴ確率スコアは０．０であるが、これは観測されたポリペプチド配列が偶然に得られる確率を示す。ＳＥＱ ＩＤ ＮＯ：１７はまた、HCO_3 輸送体ファミリドメインを有するが、 これは、隠れマルコフモデル（ＨＭＭ）を基にした保存されたタンパク質ファミリードメインのＰＦＡＭデータベースにおいて、統計的に有意な一致を検索して決定された（表３参照）。ＢＬＩＭＰＳ、ＭＯＴＩＦＳ、及びＰＲＯＦＩＬＥＳＣＡＮ解析よりのデータは、ＳＥＱ ＩＤ ＮＯ：１７ がアニオン輸送体である、さらに確証的な証拠を提供する。ＳＥＱ ＩＤ ＮＯ：１， ＳＥＱ ＩＤ ＮＯ：３－１２， ＳＥＱ ＩＤ ＮＯ：１４－１６， ＳＥＱ ＩＤ ＮＯ：１８，および ＳＥＱ ＩＤ ＮＯ：２０－２５ については、同様の方法で分析し、注釈を付けた。ＳＥＱ ＩＤ ＮＯ：１－２６ の解析のためのアルゴリズム及びパラメータが表７で記述されている。

【０１３４】

表４に示すように、本発明の完全長ポリヌクレオチド配列は、ｃＤＮＡ配列またはゲノムＤＮＡ由来のコード（エキソン）配列を用いて、或いはこれら２種類の配列を任意に組み合わせて構築した。列１および列２はそれぞれ、本発明の各ポリヌクレオチドのポリヌクレオチド配列識別番号（Ｐｏｌｙｎｕｃｌｅｏｔｉｄｅ ＳＥＱ ＩＤ ＮＯ ：）およびそれに対応するＩｎｃｙｔｅ ポリヌクレオチドコンセンサス配列番号（Ｉｎｃｙｔｅ Ｐｏｌｙｎｕｃｌｅｏｔｉｄｅ ＩＤ）を示す。列３は、各ポリヌクレオチド配列の長さ（塩基対単位）を示す。列４は、例えば、ＳＥＱ ＩＤ ＮＯ：２７－５２を同定するため、或いはＳＥＱ ＩＤ ＮＯ：２７－５２と関連するポリヌクレオチド配列とを区別するためのハイブリダイゼーションまたは増幅技術に有用なポリヌクレオチド配列の断片を示す。列５はｃＤＮＡ配列、ゲノムＤＮＡから予想されたコード配列（エキソン）及び／またはｃＤＮＡ及びゲノムＤＮＡを共に有する配列集合に対応する識別番号を示している。これらの配列は、本発明の完全長ポリヌクレオチド配列を構築するのに用いた。表４の列６および列７はそれぞれ完全長配列に対して、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。

【０１３５】

表４の列５の識別番号は、特に例えばＩｎｃｙｔｅ ｃＤＮＡとそれに対応するｃＤＮＡライブラリを指す場合もある。例えば、７２５１２６６Ｆ７ はインサイトｃＤＮＡ配列の識別番号であり、ＰＲＯＳＴＭＹ０１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、７０６１４０２１Ｖ１）に由来する。または、列５の識別番号は、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、ｇ４６８９８０１）の識別番号の場合もある。また、列５の識別番号は、ＥＮＳＥＭＢＬ （Ｔｈｅ Ｓａｎｇｅｒ Ｃｅｎｔｒｅ， Ｃａｍｂｒｉｄｇｅ， ＵＫ）データベースに由来する配列（すなわち、名前に「ＥＮＳＴ」を含む配列）の識別番号である場合もある。また、列５の識別番号は、ＮＣＢＩ ＲｅｆＳｅｑ ヌクレオチド配列記録データベースに由来する配列（すなわち、名前に「ＮＭ」または「ＮＴ」を含む配列）あるいはＮＣＢＩ ＲｅｆＳｅｑタンパク質配列記録に由来する配列（すなわち、名前に「ＮＰ」を含む配列）の識別番号である場合もある。または列５の識別番号は、「エキソンスティッチング（ｅｘｏｎ－ｓｔｉｔｃｈｉｎｇ）」アルゴリズムにより結び合わせたｃＤＮＡ及びＧｅｎｓｃａｎ予想エキソンの両方の集合を指す場合もある。例えば、ＦＬ＿ＸＸＸＸＸＸ＿N_1＿N_2＿ＹＹＹＹＹ＿N_3＿N_4 はアルゴリズムが適用される配列のクラスターの識別番号がＸＸＸＸＸＸであり、アルゴリズムにより生成される予測の番号がＹＹＹＹＹ であり、（もし存在すれば）$N_{1,2,3..}$ が解析中に手動で編集された可能性のある特定のエキソンであるような「縫合された」配列である（ 実施例５参照 ）。または列５の識別番号は、「エキソンストレッチング（ｅｘｏｎ－ｓｔｒｅｔ

ｃｈｉｎｇ）」アルゴリズムにより結び合わせたエキソンの集合を指す場合もある。例えば、ＦＬＸＸＸＸＸＸ＿ｇＡＡＡＡＡＡ＿ｇＢＢＢＢＢＢ＿１＿Ｎ は「ストレッチ」配列の識別番号である。ここでＸＸＸＸＸＸはＩｎｃｙｔｅプロジェクト識別番号、ｇＡＡＡＡＡは「エキソンストレッチング」アルゴリズムを適用したヒトゲノム配列のＧｅｎＢａｎｋ識別番号、ｇＢＢＢＢＢＢは一番近いＧｅｎＢａｎｋタンパク質相同体のＧｅｎＢａｎｋ識別番号またはＮＣＢＩ ＲｅｆＳｅｑ 識別番号である（実施例５を参照）。ＲｅｆＳｅｑ配列が「エキソンストレッチング」アルゴリズム用のタンパク質相同体として津筒されているときは、ＲｅｆＳｅｑ識別子（「ＮＭ」、「ＮＰ」、または「ＮＴ」で表示）をＧｅｎＢａｎｋ識別し（たとえばｇＢＢＢＢＢＢ）の代わりに使うことができる。

【０１３６】

あるいはまた、手動で編集された、あるいはゲノムＤＮＡ配列から予想された、あるいは配列解析法の組み合わせにより得られたコンポーネント配列は接頭コードで識別される。以下の表はコンポーネント配列の接頭コードおよびその接頭コードに対応する配列解析法の例を示している（実施例４および実施例５を参照）。

接頭コード	解析タイプおよび/またはプログラム例
GNN, GFG, ENST	たとえばGENSCAN (Stanford University, CA, USA) またはFGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK)を使ったゲノム配列からのエキソン予測。
GBI	ゲノム配列の手動編集解析。
FL	縫合（ステッチ）またはストレッチされたゲノム配列（実施例５を参照)。
INCY	ゲノムへのEST配列マッピングに由来する完全長転写およびエキソン予測。ゲノム位置とEST組成物データを結合して、エキソンと生じる転写を予想する。

【０１３７】

場合によっては、最終コンセンサスポリヌクレオチド配列を確認するための列５に示すような配列の適用範囲と重複するＩｎｃｙｔｅ ｃＤＮＡの適用範囲が得られたが、関連するＩｎｃｙｔｅ ｃＤＮＡ識別番号は示さなかった。

【０１３８】

表５は、Ｉｎｃｙｔｅ ｃＤＮＡ配列を用いて構築された完全長ポリヌクレオチド配列の

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

【０１３９】

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。

【０１４０】

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：２７－５２からなる一群から選択された配列を含むポリヌクレオチド配列を提供する。ＳＥＱ ＮＯ ＩＤ：２７－５２のポリヌクレオチド配列は、配列表に示されているように等価ＲＮＡ配列と同等の価値を有しているが、窒素塩基チミンの出現はウラシルに置換され、糖のバックボーンはデオキシリボースではなくシリボースから構成されている。

【０１４１】

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：２７－５２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：２７－５２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。

【０１４２】

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、可能コドン選択に基づく組合せの選択によって産出し得るようなありとあらゆる可能性のあるポリヌクレオチド配列変異体を網羅し得る。これらの組み合わせは、天然の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。

【０１４３】

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ことは有利となり得る。宿主が特定のコドンを利用する頻度に基づいて、特定の真核又は原核宿主に発生するペプチドの発現率を高めるようにコドンを選択することが可能である。コードされたアミノ酸配列を変えないで、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。

【０１４４】

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。

【０１４５】

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：２７－５２ 及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる（例えば、Ｗａｈｌ，Ｇ．Ｍ．及びＳ．Ｌ．Ｂｅｒｇｅｒ（１９８７）Ｍｅｔｈｏｄｓ Ｅｎｚｙｍｏｌ．１５２：３９９－４０７、Ｋｉｍｍｅｌ，Ａ．Ｒ．（１９８７）Ｍｅｔｈｏｄｓ Ｅｎｚｙｍｏｌ．１５２：５０７－５１１等を参照）。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。

【０１４６】

ＤＮＡシークエンシングの方法は当分野でよく知られており、本発明の何れの実施例もＤＮＡシークエンシング方法を用いて実施可能である。ＤＮＡシークエンシング方法には酵素を用いることができ、例えばＤＮＡポリメラーゼＩのクレノウ断片、ＳＥＱＵＥＮＡＳＥ（ＵＳ Ｂｉｏｃｈｅｍｉｃａｌ，Ｃｌｅｖｅｌａｎｄ ＯＨ）、Ｔａｑポリメラーゼ（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）、熱安定性Ｔ７ポリメラーゼ（Ａｍｅｒｓｈａｍ，Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ，Ｐｉｓｃａｔａｗａｙ ＮＪ）を用いることができる。 或いは、例えばＥＬＯＮＧＡＳＥ増幅システム（Ｌｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ，Ｇａｉｔｈｅｒｓｂｕｒｇ ＭＤ）において見られるように、ポリメラーゼと校正エキソヌクレアーゼを併用することができる。好適には、ＭＩＣＲＯＬＡＢ２２００液体転移システム（Ｈａｍｉｌｔｏｎ，Ｒｅｎｏ，ＮＶ）、ＰＴＣ２００サーマルサイクラー（ＭＪ Ｒｅｓｅａｒｃｈ，Ｗａｔｅｒｔｏｗｎ ＭＡ）及びＡＢＩ ＣＡＴＡＬＹＳＴ ８００サーマルサイクラー（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）等の装置を用いて配列の調製を自動化する。次に、ＡＢＩ ３７３ 或いは ３７７ ＤＮＡシークエンシングシステム（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）、ＭＥＧＡＢＡＣＥ １０００ ＤＮＡシークエンシングシステム（Ｍｏｌｅｃｕｌａｒ Ｄｙｎａｍｉｃｓ，Ｓｕｎｎｙｖａｌｅ ＣＡ）または当分野でよく知られている他の方法を用いてシークエンシングを行う。 結果として得られた配列を当分野でよく知られている種々のアルゴリズムを用いて分析する。（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，８５６－８５３ページ．等を参照）。

【０１４７】

当分野で周知のＰＣＲ法をベースにした種々の方法で、部分的なヌクレオチド配列を利用して、ＴＲＩＣＨをコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば、使用し得る方法の１つである制限部位ＰＣＲ法は、ユニバーサルプライマー及びネステッドプライマーを用いてクローニングベクター内のゲノムＤＮＡから未知の配列を増幅する方法である（例えば、Ｓａｒｋａｒ，Ｇ．（１９９３）ＰＣＲ Ｍｅｔｈｏｄｓ Ａｐｐｌｉｃ．２：３１８３２２を参照。）別の方法に逆ＰＣＲ法があり、これは広範な方向に伸長させたプライマーを用いて環状化した鋳型から未知の配列を増幅する方法である。鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限酵素断片から得る（例えば、Ｔｒｉｇｌｉａ，Ｔ．他（１９８８）Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．１６：８１８６。）第３の方法としてキャプチャＰＣＲ法があり、これはヒト及び酵母菌人工染色体ＤＮＡの既知の配列に隣接するＤＮＡ断片をＰＣＲ増幅する方法に関与している。（Ｌａｇｅｒｓｔｒｏｍ，Ｍ．ら（１９９１）ＰＣＲ Ｍｅｔｈｏｄｓ Ａｐｐｌｉｃ １：１１１－１１９等を参照）。．）

この方法では、ＰＣＲを行う前に複数の制限酵素の消化及び連結反応を用いて未知の配列領域内に組換え二本鎖配列を挿入することが可能である。また、未知の配列を検索するために用い得る別の方法については当分野で知られている。（Ｐａｒｋｅｒ，Ｊ．Ｄ．ら

（１９９１）Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．１９：３０５５－３０６０等を参照）。更に、ＰＣＲ、ネステッドプライマー及びＰｒｏｍｏｔｅｒＦｉｎｄｅｒ（商標）ライブラリ（Ｃｌｏｎｔｅｃｈ，Ｐａｌｏ Ａｌｔｏ ＣＡ）を用いてゲノムＤＮＡをウォーキングすることができる。この手順は、ライブラリをスクリーニングする必要がなく、イントロン／エキソン接合部を見付けるのに有用である。全てのＰＣＲベースの方法に対して、市販されているソフトウェア、例えばＯＬＩＧＯ ４．０６プライマー分析ソフトウェア（Ｎａｔｉｏｎａｌ Ｂｉｏｓｃｉｅｎｃｅｓ，Ｐｌｙｍｏｕｔｈ ＭＮ）或いは別の好適なプログラムを用いて、長さが約２２～３０ヌクレオチド、ＧＣ含有率が約５０％以上、温度約６８℃～７２℃で鋳型に対してアニーリングするようにプライマーを設計し得る。

【０１４８】

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

【０１４９】

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

【０１５０】

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。

【０１５１】

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。遺伝子断片及び合成オリゴヌクレオチドのランダムなフラグメンテーション及びＰＣＲ再アセンブリによるＤＮＡシャッフリングを用い、ヌクレオチド配列を組み換えることが可能である。例えば、オリゴヌクレオチド仲介特定部位突然変異誘導を利用して、新規な制限部位の生成、グリコシル化パターンの変更、コドン優先の変更、スプライス変異体の生成等を行う突然変異を導入し得る。

【０１５２】

本発明のヌクレオチドを、ＭＯＬＥＣＵＬＡＲＢＲＥＥＤＩＮＧ（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；Ｃｒａｍｅｒｉ，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シャッフリングは、遺伝子断片のＰＣＲ仲介再組換えを用いて遺伝子変異体のライブラリを生成するプロセスである。ライブラリはその後、所望の特性を有する遺伝子変異体を同定するための選択またはスクリーニングにかける。次にこれらの好適な変異体をプールし、更に反復してＤＮＡシャッフリング及び選択／スクリーニングを行ってもよい。このように、遺伝の多様性は「人為的」品種改良及び急速な分子の進化を経て創生される。例えば、ランダムポイント突然変異を有する単一の遺伝子の断片を組み換えて、スクリーニングし、その後所望の特性が最適化されるまでシャッフリングすることができる。或いは、所定の遺伝子の断片を同種または異種のいずれかから得た同一遺伝子ファミリーの相同遺伝子と組み換えて、それによって天然に存在する複数の遺伝子の遺伝多様性を、指図された制御可能な方法で最大化させることができる。

【０１５３】

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、種々の液相または固相技術を用いてペプチド合成を行うことができる（たとえば、Ｃｒｅｉｇｈｔｏｎ，Ｔ．（１９８４）Ｐｒｏｔｅｉｎｓ，Ｓｔｒｕｃｔｕｒｅｓ ａｎｄ Ｍｏｌｅｃｕｌａｒ Ｐｒｏｐｅｒｔｉｅｓ，ＷＨ Ｆｒｅｅｍａｎ，Ｎｅｗ Ｙｏｒｋ ＮＹ，５５－６０ページ；Ｒｏｂｅｒｇｅ，Ｊ．Ｙ．他（１９９５）Ｓｃｉｅｎｃｅ ２６９：２０２２０４を参照。）自動合成はＡＢＩ ４３１Ａペプチドシンセサイザ（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）を用いて達成し得る。更にＴＲＩＣＨのアミノ酸配列または任意のその一部を用い、直接的な合成の際の変更、及び／または化学的方法を用いた他のタンパク質または任意のその一部からの配列との組み合わせにより、天然のポリペプチド配列を有するポリペプチドまたは変異体ポリペプチドを作製することが可能である。

【０１５４】

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

【０１５５】

生物学的に活性なＴＲＩＣＨを発現させるために、ＴＲＩＣＨをコードするヌクレオチド配列またはその誘導体を好適な発現ベクターに挿入する。この発現ベクターは、好適な宿主に挿入されたコーディング配列の転写及び翻訳の調節に必要なエレメントを含む。これらのエレメントには、ベクター及びＴＲＩＣＨをコードするポリヌクレオチド配列におけるエンハンサー、構成型及び発現誘導型のプロモーター、５’及び３’の非翻訳領域などの調節配列が含まれる。このような要素は、長さ及び特異性が様々である。特定の開始シグナルによって、ＴＲＩＣＨをコードする配列のより効果的な翻訳を達成することが可能である。このようなシグナルには、ＡＴＧ開始コドンと、コザック配列などの近傍の配列が含まれる。ＴＲＩＣＨをコードする配列及びその開始コドン、上流の調節配列が好適な発現ベクターに挿入された場合は、更なる転写調節シグナルや翻訳調節シグナルは必要なくなるであろう。しかしながら、コーディング配列或いはその断片のみが挿入された場合は、インフレームのＡＴＧ開始コドンを含む外来性の翻訳調節シグナルが発現ベクターに含まれるようにすべきである。外来性の翻訳要素及び開始コドンは、様々な天然物及び合成物を起源とし得る。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である。（例えば、Ｓｃｈａｒｆ，Ｄ．他（１９９４）Ｒｅｓｕｌｔｓ Ｐｒｏｂｌ．Ｃｅｌｌ Ｄｉｆｆｅｒ．２０１－１８－１６２．

を参照）。

【０１５６】

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

【０１５７】

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

【０１５８】

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は、クローニングされた

配列におけるｉｎ ｖｉｔｒｏ転写、ジデオキシのシークエンシング、ヘルパーファージによる一本鎖の救出、入れ子状態の欠失の生成にも有用であろう（例えば、Ｖａｎ Ｈｅｅｋｅ，Ｇ．およびＳ．Ｍ．Ｓｃｈｕｓｔｅｒ（１９８９）Ｊ．Ｂｉｏｌ．Ｃｈｅｍ．２６４：５５０３５５０９を参照）。例えば、抗体の産生のためなどに多量のＴＲＩＣＨが必要な場合は、ＴＲＩＣＨの発現をハイレベルで誘導するベクターが使用できる。例えば、強力な誘導ＳＰ６バクテリオファージプロモーターまたは誘導Ｔ７バクテリオファージプロモーターを含むベクターが使用できる。

【０１５９】

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

【０１６０】

植物系もＴＲＩＣＨの発現に使用可能である。ＴＲＩＣＨをコードする配列の転写は、ウイルスプロモーター、例えば単独或いはＴＭＶ（Ｔａｋａｍａｔｓｕ，Ｎ．（１９８７）ＥＭＢＯ Ｊ ６：３０７－３１１）由来のオメガリーダー配列と組み合わせて用いられるようなＣａＭＶ由来の３５Ｓ及び１９Ｓプロモーターによって促進される。或いは、ＲＵＢＩＳＣＯの小サブユニット等の植物プロモーターまたは熱ショックプロモーターを用いてもよい（例えば、Ｃｏｒｕｚｚｉ，Ｇ．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）これらの構成物は、直接ＤＮＡ形質転換または病原体を媒介とする形質移入によって、植物細胞内に導入可能である。（『マグローヒル科学技術年鑑』（Ｔｈｅ ＭｃＧｒａｗ Ｈｉｌｌ Ｙｅａｒｂｏｏｋ ｏｆ Ｓｃｉｅｎｃｅ ａｎｄ Ｔｅｃｈｎｏｌｏｇｙ）（１９９２）ＭｃＧｒａｗ Ｈｉｌｌ Ｎｅｗ Ｙｏｒｋ ＮＹ，１９１－１９６ページ等を参照。

【０１６１】

哺乳動物細胞においては、多数のウイルスベースの発現系を利用し得る。アデノウイルスが発現ベクターとして用いられる場合、後発プロモーター及び３連リーダー配列からなるアデノウイルス転写物／翻訳複合体にＴＲＩＣＨをコードする配列を結合し得る。可欠Ｅ１またはＥ３領域へウイルスのゲノムを挿入し、宿主細胞でＴＲＩＣＨを発現する感染ウイルスを得ることが可能である。（例えば、Ｌｏｇａｎ，Ｊ．およびＴ．Ｓｈｅｎｋ（１９８４）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８１：３６５５３６５９を参照）。更に、ラウス肉腫ウイルス（ＲＳＶ）エンハンサー等の転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させ得る。ＳＶ４０またはＥＢＶをベースにしたベクターを用いてタンパク質を高レベルで発現させることもできる。

【０１６２】

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

【０１６３】

哺乳動物系の組換えタンパク質の長期にわたる産生のためには、株化細胞におけるＴＲＩＣＨの安定した発現が望ましい。例えば、発現ベクターを用いて、ＴＲＩＣＨをコードする配列を株化細胞に形質転換することが可能である。 このような発現ベクターは、ウイ

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

【０１６４】

任意の数の選択系を用いて、形質転換細胞株を回収できる。限定するものではないがこのような選択系には、ｔｋ－単純細胞のために用いられるヘルペスウイルスチミジンキナーゼ遺伝子と、ａｐｒ－細胞のために用いられるアデニンホスホリボシルトランスフェラーゼ遺伝子がある（例えば、Ｗｉｇｌｅｒ，Ｍ．他（１９７７）Ｃｅｌｌ １１：２２３－２３２；及びＬｏｗｙ，Ｉ．他（１９８０）Ｃｅｌｌ ２２：８１７－８２３を参照）。また、選択の基礎として代謝拮抗物質、抗生物質或いは除草剤への耐性を用いることができる。例えば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はクロルスルフロンに対する耐性を、ｐａｔはホスフィノトリシンアセチルトランスフェラーゼに対する耐性を各々与える（Ｗｉｇｌｅｒ，Ｍ．ら（１９８０）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ７７：３５６７３５７０；ＣｏｌｂｅｒｅＧａｒａｐｉｎ，Ｆ．ら（１９８１）Ｊ．Ｍｏｌ．Ｂｉｏｌ．１５０：１１４ 等を参照。）この他の選択可能な遺伝子、例えば、代謝のための細胞の必要条件を変えるｔｒｐＢ及びｈｉｓＤは、文献に記載されている（Ｈａｒｔｍａｎ，Ｓ．Ｃ．およびＲ．Ｃ．Ｍｕｌｌｉｇａｎ（１９８８）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８５：８０４７８０５１を参照。）可視マーカー、例えばアントシアニン、緑色蛍光タンパク質（ＧＦＰ；Ｃｌｏｎｔｅｃｈ）、βグルクロニダーゼ及びその基質βグルクロニド、またはルシフェラーゼ及びその基質ルシフェリン等を用いてもよい。これらのマーカーを用いて、トランスフォーマントを特定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することが可能である（Ｒｈｏｄｅｓ，Ｃ．Ａ．（１９９５）Ｍｅｔｈｏｄｓ Ｍｏｌ．Ｂｉｏｌ．５５：１２１１３１等を参照）。

【０１６５】

マーカー遺伝子発現の存在／不存在によって目的の遺伝子の存在が示唆されても、その遺伝子の存在及び発現の確認が必要な場合もある。例えば、ＴＲＩＣＨをコードする配列がマーカー遺伝子配列の中に挿入された場合、ＴＲＩＣＨをコードする配列を含む形質転換された細胞は、マーカー遺伝子機能の欠落により同定可能である。または、１つのプロモーターの制御下でマーカー遺伝子がＴＲＩＣＨをコードする配列と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。

【０１６６】

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

【０１６７】

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な方法は、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 ｉｍｍｕｎｏａ

ｓｓａｙ）が好ましいが、競合の結合アッセイも用いることもできる。これらのアッセイ及びこれ以外のアッセイは、当分野で公知である（Ｈａｍｐｔｏｎ．Ｒ．ら（１９９０）Ｓｅｒｏｌｏｇｉｃａｌ Ｍｅｔｈｏｄｓ，ａ Ｌａｂｏｒａｔｏｒｙ Ｍａｎｕａｌ．ＡＰＳ Ｐｒｅｓｓ．Ｓｔ Ｐａｕｌ．ＭＮ，Ｓｅｃｔ．ＩＶ、Ｃｏｌｉｇａｎ，Ｊ．Ｅ．ら（１９９７）Ｃｕｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｉｍｍｕｎｏｌｏｇｙ，Ｇｒｅｅｎｅ Ｐｕｂ．Ａｓｓｏｃｉａｔｅｓ ａｎｄ Ｗｉｌｅｙ－Ｉｎｔｅｒｓｃｉｅｎｃｅ，Ｎｅｗ Ｙｏｒｋ ＮＹ、Ｐｏｕｎｄ，Ｊ．Ｄ．（１９９８）Ｉｍｍｕｎｏｃｈｅｍｉｃａｌ Ｐｒｏｔｏｃｏｌｓ，Ｈｕｍａｎｓ Ｐｒｅｓｓ，Ｔｏｔｏｗａ ＮＪ等を参照）。

【０１６８】

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

【０１６９】

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

【０１７０】

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

【０１７１】

本発明の別の実施例では、ＴＲＩＣＨをコードする自然或いは変更された、または組換えの核酸配列を上記した任意の宿主系の融合タンパク質の翻訳となる異種配列に結合させる。例えば、市販の抗体によって認識できる異種部分を含むキメラＴＲＩＣＨタンパク質が、ＴＲＩＣＨの活性のインヒビターに対するペプチドライブラリのスクリーニングを促進し得る。また、異種タンパク質部分及び異種ペプチド部分も、市販されている親和性基質を用いて融合タンパク質の精製を促進し得る。限定されるものではないがこのような部分には、グルタチオンＳトランスフェラーゼ（ＧＳＴ）、マルトース結合タンパク質（ＭＢ

Ｐ）、チオレドキシン（Ｔｒｘ）、カルモジュリン結合ペプチド（ＣＢＰ）、６－Ｈｉｓ、ＦＬＡＧ、ｃ－ｍｙｃ、赤血球凝集素（ＨＡ）がある。ＧＳＴは固定化グルタチオン上で、ＭＢＰはマルトース上で、Ｔｒｘはフェニルアルシンオキシド上で、ＣＢＰはカルモジュリン上で、そして６－Ｈｉｓは金属キレート樹脂上で、同族の融合タンパク質の精製を可能にする。ＦＬＡＧ、ｃ－ｍｙｃ及び赤血球凝集素（ＨＡ）は、これらのエピトープ標識を特異的に認識する市販されているモノクローナル抗体及びポリクローナル抗体を用いて、融合タンパク質の免疫親和性精製を可能にする。また、ＴＲＩＣＨをコードする配列と異種タンパク質配列との間にあるタンパク質分解切断部位を融合タンパク質が含むように遺伝子操作すると、ＴＲＩＣＨが精製の後に異種部分から切断され得る。融合タンパク質の発現及び精製方法は、前出のＡｕｓｕｂｅｌ（１９９５）１０章に記載されている。市販されている種々のキットを用いて融合タンパク質の発現及び精製を促進することもできる。

【０１７２】

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

【０１７３】

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。

【０１７４】

一実施例では、このように同定された化合物は、例えばリガンドやその断片などのＴＲＩＣＨの天然のリガンド、または天然の基質、構造的または機能的な擬態性または自然結合パートナーに密接に関連している（Ｃｏｌｉｇａｎ，Ｊ．Ｅ．他（１９９１）Ｃｕｒｒｅｎｔ Ｐｒｏｔｏｃｏｌｓ ｉｎ Ｉｍｍｕｎｏｌｏｇｙ １（２）の５章等を参照）。同様に、化合物は、ＴＲＩＣＨが結合する天然受容体、或いは例えばリガンド結合部位などの少なくとも受容体のある断片に密接に関連し得る。何れの場合も、既知の技術を用いてこの化合物を合理的に設計することができる。一実施例では、このような化合物に対するスクリーニングには、分泌タンパク質或いは細胞膜上のタンパク質の何れか一方としてＴＲＩＣＨを発現する好適な細胞の作製が含まれる。好適な細胞には、哺乳動物、酵母、ショウジョウバエ、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。

【０１７５】

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、蛍光色素、放射性同位体、酵素抱合体またはその他の検出可能な標識により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、無細胞再構成標本、化学ライブラリまたは天然の生成混合物を用いて実施することができ、試験化合物は、溶液中で遊離させるか固体支持体に固定させる。

【０１７６】

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の活性が許容される条件下でアッセイを実施し、そのアッセイでは少なくとも１つの試験化合物を

ＴＲＩＣＨと混合し、試験化合物の存在下でＴＲＩＣＨの活性を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。

【０１７７】

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

【０１７８】

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）。

【０１７９】

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

【０１８０】

（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。更にＴＲＩＣＨの発現は、脳、肺、前立腺、膀胱、骨、視床下部、乳房、回腸、胃、膵臓、および胃腸の組織および脳と前立腺の腫瘍に密接に関連する。従って、ＴＲＩＣＨは、輸送障害、神経の疾患、筋疾患、及び免疫疾患や細胞異常増殖においてある役割を果たすと考えられる。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。

【０１８１】

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、このような疾患のうち、輸送障害には運動不能症、筋萎縮性側索硬化症、毛細血管拡張性運動失調、ベッカー筋ジストロフィー、顔面麻痺、シャルコー‐マリー‐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、トゥーレット病、妄想性精神病、及び分裂病などの輸送に関連した神経障害と、神経線維腫症、帯状疱疹後神経痛、３叉神経ニューロパシー、サルコイドーシス、鎌状赤血球性貧血、ウィルソン病、白内障、不妊症、肺動脈狭窄症、常染色体性感音性難聴、高／低血糖症、グレーブス病、甲状腺腫、クッシング病と、副腎機能不全、グルコース‐ガラクトース吸収不全症候群、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、ウイルス性中枢神経系疾患と、クールー及びクロイツフェルト‐ヤコブ病、ゲルストマン症候群、Ｇｅｒｓｔｍａｎｎ－Ｓｔｒａｕｓｓｌｅｒ－Ｓｃｈｅｉｎｋｅｒ症候群を含むプリオン病（ｐｒｉｏｎ ｄｉｓｅａｓｅ）と、致死性家族性不眠症、神経系性栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜血管芽腫（ｃｅｒｅｂｅｌｌｏｒｅｔｉｎａｌ ｈｅｍａｎｇｉｏｂｌａｓｔｏｍａｔｏｓｉｓ）、脳３叉神経血管症候群、ダウン症を含む中枢神経系性精神薄弱及び他の発生障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィー及び他の神経筋障害、末梢神経疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性、及び中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分性及び不安性精神障害、及び妄想性精神病と、季節性の感情の障害（ＳＡＤ）、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、帯状疱疹後神経痛、及びトゥーレット病が含まれ、筋疾患の中には、心筋症、心筋炎、デュシェンヌ型筋ジストロフィー、ベッカー型偽肥大性筋ジストロフィー、筋緊張性ジストロフィー、中心コア病、ネマリンミオパシーラネマリン筋障害、中心核ミオパシー、脂質ミオパシー、ミトコンドリアミオパチー、感染性節炎、多発性筋炎、皮膚筋炎、封入体筋炎、甲状腺中毒性ミオパシー、エタノールミオパシー（ｅｔｈａｎｏｌ ｍｙｏｐａｔｈｙ）、口峡炎、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性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎症、骨関節炎、骨粗しょう症、膵炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷、および細胞異常増殖には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病（ＭＣＴＤ）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頚部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌が含まれる。

【０１８２】

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。

【０１８３】

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。

【０１８４】

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。

【０１８５】

更なる実施例では、ＴＲＩＣＨの発現または活性の増大に関連した疾患の治療または予防のために、患者にＴＲＩＣＨのアンタゴニストを投与することが可能である。限定するものではないが、このような疾患の例には、上記した輸送障害、神経疾患、筋疾患、免疫疾患および細胞の増殖異常が含まれる。一実施態様では、ＴＲＩＣＨと特異的に結合する抗体が直接アンタゴニストとして、或いはＴＲＩＣＨを発現する細胞または組織に薬剤を運ぶターゲッティング或いは運搬機構として間接的に用いられ得る。

【０１８６】

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。

【０１８７】

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

【０１８８】

ＴＲＩＣＨのアンタゴニストは、当分野で一般的な方法を用いて製造することが可能である。 詳しくは、精製されたＴＲＩＣＨを用いて抗体を作ったり、治療薬のライブラリをスクリーニングしてＴＲＩＣＨと特異的に結合するものを同定が可能である。ＴＲＩＣＨの抗体も、当分野で一般的な方法を用いて製造することが可能である。 このような抗体には、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、一本鎖、Ｆａｂフラグメント、及びＦａｂ発現ライブラリによって作られたフラグメントが含まれる。 但し、こ

れらに限定されるものではない。中和抗体（即ち二量体の形成を阻害する抗体）は通常、治療用に好適である。

【０１８９】

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

【０１９０】

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

【０１９１】

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

【０１９２】

更に、「キメラ抗体」作製のために発達したヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの技術が、好適な抗原特異性及び生物学的活性を備える分子を得るために用いられる（例えば、Ｍｏｒｒｉｓｏｎ，Ｓ．Ｌ．他．（１９８４）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．８１：６８５１６８５５；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。関連特異性を有するがイディオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリからチェーンシャッフリングによって産生することもできる（Ｂｕｒｔｏｎ Ｄ．Ｒ．（１９９１）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８８：１０１３４－１０１３７等を参照）。

【０１９３】

抗体の産生は、リンパ球集団における<u>ｉｎ ｖｉｖｏ</u>産生の誘導によって、或いは免疫グロブリンライブラリのスクリーニングまたは文献に開示されているような高特異結合試薬のパネルのスクリーニングによっても行い得る（Ｏｒｌａｎｄｉ，Ｒ．ら（１９８９）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ ８６：３８３３－３８３７、Ｗｉｎｔｅｒ，Ｇ．ら（１９９１）Ｎａｔｕｒｅ ３４９：２９３－２９９等を参照）。

【０１９４】

ＴＲＩＣＨに対する特異的な結合部位を含む抗体も得ることができる。例えば、限定するものではないが、このような断片には、抗体分子のペプシン消化によって作製されるＦ（

ａｂ’）$_2$ 断片と、Ｆ（ａｂ’）$_2$ 断片のジスルフィド架橋を還元することによって作製されるＦａｂ断片とがある。或いは、Ｆａｂ発現ライブラリを作製することによって、モノクローナルＦａｂ断片を所望の特異性と迅速且つ容易に同定することが可能となる（Ｈｕｓｅ，Ｗ．Ｄ．他（１９８９）Ｓｃｉｅｎｃｅ ２５６：１２７５－１２８１等を参照）。

【０１９５】

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

【０１９６】

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10^9～10^{12}ｌｉｔｅｒ／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10^6～10^7ｌｉｔｅｒ／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）。

【０１９７】

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

【０１９８】

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（ＤＮＡ及びＲＮＡ、修飾ヌクレオチド）を設計して遺伝子発現を変更することができる。このような技術は当分野では周知であり、センスまたはアンチセンスオリゴヌクレオチドまたは大きな断片が、ＴＲＩＣＨをコードする配列の制御領域から、またはコード領域に沿ったさまざまな位置から設計可能である（Ａｇｒａｗａｌ，Ｓ．，編集（１９９６）Ａｎｔｉｓｅｎｓｅ Ｔｈｅｒａｐｅｕｔｉｃｓ，Ｈｕｍａｎａ Ｐｒｅｓｓ Ｉｎｃ．，Ｔｏｔａｗａ ＮＪを参照）。

【０１９９】

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輸送することが可能である。（Ｓｌａｔｅｒ，Ｊ．Ｅ．他（１９９８）Ｊ．Ａｌｌｅｒｇｙ Ｃｌｉｎ．Ｉｍｍｕｎｏｌ．１０２（３）：４６９－４７５；およびＳｃａｎｌｏｎ，Ｋ．Ｊ．他（１９９５）９（１３）：１２８８－１２９６を参照。）アンチセンス配列はまた、例えばレトロウイルスやアデノ関連ウイルスベクター等のウイルスベクターを用いて細胞内に導入することもできる（Ｍｉｌｌｅｒ，Ａ．Ｄ．（１９９０）Ｂｌｏｏｄ ７６：２７１、前出のＡｕｓｕｂｅｌ、Ｕｃｋｅｒｔ，Ｗ．ａｎｄ Ｗ．Ｗａｌｔｈｅｒ（１９９４）Ｐｈａｒｍａｃｏｌ．Ｔｈｅｒ．６３（３）：３２３－３４７等を参照）。その他の遺伝送達機構には、リポソーム系、人工的なウイルスエンベロープ及び当分野で公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．ら（１９９８）Ｊ．Ｐｈａｒｍ．Ｓｃｉ．８７（１１）：１３０８－１３１５、Ｍｏｒｒｉｓ，Ｍ．Ｃ．ら（１９９７）Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．２５（１４）：２７３０－２７３６．等を参照）。

【０２００】

本発明の別の実施例では、ＴＲＩＣＨをコードするポリヌクレオチドを、体細胞若しくは生殖細胞の遺伝子治療に用いることが可能である。遺伝子治療を行うことにより、（ｉ）遺伝子欠損症（例えばＸ染色体鎖遺伝（Ｃａｖａｚｚａｎａ－Ｃａｌｖｏ，Ｍ．ら（２０００）Ｓｃｉｅｎｃｅ ２８８：６６９－６７２）により特徴付けられる重度の複合型免疫欠損（ＳＣＩＤ）－Ｘ１の場合）、先天性アデノシンデアミナーゼ（ＡＤＡ）欠損症に関連する重度の複合型免疫欠損（Ｂｌａｅｓｅ，Ｒ．Ｍ．ら（１９９５）Ｓｃｉｅｎｃｅ ２７０：４７５－４８０、Ｂｏｒｄｉｇｎｏｎ，Ｃ．ら（１９９５）Ｓｃｉｅｎｃｅ ２７０：４７０－４７５）、嚢胞性繊維症（Ｚａｂｎｅｒ，Ｊ．ら（１９９３）Ｃｅｌｌ ７５：２０７－２１６：Ｃｒｙｓｔａｌ、Ｒ．Ｇ．ら（１９９５）Ｈｕｍ．Ｇｅｎｅ Ｔｈｅｒａｐｙ ６：６４３－６６６、Ｃｒｙｓｔａｌ，Ｒ．Ｇ．ら．（１９９５）Ｈｕｍ．Ｇｅｎｅ Ｔｈｅｒａｐｙ ６：６６７－７０３）、サラセミア（ｔｈａｌａｓｓａｍｉａ）、家族性高コレステロール血症、第ＶＩＩＩ因子若しくは第ＩＸ因子欠損に起因する血友病（Ｃｒｙｓｔａｌ，３５ Ｒ．Ｇ．（１９９５）Ｓｃｉｅｎｃｅ ２７０：４０４－４１０、Ｖｅｒｍａ，Ｉ．Ｍ．およびＮ．Ｓｏｍｉａ（１９９７）Ｎａｔｕｒｅ ３８９：２３９－２４２）を治療し、（ｉｉ）条件的致死性遺伝子産物を発現させ（例えば制御不能な細胞増殖に起因する癌の場合）、（ｉｉｉ）細胞内の寄生虫（例えばヒト免疫不全ウイルス（ＨＩＶ）（Ｂａｌｔｉｍｏｒｅ，Ｄ．（１９８８）Ｎａｔｕｒｅ ３３５：３９５－３９６、Ｐｏｅｓｃｂｌａ，Ｅ．ら（１９９６）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ．９３：１１３９５－１１３９９）、Ｂ型若しくはＣ型肝炎ウイルス（ＨＢＶ、ＨＣＶ）、Ｃａｎｄｉｄａ ａｌｂｉｃａｎｓ及びＰａｒａｃｏｃｃｉｄｉｏｉｄｅｓ ｂｒａｓｉｌｉｅｎｓｉｓ等の真菌寄生虫、並びにＰｌａｓｍｏｄｉｕｍ ｆａｌｃｉｐａｒｕｍ及びＴｒｙｐａｎｏｓｏｍａ ｃｒｕｚｉ等の原虫寄生体に対する防御機能を有するタンパク質を発現させることができる。ＴＲＩＣＨの発現若しくは調節に必要な遺伝子の欠損が疾患を引き起こす場合、導入した細胞の好適な集団からＴＲＩＣＨを発現させて、遺伝子欠損によって起こる症状の発現を緩和することが可能である。

【０２０１】

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導入技術には、（ｉ）個々の細胞内への直接的なＤＮＡ微量注射法、（ｉｉ）遺伝子銃、（ｉｉｉ）リポソームを介した形質移入、（ｉｖ）受容体を介した遺伝子導入、及び（ｖ）ＤＮＡトランスポゾンの使用（Ｍｏｒｇａｎ，Ｒ．Ａ．および

W. F. Anderson（1993）Annu. Rev. Biochem. 62：191－217、Ivics, Z.（1997）Cell 91：501－510；Boulay, J－L. およびH. Recipon（1998）Curr. Opin. Biotechnol. 9：445－450）がある。

【0202】

限定するものではないがTRICHの発現に影響を及ぼし得る発現ベクターには、PCDNA 3.1、EPITAG、PRCCMV2、PREP、PVAX、PCR2－TOPOTAベクター（Invitrogen, Carlsbad CA）、PCMV－SCRIPT、PCMV－TAG、PEGSH／PERV（Stratagene, La Jolla CA）及びPTET－OFF、PTET－ON、PTRE2、PTRE2－LUC、PTK－HYG（Clontech, Palo Alto CA）がある。TRICHを発現させるために、（i）恒常的に活性なプロモーター（例えば、サイトメガロウイルス（CMV）、ラウス肉腫ウイルス（RSV）、SV40ウイルス、チミジンキナーゼ（TK）、若しくはβ－アクチン遺伝子等）、（ii）誘導性プロモーター（例えば、市販されているT－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／ミフェプリストーン誘導性プロモーター（Rossi, F. M. V. and H. M. Blau, 前出）、または（iii）正常な個体に由来する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。

【0203】

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

【0204】

本発明の別の実施例では、TRICHの発現に関連する遺伝子欠損によって起こる疾患や異常症は、（i）レトロウイルス末端反復配列（LTR）プロモーター若しくは独立したプロモーターのコントロール下でTRICHをコードするポリヌクレオチドと、（ii）好適なRNAパッケージングシグナルと、（iii）追加のレトロウイルス・シス作用性RNA配列及び効率的なベクターの増殖に必要なコーディング配列を伴うRev応答性エレメント（RRE）とからなるレトロウイルスベクターを作製して治療することができる。レトロウイルスベクター（例えばPFB及びPFBNEO）はStratagene社から市販されており、刊行データ（Riviere, I. ら.（1995）Proc. Natl. Acad. Sci. U. S. A. 92：6733－6737）に基づいている。 上記データを引用することをもって本明細書の一部とする。ベクターは、好適なベクター産生細胞系（VPCL）において増殖され、VPCLは、標的細胞上の受容体に対する親和性を有するエンベロープ遺伝子またはVSVg等の 汎親和性エンベロープタンパク質を発現する（Armentano, D. ら（1987）J. Virol. 61：1647－1650、Bender, M. A. ら（1987）J. Vi

ｒｏｌ．６１：１６３９－１６４６、Ａｄａｍ，Ｍ．Ａ．およびＡ．Ｄ．Ｍｉｌｌｅｒ（１９８８）Ｊ．Ｖｉｒｏｌ．６２：３８０２－３８０６、Ｄｕｌｌ，Ｔ．ら（１９９８）Ｊ．Ｖｉｒｏｌ．７２：８４６３－８４７１、Ｚｕｆｆｅｒｅｙ，Ｒ．ら（１９９８）Ｊ．Ｖｉｒｏｌ．７２：９８７３－９８８０）。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、細胞集団（例えばＣＤ４$^{+}$ Ｔ細胞）の形質導入、及び形質導入した細胞の患者への戻しは、遺伝子治療の分野では当業者に公知の方法であり、多数の文献に記載されている（Ｒａｎｇａ，Ｕ．ら．（１９９７）Ｊ．Ｖｉｒｏｌ．７１：７０２０－７０２９、Ｂａｕｅｒ，Ｇ．ら（１９９７）Ｂｌｏｏｄ ８９：２２５９－２２６７、Ｂｏｎｙｈａｄｉ，Ｍ．Ｌ．（１９９７）Ｊ．Ｖｉｒｏｌ．７１：４７０７－４７１６、Ｒａｎｇａ，Ｕ．ら（１９９８）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．Ｕ．Ｓ．Ａ．９５：１２０１－１２０６、Ｓｕ，Ｌ．（１９９７）Ｂｌｏｏｄ ８９：２２８３－２２９０）。

【０２０５】

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

【０２０６】

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、当業者に公知である。 複製適格性単純ヘルペスウイルス（ＨＳＶ）Ｉ型系のベクターは、レポーター遺伝子を霊長類の眼に送達するために用いられてきた（Ｌｉｕ，Ｘ．ら（１９９９）Ｅｘｐ．Ｅｙｅ Ｒｅｓ．１６９：３８５－３９５）。ＨＳＶ－１ウイルスベクターの作製についても、ＤｅＬｕｃａに付与された米国特許第５，８０４，４１３号（「Ｈｅｒｐｅｓ ｓｉｍｐｌｅｘ ｖｉｒｕｓ ｓｗａｉｎｓ ｆｏｒ ｇｅｎｅ ｔｒａｎｓｆｅｒ」）に開示されており、該特許の引用をもって本明細書の一部とする。 米国特許第５，８０４，４１３号には、ヒト遺伝子治療を含む目的のために好適なプロモーターの制御下において細胞に導入される少なくとも１つの外在性遺伝子を有するゲノムを含む組換えＨＳＶ ｄ９２についての記載がある。上記特許はまた、ＩＣＰ４、ＩＣＰ２７及びＩＣＰ２２のために除去される組換えＨＳＶ系統の作製及び使用について開示している。ＨＳＶベクターについては、Ｇｏｉｎｓ，Ｗ．Ｆ．ら（１９９９）Ｊ．Ｖｉｒｏｌ．７３：５１９－５３２及びＸｕ，Ｈ．ら（１９９４）Ｄｅｖ．Ｂｉｏｌ．１６３：１５２－１６１も参照されたい。 両文献は、引用をもって本明細書の一部とする。クローン化ヘルペスウイルス配列の操作、巨大ヘルペスウイルスのゲノムの異なった

部分を含む多数のプラスミドを形質移入した後の組換えウイルスの産生、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は、当業者に公知の技術である。

【０２０７】

別法では、αウイルス（正の一本鎖ＲＮＡウイルス）ベクターを用いてＴＲＩＣＨをコードするポリヌクレオチドを標的細胞に送達する。プロトタイプのαウイルスであるセムリキ森林熱ウイルス（Ｓｅｍｌｉｋｉ Ｆｏｒｅｓｔ Ｖｉｒｕｓ，ＳＦＶ）の生物学的研究が広範に行われており、遺伝子導入ベクターがＳＦＶゲノムに基づいていることが分かった（Ｇａｒｏｆｆ，Ｈ．およびＫ．－Ｊ．Ｌｉ（１９９８）Ｃｕｎ．Ｏｐｉｎ．Ｂｉｏｔｅｃｈ．９：４６４－４６９）。αウイルスＲＮＡの複製中に、通常はウイルスのキャプシッドタンパク質をコードするサブゲノムＲＮＡが作り出される。このサブゲノムＲＮＡは、完全長のゲノムＲＮＡより高いレベルに複製されるため、酵素活性（例えばプロテアーゼ及びポリメラーゼ）を有するウイルスタンパク質に比べてキャプシッドタンパク質が過剰産生される。同様に、ＴＲＩＣＨをコードする配列をαウイルスゲノムのカプシドをコードする領域に導入することによって、ベクター導入細胞において多数のＴＲＩＣＨをコードするＲＮＡが産生され、高いレベルでＴＲＩＣＨが合成される。通常はαウイルスの感染が数日以内での細胞溶解に関係する一方で、シンドビスウイルス（ＳＩＮ）の変異体を有するハムスター正常腎臓細胞（ＢＨＫ－２１）の持続的な感染を確立する能力は、αウイルスの溶解複製を遺伝子治療に適用できるように好適に変更可能であることを示唆している（Ｄｒｙｇａ，Ｓ．Ａ．ら．（１９９７）Ｖｉｒｏｌｏｇｙ ２２８：７４－８３）。様々な宿主にαウイルスを導入できることから、様々なタイプの細胞にＴＲＩＣＨを導入することできる。或る集団におけるサブセットの細胞の特定形質導入は、形質導入前に細胞の選別を必要とし得る。αウイルスの感染性ｃＤＮＡクローンの処置方法、αウイルスのｃＤＮＡ及びＲＮＡの形質移入方法及びαウイルスの感染方法は、当業者に公知である。

【０２０８】

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

【０２０９】

リボザイムは酵素性ＲＮＡ分子であり、ＲＮＡの特異的切断を触媒するためにリボザイムを用いることもできる。リボザイム作用のメカニズムは、相補的標的ＲＮＡへのリボザイム分子の配列特異性ハイブリダイゼーションおよびその後に起こる内ヌクレオチド鎖切断に関与している。例えば、遺伝子操作で作られたハンマーヘッド型リボザイム分子が、ＴＲＩＣＨをコードする配列の内ヌクレオチド鎖分解性の切断を特異的且つ効果的に触媒できる可能性がある。

【０２１０】

任意のＲＮＡ標的内の特異的リボザイム切断部位を、ＧＵＡ、ＧＵＵ、ＧＵＣ配列を含めたリボザイム切断部位に対する標的分子をスキャンすることによって先ず同定する。一度同定すると、切断部位を含む標的遺伝子の領域に対応する１５～２０リボヌクレオチドの短いＲＮＡ配列が、そのオリゴヌクレオチドを機能不全にするような２次構造の特徴をもっていないかを、評価することが可能になる。候補標的の適合性の評価も、リボヌクレア

ーゼ保護アッセイを用いて相補的オリゴヌクレオチドとのハイブリダイゼーションのアクセス可能性をテストすることによって行うことができる。

【０２１１】

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

【０２１２】

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

【０２１３】

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

【０２１４】

特異ポリヌクレオチドの変異発現における有効性に対して、少なくとも１個から複数個の試験化合物をスクリーニングし得る。試験化合物は、当分野で通常知られている任意の方法により得られる。 このような方法には、ポリヌクレオチドの発現を変異させる場合と、既存の、市販のまたは専売の、天然または非天然の化合物ライブラリから選択する場合と、標的ポリヌクレオチドの化学的及び／または構造的特性に基づく化合物を合理的にデザインする場合と、組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合に有効であることが知られているような化合物の化学修飾がある。ＴＲＩＣＨをコードするポリヌクレオチドを含むサンプルは、このようにして得られた試験化合物の少なくとも１つに曝露する。サンプルには例えば、無傷細胞、透過化処理した細胞、ｉｎ ｖｉｔｒｏ 無細胞系または再構成生化学系があり得る。ＴＲＩＣＨをコードするポリヌクレオチドの発現における変化は、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ョンにより、特定のヌクレオチドの発現を検出する。ハイブリダイゼーション量を定量し、それによって１つ以上の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較に対する基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現変化の検出は、ポリヌクレオチドの発現を変異する際に試

験化合物が有効であることを示している。特異ポリヌクレオチドの変異発現に有効な化合物に対して、例えばＳｃｈｉｚｏｓａｃｃｈａｒｏｍｙｃｅｓ ｐｏｍｂｅ遺伝子発現系（Ａｔｋｉｎｓ，Ｄ．ら（１９９９）米国特許第５，９３２，４３５号、Ａｒｎｄｔ，Ｇ．Ｍ．ら（２０００）Ｎｕｃｌｅｉｃ Ａｃｉｄｓ Ｒｅｓ．２８：Ｅ１５）またはＨｅＬａ細胞等のヒト細胞系（Ｃｌａｒｋｅ，Ｍ．Ｌ．ら（２０００）Ｂｉｏｃｈｅｍ．Ｂｉｏｐｈｙｓ．Ｒｅｓ．Ｃｏｍｍｕｎ．２６８：８－１３）を用いてスクリーニングを実行する。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性のためのオリゴヌクレオチド（デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾オリゴヌクレオチド）の組み合わせライブラリをスクリーニングすることに関与している（Ｂｒｕｉｃｅ，Ｔ．Ｗ．ら（１９９７）米国特許第５，６８６，２４２号、Ｂｒｕｉｃｅ，Ｔ．Ｗ．ら（２０００）米国特許第６，０２２，６９１号）。

【０２１５】

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

【０２１６】

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

【０２１７】

本発明の追加実施例は、通常薬剤として許容できる賦形剤で処方される活性成分を有する成分の投与に関連する。賦形剤には例えば、糖、でんぷん、セルロース、ゴム及びタンパク質がある。様々な処方が通常知られており、詳細はＲｅｍｉｎｇｔｏｎ’ｓ Ｐｈａｒｍａｃｅｕｔｉｃａｌ Ｓｃｉｅｎｃｅｓ（Ｍａａｃｋ Ｐｕｂｌｉｓｈｉｎｇ，Ｅａｓｔｏｎ ＰＡ）の最新版に記載されている。このような組成物は、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。

【０２１８】

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

【０２１９】

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

【０２２０】

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

【０２２１】

ＴＲＩＣＨまたはその断片を含む高分子を直接細胞内に輸送するべく、特殊な形態に組成物が調製されるのが好ましい。例えば、細胞不透過性高分子を含むリポソーム製剤は、細

胞融合及び高分子の細胞内送達を促進し得る。別法では、ＴＲＩＣＨまたはその断片をＨＩＶ Ｔａｔ－１タンパク質の陽イオンＮ末端部に結合することもできる。このようにして生成された融合タンパク質は、マウスモデル系の脳を含む全ての組織の細胞に形質導入することがわかっている（Ｓｃｈｗａｒｚｅ，Ｓ．Ｒ．ら（１９９９）Ｓｃｉｅｎｃｅ ２８５：１５６９－１５７２）。

【０２２２】

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

【０２２３】

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。治療有効度及び毒性は、細胞培養または動物実験における標準的な薬剤手法によって、例えばED_{50}（集団の５０％の医薬的有効量）またはLD_{50}（集団の５０％の致死量）を測定するなどして決定することができる。毒性効果の薬用効果に対する投与量の比は、治療指数であり、LD_{50}／ED_{50}比として表すことができる。高い治療指数を示すような成分が望ましい。細胞培養アッセイ及び動物実験から得られたデータは、ヒトに用いるための投与量の範囲を調剤するのに用いられる。このような組成物が含まれる投与量は、毒性を殆ど或いは全く含まず、ED_{50}を含むような血中濃度の範囲にあることが好ましい。用いられる投与形態、患者の感受性及び投与の経路によって、投与量はこの範囲内で様々に変わる。

【０２２４】

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

【０２２５】

通常の投与量は、投与の経路にもよるが約０．１～１００，０００μｇであり、合計で約１ｇまでとする。特定の投与量及び送達方法に関するガイダンスは文献に記載されており、現場の医者は通常それを利用することができる。 当業者は、タンパク質またはインヒビターに対する処方とは異なる、ヌクレオチドに対する処方を利用することになる。同様に、ポリヌクレオチドまたはポリペプチドの送達は、特定の細胞、状態、位置等に特異的なものとなる。

【０２２６】

（診断）

別の実施例では、ＴＲＩＣＨに特異的に結合する抗体が、ＴＲＩＣＨの発現によって特徴付けられる疾患の診断、またはＴＲＩＣＨやＴＲＩＣＨのアゴニストまたはアンタゴニスト、インヒビターで治療を受けている患者をモニターするためのアッセイに用いられる。診断目的に有用な抗体は、上記の治療の箇所で記載した方法と同じ方法で調合される。ＴＲＩＣＨの診断アッセイには、抗体及び標識を用いてヒトの体液或いは細胞や組織から採取されたものからＴＲＩＣＨを検出する方法が含まれる。この抗体は修飾されたものもされていないものも可能であり、レポーター分子との共有結合または非共有結合で標識化できる。レポーター分子としては広くさまざまな種類が本分野で知られており、また使用可能であるが、そのうちのいくつかは上記で説明されている。

【０２２７】

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。標準複合体形成量は、種々の方法、例えば測光法で定量できる。被験者のＴＲＩＣＨの発現の量、制御及び疾患、生検組織からのサンプルが基準値と比較される。標準値と被験者との偏差が疾患を診断するパラメータとなる。

【０２２８】

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。用いることができるポリヌクレオチドには、オリゴヌクレオチド配列、相補的ＲＮＡ及びＤＮＡ分子、そしてＰＮＡが含まれる。このポリヌクレオチドを用いて、疾患と相関し得る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。

【０２２９】

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

【０２３０】

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：２７－５２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。

【０２３１】

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

【０２３２】

ＴＲＩＣＨをコードするポリヌクレオチド配列を用いて、ＴＲＩＣＨの発現に関連する疾患を診断することが可能である。限定するものではないが、このような疾患のうち、輸送障害には運動不能症、筋萎縮性側索硬化症、毛細血管拡張性運動失調、ベッカー筋ジストロフィー、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、トゥーレット病、妄想性精神病、及び分裂病などの輸送に関連した神経障害と、神経線維腫症、帯状疱疹後神経痛、３叉神経ニューロパシー、サルコイドーシス、鎌状赤血球性貧血、ウィルソン病、白内障、不妊症、肺動脈狭窄症、常染色体性感音性難聴、高／低血糖症、グレーブス病、甲状腺腫、クッシング病と、副腎機能不全、グルコース‐ガラクトース吸収不全症候群、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、ウイルス性中枢神経系疾患と、クールー及びクロイツフェルト‐ヤコブ病、ゲルストマン症候群、Ｇｅｒｓｔｍａｎｎ－Ｓｔｒａｕｓｓｌｅｒ－Ｓｃｈｅｉｎｋｅｒ症候群を含むプリオン病（ｐｒｉｏｎ ｄｉｓｅａｓｅ）と、致死性家族性不眠症、神経系性栄養病及び代謝病、神経線維腫症、結節硬化症、小脳網膜血管芽腫（ｃｅｒｅｂｅｌｌｏｒｅｔｉｎａｌ ｈｅｍａｎｇｉｏｂｌａｓｔｏｍａｔｏｓｉｓ）、脳３叉神経血管症候群、ダウン症を含む中枢神経系性精神薄弱及び他の発生障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄病、筋ジストロフィー及び他の神経筋障害、末梢神経疾患、皮膚筋炎及び多発性筋炎と、遺伝性、代謝性、内分泌性、及び中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺と、気分性及び不安性精神障害、及び妄想性精神病と、季節性の感情の障害（ＳＡＤ）、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、錐体外路性終末欠陥症候群、ジストニー、分裂病性精神障害、帯状疱疹後神経痛、及びトゥーレット病が含まれ、筋疾患の中には、心筋症、心筋炎、デュシェンヌ型筋ジストロフィー、ベッカー型偽肥大性筋ジストロフィー、筋緊張性ジストロフィー、中心コア病、ネマリンミオパシーラネマリン筋障害、中心核ミオパシー、脂質ミオパシー、ミトコンドリアミオパチー、感染性節炎、多発性筋炎、皮膚筋炎、封入体筋炎、甲状腺中毒性ミオパシー、エタノールミオパシー（ｅｔｈａｎｏｌ ｍｙｏｐａｔｈｙ）、口峡炎、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性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎症、骨関節炎、骨粗しょう症、膵炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷、および細胞異常増殖には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病（ＭＣＴＤ）、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頚部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓

、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌が含まれる。ＴＲＩＣＨをコードするポリヌクレオチド配列は、サザーン法やノーザン法、ドットブロット法、或いはその他の膜系の技術、ＰＣＲ法、ディップスティック（ｄｉｐｓｔｉｃｋ）、ピン（ｐｉｎ）、ＥＬＩＳＡ式アッセイ、及び変異ＴＲＩＣＨの発現を検出するために患者から採取した体液或いは組織を利用するマイクロアレイに使用することが可能である。このような定性方法または定量方法は、当分野で公知である。

【０２３３】

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

【０２３４】

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

【０２３５】

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

【０２３６】

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

【０２３７】

ＴＲＩＣＨをコードする配列から設計されたオリゴヌクレオチドのさらなる診断への利用には、ＰＣＲの利用が含まれ得る。これらのオリゴマーは、化学的に合成するか、酵素により生産するか、或いはｉｎ ｖｉｔｒｏで産出し得る。オリゴマーは、好ましくはＴＲＩＣＨをコードするポリヌクレオチドの断片、或いはＴＲＩＣＨをコードするポリヌクレオチドと相補的なポリヌクレオチドの断片を含み、最適な条件の下、特定の遺伝子や条件を識別するために利用される。また、オリゴマーは、やや緩いストリンジェント条件下で、近縁のＤＮＡ或いはＲＮＡ配列の検出、定量、或いはその両方のため用いることが可能である。

【０２３８】

或る実施態様において、ＴＲＩＣＨをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いて、一塩基多型（ＳＮＰ）を検出し得る。ＳＮＰは、多くの場合にヒトの先天性または後天性遺伝病の原因となるような置換、挿入及び欠失である。限

定するものではないがＳＮＰの検出方法には、ＳＳＣＰ（ｓｉｎｇｌｅ－ｓｔｒａｎｄｅｄ ｃｏｎｆｏｒｍａｔｉｏｎ ｐｏｌｙｍｏｒｐｈｉｓｍ）及び蛍光ＳＳＣＰ（ｆＳＳＣＰ）法がある。ＳＳＣＰでは、ＴＲＩＣＨをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いたポリメラーゼ連鎖反応（ＰＣＲ）でＤＮＡを増幅する。ＤＮＡは例えば、病変組織または正常組織、生検サンプル、体液その他に由来し得る。ＤＮＡ内のＳＮＰは、一本鎖形状のＰＣＲ生成物の２次及び３次構造に差異を生じさせる。差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。ｆＳＣＣＰでは、オリゴヌクレオチドプライマーを蛍光性に標識する。それによってＤＮＡシークエンシング機などの高処理機器でアンプリマー（ａｍｐｌｉｍｅｒ）の検出が可能になる。更に、インシリコＳＮＰ（ｉｎ ｓｉｌｉｃｏ ＳＮＰ，ｉｓＳＮＰ）と呼ばれる配列データベース分析法は、一般的なコンセンサス配列に配列されるような個々の重畳するＤＮＡ断片の配列を比較することにより、多型性を同定し得る。これらのコンピュータベースの方法は、ＤＮＡの実験室での調整及び統計モデル及びＤＮＡ配列クロマトグラムの自動分析を用いたシークエンシングのエラーに起因する配列の変異をフィルタリングして除去する。別の態様では、例えば高処理ＭＡＳＳＡＲＲＡＹシステム（Ｓｅｑｕｅｎｏｍ，Ｉｎｃ．，Ｓａｎ Ｄｉｅｇｏ ＣＡ）を用いた質量分析によりＳＮＰを検出し、特徴付ける。

【０２３９】

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

【０２４０】

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

【０２４１】

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

【０２４２】

或る実施例は、或る組織または細胞タイプの転写イメージを生成するような本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞タイプにより遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所定の条件下で所定の時間に発現した遺伝子の数及び相対存在量を定量することにより分析される（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タイプの転写または逆転写全体に本発明のポリヌクレオチドまたはその補体をハイブリダイズすることにより、転写イメージを生成し得る。或る実施例では、本発明のポリヌクレオチドまたはその補体がマイクロアレイ上のエレメントのサブセットを複数含むような高処理フォーマットでハイブリダイゼーションを発生させる。結果として得られる転写イメージは、遺伝子活性のプロフィールを提供し得る。

【０２４３】

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

【０２４４】

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

【０２４５】

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

【０２４６】

別の実施例は、本発明のポリペプチド配列を用いて組織または細胞タイプのプロテオームを分析することに関連する。プロテオームの語は、特定の組織または細胞タイプでのタンパク質発現の包括的パターンを指す。プロテオームの各タンパク質成分は、個々に更に分析の対象とすることができる。プロテオーム発現パターン即ちプロフィールは、所与の条件下で所与の時間に発現したタンパク質の数及び相対存在量を定量することにより分析し得る。従って細胞のプロテオームのプロフィールは、特定の組織または細胞タイプのポリペプチドを分離及び分析することにより作成し得る。或る実施例では、１次元等電点電気泳動によりサンプルからタンパク質を分離し、２次元ドデシル硫酸ナトリウムスラブゲル

電気泳動により分子量に応じて分離するような２次元ゲル電気泳動により分離が達成される（前出のＳｔｅｉｎｅｒ および Ａｎｄｅｒｓｏｎ）。タンパク質は、通常クーマシーブルーまたはシルバーまたは蛍光染色などの物質を用いてゲルで染色することにより、分散した、独特な位置にある点としてゲル中で可視化される。各タンパク質スポットの光学密度は、通常サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処理または非処理のいずれかの生物学的サンプルから得られるタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を同定する。スポット内のタンパク質は、例えば化学的または酵素的切断とそれに続く質量分析を用いる標準的な方法を用いて部分的にシークエンシングする。スポット内のタンパク質の素性は、その部分配列を、好適には少なくとも５個の連続するアミノ酸残基を、本発明のポリペプチド配列と比較することにより決定し得る。場合によっては、決定的なタンパク質同定のための更なる配列が得られる。

【０２４７】

プロテオームのプロファイルは、ＴＲＩＣＨに特異的な抗体を用いてＴＲＩＣＨ発現レベルを定量することによっても作成可能である。或る実施例では、マイクロアレイ上でエレメントとして抗体を用い、マイクロアレイをサンプルに曝して各アレイ要素へのタンパク質結合レベルを検出することによりタンパク質発現レベルを定量する（Ｌｕｅｋｉｎｇ，Ａ．ら．（１９９９）Ａｎａｌ．Ｂｉｏｃｈｅｒｎ．２７０：１０３－１１１、Ｍｅｎｄｏｚｅ，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、チオールまたはアミノ反応性蛍光化合物を用いてサンプル中のタンパク質を反応させ、各アレイのエレメントにおける蛍光結合の量を検出し得る。

【０２４８】

プロテオームレベルでの毒性シグネチャも中毒学的スクリーニングに有用であり、転写レベルでの毒性シグネチャと並行に分析するべきである。或る組織の或るタンパク質に対しては、転写とタンパク質の存在量の相関が乏しいこともあるので（Ａｎｄｅｒｓｏｎ，Ｎ．Ｌ．および Ｊ．Ｓｅｉｌｈａｍｅｒ（１９９７）Ｅｌｅｃｔｒｏｐｈｏｒｅｓｉｓ １８：５３３－５３７）、転写イメージにはそれ程影響しないがタンパク質のプロフィールを変化させるような化合物の分析においてプロテオーム毒性シグネチャは有用たり得る。更に、体液中での転写の分析はｍＲＮＡ急速な分解により困難であるので、タンパク質のプロフィール作成はこのような場合により信頼でき、情報価値がある。

【０２４９】

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

【０２５０】

別の実施例では、タンパク質を含有する生物学的サンプルを試験化合物で処理することにより試験化合物の毒性を算定する。生物学的サンプルから得たタンパク質は、本発明のポリペプチドに特異的な抗体を用いてインキュベートする。抗体により認識されたタンパク質の量を定量する。処理された生物学的サンプル中のタンパク質の量を、非処理生物学的サンプル中のタンパク質の量と比較する。両サンプルのタンパク質量の差は、処理サンプル中の試験化合物に対する毒性反応を示す。

【０２５１】

マイクロアレイは、本技術分野でよく知られている方法を用いて調製し、使用し、そして分析する（Ｂｒｅｎｎａｎ，Ｔ．Ｍ．ら（１９９５）の米国特許第５，４７４，７９６号、Ｓｃｈｅｎａ，Ｍ．ら（１９９６）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃ

ｉ．ＵＳＡ９３：１０６１４－１０６１９、Ｂａｌｄｅｓｃｈｗｅｉｌｅｒらの（１９９５）ＰＣＴ出願第ＷＯ９５／２５１１１６号、Ｓｈａｌｏｎ，Ｄ．らの（１９９５）ＰＣＴ出願第ＷＯ９５／３５５０５号、Ｈｅｌｌｅｒ，Ｒ．Ａ．ら（１９９７）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ９４：２１５０－２１５５、Ｈｅｌｌｅｒ，Ｍ．Ｊ．らの（１９９７）米国特許第５，６０５，６６２号等を参照）。様々なタイプのマイクロアレイが公知であり、詳細については、ＤＮＡ Ｍｉｃｒｏａｒｒａｙｓ：Ａ Ｐｒａｃｔｉｃａｌ Ａｐｐｒｏａｃｈ，Ｍ．Ｓｃｈｅｎａ，編集．（１９９９）Ｏｘｆｏｒｄ Ｕｎｉｖｅｒｓｉｔｙ Ｐｒｅｓｓ，Ｌｏｎｄｏｎに記載されている。該文献は、特に引用することを以って本明細書の一部となす。

【０２５２】

本発明の別の実施例ではまた、ＴＲＩＣＨをコードする核酸配列を用いて、天然のゲノム配列をマッピングするのに有用なハイブリダイゼーションプローブを作製することが可能である。コード配列または非コード配列のいずれかを用いることができ、或る例では、コード配列全体よりも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内でのコード配列の保存により、染色体マッピング中に望ましくないクロスハイブリダイゼーションが生じる可能性がある。核酸配列は、特定の染色体、染色体の特定領域または人工形成の染色体、例えば、ヒト人工染色体（ＨＡＣ）、酵母人工染色体（ＹＡＣ）、細菌人工染色体（ＢＡＣ）、細菌Ｐ１産物、或いは単一染色体ｃＤＮＡライブラリに対してマッピングされる（例えば、Ｈａｒｒｉｎｇｔｏｎ，Ｊ．Ｊ．他（１９９７）Ｎａｔ．Ｇｅｎｅｔ．１５：３４５－３５５；Ｐｒｉｃｅ，Ｃ．Ｍ．（１９９３）Ｂｌｏｏｄ Ｒｅｖ．７：１２７１３４；ａｎｄ Ｔｒａｓｋ，Ｂ．Ｊ．（１９９１）Ｔｒｅｎｄｓ Ｇｅｎｅｔ．７：１４９１５４を参照。）一度マッピングすると、本発明の核酸配列を用いて例えば病状の遺伝を特定の染色体領域の遺伝または制限酵素断片長多型（ＲＦＬＰ）と相関させるような遺伝子連鎖地図を発生させ得る。（例えば、Ｌａｎｄｅｒ，Ｅ．Ｓ．ａｎｄ Ｄ．Ｂｏｔｓｔｅｉｎ（１９８６）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ８３：７３５３－７３５７を参照。）

蛍光ｉｎ ｓｉｔｕハイブリッド形成法（ＦＩＳＨ）は、他の物理的及び遺伝地図データと相関し得る（前出のＨｅｉｎｚ－Ｕｌｒｉｃｈ，ら（１９９５）ｉｎ Ｍｅｙｅｒｓ，９６５－９６８ページ．等を参照）遺伝地図データの例は、種々の科学雑誌あるいはＯｎｌｉｎｅ Ｍｅｎｄｅｌｉａｎ Ｉｎｈｅｒｉｔａｎｃｅ ｉｎ Ｍａｎ（ＯＭＩＭ）のウェブサイトに見ることができる。物理的な染色体地図上のＴＲＩＣＨをコードする遺伝子の位置と特定の疾患との相関性、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。

【０２５３】

確定した染色体マーカーを用いた結合分析等の物理的マッピング技術及び染色体標本原位置ハイブリッド形成法を用いて、遺伝地図を拡張することができる。例えばマウスなど別の哺乳動物の染色体上に遺伝子を配置することにより、正確な染色体の遺伝子座がわかっていない場合でも関連するマーカーを明らかにし得る。この情報は、位置クローニングその他の遺伝子発見技術を用いて疾患遺伝子を探す研究者にとって価値がある。疾患または症候群に関与する遺伝子が、血管拡張性失調症の１１ｑ２２－２３領域等、特定の遺伝子領域への遺伝的結合によって大まかに位置決めがなされると、該領域に対するいかなる配列マッピングも、更なる調査のための関連遺伝子或いは調節遺伝子を表すことができる（Ｇａｔｔｉ，Ｒ．Ａ．ら（１９８８）Ｎａｔｕｒｅ ３３６：５７７－５８０等を参照）転座、反転等に起因する、健常者、保有者、感染者の三者間における染色体位置の相違を発見するために本発明のヌクレオチド配列を用いてもよい。

【０２５４】

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リゴペプチドを、種々の任意の薬剤スクリーニング技術における化合物のライブラリのスクリーニングに用いることができる。薬剤スクリーニングに用いる断片は、溶液中に遊離

しているか、固体支持物に固定されるか、細胞表面上に保持されるか、細胞内に位置することになろう。ＴＲＩＣＨ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。

【０２５５】

別の薬剤スクリーニング方法は、目的のタンパク質に対して好適な結合親和性を有する化合物を高い処理能力でスクリーニングするために用いられる（Ｇｅｙｓｅｎ，他（１９８４）ＰＣＴ ａｐｐｌｉｃａｔｉｏｎ ＷＯ８４／０３５６４等を参照。）この方法においては、多数の異なる小さな試験用化合物を固体基板上で合成する。試験用化合物は、ＴＲＩＣＨ、或いはその断片と反応してから洗浄される。次ぎに、結合されたＴＲＩＣＨが、当分野で周知の方法で検出される。精製されたＴＲＩＣＨはまた、前記した薬剤をスクリーニングする技術に用いられるプレート上で直接被覆することもできる。別法では、非中和抗体を用いてペプチドを捕捉し、ペプチドを固体支持物に固定することもできる。

【０２５６】

別の実施例では、ＴＲＩＣＨと結合可能な中和抗体がＴＲＩＣＨ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、ＴＲＩＣＨ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。

【０２５７】

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

【０２５８】

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

【０２５９】

本明細書において開示した全ての特許、特許出願及び刊行物、特に米国特許第６０／２３２，６８５号、第６０／２３４，８４２号、第６０／２３６，８８２号、第６０／２３９，０５７号、第６０／２４０，５４０号および第６０／２４１，７００号は、言及することをもって本明細書の一部となす。

【０２６０】

（実施例）

１　ｃＤＮＡライブラリの作製

Ｉｎｃｙｔｅ ｃＤＮＡは、ＬＩＦＥＳＥＱ ＧＯＬＤデータベース（Ｉｎｃｙｔｅ Ｇｅｎｏｍｉｃｓ，Ｐａｌｏ Ａｌｔｏ ＣＡ）に記載されたｃＤＮＡライブラリに由来するものであり、表４の列５に列記した。ホモジナイズしてグアニジニウムイソチオシアネート溶液に溶解した組織もあり、また、ホモジナイズしてフェノールまたは好適な変性剤の混合液に溶解した組織もある。変性剤の混合液は、例えばフェノールとグアニジニウムイソチオシアネートの単相溶液であるＴＲＩＺＯＬ（Ｌｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ）等である。結果として得られた溶解物は、塩化セシウムクッション上で遠心分離するかクロロホルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方、或いは別の方法を用いて、溶解物からＲＮＡを沈殿させた。

【０２６１】

ＲＮＡの純度を高めるため、ＲＮＡのフェノールによる抽出及び沈殿を必要な回数繰り返した。場合によっては、ＤＮアーゼでＲＮＡを処理した。殆どのライブラリでは、オリゴｄ（Ｔ）連結常磁性粒子（Ｐｒｏｍｅｇａ）、ＯＬＩＧＯＴＥＸラテックス粒子（ＱＩＡＧＥＮ，Ｖａｌｅｎｃｉａ ＣＡ）またはＯＬＩＧＯＴＥＸ ｍＲＮＡ精製キット（ＱＩＡＧＥＮ）を用いて、ポリ（Ａ＋）ＲＮＡを単離した。別法では、別のＲＮＡ単離キッ

ト、例えばＰＯＬＹ（Ａ）ＰＵＲＥ ｍＲＮＡ精製キット（Ａｍｂｉｏｎ，Ａｕｓｔｉｎ ＴＸ）を用いて組織溶解物からＲＮＡを直接単離した。

【０２６２】

場合によってはＳｔｒａｔａｇｅｎｅ社へのＲＮＡ提供を行い、対応するｃＤＮＡライブラリをＳｔｒａｔａｇｅｎｅ社が作製することもあった。そうでない場合は、ＵＮＩＺＡＰベクターシステム（Ｓｔｒａｔａｇｅｎｅ）またはＳＵＰＥＲＳＣＲＩＰＴプラスミドシステム（Ｌｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ）を用いて本技術分野で公知の推奨方法または類似の方法でｃＤＮＡを合成し、ｃＤＮＡライブラリを作製した（前出のＡｕｓｕｂｅｌ，１９９７，ｕｎｉｔ ５．１－６．６等を参照）。逆転写は、オリゴｄ（Ｔ）またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖ｃＤＮＡに連結反応させ、好適な制限酵素または酵素でｃＤＮＡを消化した。殆どのライブラリに対して、ｃＤＮＡのサイズ（３００～１０００ｂｐ）選択は、ＳＥＰＨＡＣＲＹＬ Ｓ１０００、ＳＥＰＨＡＲＯＳＥ ＣＬ２ＢまたはＳＥＰＨＡＲＯＳＥ ＣＬ４Ｂカラムクロマトグラフィー（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）、或いは調製用アガロースゲル電気泳動法を用いて行った。 合成オリゴヌクレオチドアダプターを二本鎖ｃＤＮＡに連結反応させ、好適な制限酵素または酵素でｃＤＮＡを消化した。 好適なプラスミドは、例えばＰＢＬＵＥＳＣＲＩＰＴプラスミド（Ｓｔｒａｔａｇｅｎｅ）、ＰＳＰＯＲＴ１ プラスミド（Ｌｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ）、ＰＣＤＮＡ２．１ プラスミド（Ｉｎｖｉｔｒｏｇｅｎ，Ｃａｒｌｓｂａｄ ＣＡ）、ＰＢＫ－ＣＭＶ プラスミド（Ｓｔｒａｔａｇｅｎｅ）、ＰＣＲ２－ＴＯＰＯＴＡ（Ｉｎｖｉｔｒｏｇｅｎ）、ＰＣＭＶ－ＩＣＩＳ（Ｓｔｒａｔａｇｅｎｅ）、ｐＩＮＣＹ（Ｉｎｃｙｔｅ Ｇｅｎｏｍｉｃｓ，Ｐａｌｏ Ａｌｔｏ ＣＡ）等およびその誘導体である。組換えプラスミドは、Ｓｔｒａｔａｇｅｎｅ社のＸＬ１－Ｂｌｕｅ、ＸＬ１－ＢＩｕｅＭＲＦまたはＳＯＬＲ、或いはＬｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ社のＤＨ５α、ＤＨ１０ＢまたはＥＬＥＣＴＲＯＭＡＸ ＤＨ１０Ｂを含む大腸菌細胞に形質転換した。

【０２６３】

２ ｃＤＮＡクローンの単離

ＵＮＩＺＡＰベクターシステム（Ｓｔｒａｔａｇｅｎｅ）を用いたｉｎ ｖｉｖｏ切除によって、或いは細胞溶解によって、実施例１のようにして得たプラスミドを宿主細胞から回収した。ＭａｇｉｃまたはＷＩＺＡＲＤ Ｍｉｎｉｐｒｅｐｓ ＤＮＡ精製システム（Ｐｒｏｍｅｇａ）、ＡＧＴＣ Ｍｉｎｉｐｒｅｐ精製キット（Ｅｄｇｅ Ｂｉｏｓｙｓｔｅｍｓ，Ｇａｉｔｈｅｒｓｂｕｒｇ ＭＤ）、ＱＩＡＧＥＮ社のＱＩＡＷＥＬＬ ８ Ｐｌａｓｍｉｄ、ＱＩＡＷＥＬＬ ８ Ｐｌｕｓ Ｐｌａｓｍｉｄ及びＱＩＡＷＥＬＬ ８ Ｕｌｔｒａ Ｐｌａｓｍｉｄ 精製システム、Ｒ．Ｅ．Ａ．Ｌ． Ｐｒｅｐ ９６プラスミドキットの中から少なくとも１つを用いて、プラスミドを精製した。プラスミドは、沈殿させた後、０．１ｍｌの蒸留水に再懸濁して、凍結乾燥して或いは凍結乾燥しないで４℃で保管した。

【０２６４】

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

【０２６５】

３ シークエンシング及び分析

実施例２に記載したようにプラスミドから回収したＩｎｃｙｔｅ ｃＤＮＡを、以下に示すようにシークエンシングした。ｃＤＮＡのシークエンス反応は、標準的方法或いは高処理装置、例えばＡＢＩ ＣＡＴＡＬＹＳＴ ８００ サーマルサイクラー（Ａｐｐｌｉｅｄ

Ｂｉｏｓｙｓｔｅｍｓ）またはＰＴＣ－２００ サーマルサイクラー（ＭＪ Ｒｅｓｅａｒｃｈ）をＨＹＤＲＡマイクロディスペンサー（Ｒｏｂｂｉｎｓ Ｓｃｉｅｎｔｉｆｉｃ）またはＭＩＣＲＯＬＡＢ ２２００（Ｈａｍｉｌｔｏｎ）液体転移システムと併用して処理した。ｃＤＮＡのシークエンス反応は、Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ社が提供する試薬またはＡＢＩシークエンシングキット、例えばＡＢＩ ＰＲＩＳＭ ＢＩＧＤＹＥ Ｔｅｒｍｉｎａｔｏｒ ｃｙｃｌｅ ｓｅｑｕｅｎｃｉｎｇ ｒｅａｄｙ ｒｅａｃｔｉｏｎ ｋｉｔ（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）に与えられた試薬を用いて準備した。ｃＤＮＡのシークエンス反応の電気泳動的分離及び標識したポリヌクレオチドの検出には、ＭＥＧＡＢＡＣＥ １０００ ＤＮＡシークエンシングシステム（Ｍｏｌｅｃｕｌａｒ Ｄｙｎａｍｉｃｓ）か、標準ＡＢＩプロトコル及び塩基対呼び出しソフトウェアを用いるＡＢＩ ＰＲＩＳＭ ３７３または３７７シークエンシングシステム（Ａｐｐｌｉｅｄ Ｂｉｏｓｙｓｔｅｍｓ）か、或いはその他の本技術分野でよく知られている配列解析システムを用いた。ｃＤＮＡ配列内のリーディングフレームは、標準的方法（前出のＡｕｓｕｂｅｌ，１９９７，ｕｎｉｔ ７．７に概説）を用いて決定した。ｃＤＮＡ配列の幾つかを選択して、実施例８に記載した方法で配列を伸長させた。

【０２６６】

Ｉｎｃｙｔｅ ｃＤＮＡ配列に由来するポリヌクレオチド配列は、ベクター、リンカー及びポリ（Ａ）配列を除去し、あいまいな塩基対をマスクすることによって有効性を確認した。その際、ＢＬＡＳＴ、動的プログラミング及び隣接ジヌクレオチド頻度分析に基づくアルゴリズム及びプログラムを用いた。公共のデータベース、例えばＧｅｎＢａｎｋの霊長類及びげっ歯類、哺乳動物、脊椎動物、真核生物のデータベースと、ＢＬＯＣＫＳ、ＰＲＩＮＴＳ、ＤＯＭＯ、ＰＲＯＤＯＭ、及びＰＦＡＭ等のような隠れマルコフモデル（ＨＭＭ）ベースのタンパク質ファミリーデータベースの選択に対して、Ｉｎｃｙｔｅ ｃＤＮＡ配列またはその翻訳を問い合わせた（ＨＭＭは、遺伝子ファミリーのコンセンサス１次構造を分析する確率的アプローチである。Ｅｄｄｙ，Ｓ．Ｒ．（１９９６）Ｃｕｒｒ．Ｏｐｉｎ．Ｓｔｒｕｃｔ．Ｂｉｏｌ．６：３６１－３６５等を参照。）問合せは、ＢＬＡＳＴ、ＦＡＳＴＡ、ＢＬＩＭＰＳ及びＨＭＭＥＲに基づくプログラムを用いて行った。Ｉｎｃｙｔｅ ｃＤＮＡ配列は、完全長のポリヌクレオチド配列を産出するように構築した。或いは、ＧｅｎＢａｎｋ ｃＤＮＡ、ＧｅｎＢａｎｋ ＥＳＴ、ステッチされた配列、ストレッチされた配列またはＧｅｎｓｃａｎ予測コード配列（実施例４及び５を参照）を用いてＩｎｃｙｔｅ ｃＤＮＡの集団を完全長まで伸長させた。Ｐｈｒｅｄ、Ｐｈｒａｐ及びＣｏｎｓｅｄに基づくプログラムを用いて構築し、ＧｅｎＭａｒｋ、ＢＬＡＳＴ及びＦＡＳＴＡに基づくプログラムを用いてｃＤＮＡの集団をオープンリーディングフレームに対してスクリーニングした。対応する完全長ポリペプチド配列を誘導するべく完全長ポリヌクレオチド配列を翻訳した。或いは、本発明のポリペプチドは完全長翻訳ポリペプチドの任意のメチオニン残基で開始し得る。引き続いて、ＧｅｎＢａｎｋタンパク質データベース（ｇｅｎｐｅｐｔ）、ＳｗｉｓｓＰｒｏｔ、ＢＬＯＣＫＳ、ＰＲＩＮＴＳ、ＤＯＭＯ、ＰＲＯＤＯＭ及びＰｒｏｓｉｔｅ等のデータベース、ＰＦＡＭ等の隠れマルコフモデル（ＨＭＭ）ベースのタンパク質ファミリーデータベースに対する問合せによって完全長ポリペプチド配列を分析した。完全長ポリヌクレオチド配列はまた、ＭＡＣＤＮＡＳＩＳ ＰＲＯソフトウェア（日立ソフトウェアエンジニアリング，Ｓｏｕｔｈ Ｓａｎ Ｆｒａｎｃｉｓｃｏ ＣＡ）及びＬＡＳＥＲＧＥＮＥソフトウェア（ＤＮＡＳＴＡＲ）を用いて分析した。ポリヌクレオチド及びポリペプチド配列アラインメントは、アラインメントした配列と配列の一致率も計算するＭＥＧＡＬＩＧＮマルチシークエンスアラインメントプログラム（ＤＮＡＳＴＡＲ）に組み込まれているようなＣＬＵＳＴＡＬアルゴリズムによって特定されるデフォルトパラメータを用いて生成する。

【０２６７】

Ｉｎｃｙｔｅ ｃＤＮＡ及び完全長配列の分析及びアセンブリに利用したツール、プログラム及びアルゴリズムの概略と、適用可能な説明、引用文献、閾値パラメータを表７に示す。用いたツール、プログラム及びアルゴリズムを表７の列１に、それらの簡単な説明を

列２に示す。 列３は好適な引用文献であり、全ての文献はそっくりそのまま引用を以って本明細書の一部となす。 適用可能な場合には、列４は２つの配列が一致する強さを評価するために用いたスコア、確率値その他のパラメータを示す（スコアが高ければ高いほどまたは確率値が低ければ低いほど、２配列間の相同性が高くなる）。

【０２６８】

完全長のポリヌクレオチド配列およびポリペプチド配列の組み立て及び分析に用いる上記のプログラムは、ＳＥＱ ＩＤ ＮＯ：２７－５２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。

【０２６９】

４ ゲノムＤＮＡからのコード配列の同定及び編集

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

【０２７０】

５ ｃＤＮＡ配列データを使ったゲノム配列データの構築

ステッチ配列（Ｓｔｉｃｈｅｄ Ｓｅｑｕｅｎｃｅ）

部分ｃＤＮＡ配列は、実施例４に記載のＧｅｎｓｃａｎ遺伝子同定プログラムにより予測されたエキソンを用いて伸長させた。実施例３に記載されたように構築された部分ｃＤＮＡは、ゲノムＤＮＡにマッピングし、関連するｃＤＮＡ及び１つ若しくは複数のゲノム配列から予測されたＧｅｎｓｃａｎエキソンを含むクラスタに分解した。ｃＤＮＡ及びゲノム情報を統合するべくグラフ理論及び動的プログラミングに基づくアルゴリズムを用いて各クラスタを分析し、引き続いて確認、編集または伸長して完全長配列を産出するような潜在的スプライス変異体を生成した。間隔全体の長さがクラスタ中の２以上の配列に存在するような配列を同定し、そのように同定された間隔は推移により等しいと考えられた。例えば、１つのｃＤＮＡ及び２つのゲノム配列に間隔が存在する場合、３つの間隔は全て等しいと考えられる。このプロセスは、無関係であるが連続したゲノム配列をｃＤＮＡ配列により結び合わせて架橋し得る。このようにして同定された区間を、親配列（ｐａｒｅ

ｎｔ ｓｅｑｕｅｎｃｅ）に沿って現われるようにステッチアルゴリズムで縫い合わせ、可能な最も長い配列および変異配列を作製する。１種類の親配列に沿って発生した間隔と間隔との連鎖（ｃＤＮＡ－ｃＤＮＡまたはゲノム配列－ゲノム配列）は、親の種類を変える連鎖（ｃＤＮＡ－ゲノム配列）に優先した。結果として得られるステッチ配列は、ＢＬＡＳＴ分析により公共データベースｇｅｎｐｅｐｔ及びｇｂｐｒｉに翻訳されて比較された。Ｇｅｎｓｃａｎにより予測された不正確なエキソンは、ｇｅｎｐｅｐｔからヒットしたトップのＢＬＡＳＴと比較することにより修正した。必要な場合には、追加ｃＤＮＡ配列を用いるかゲノムＤＮＡの検査により配列を更に伸長させた。

【０２７１】

ストレッチ配列（Ｓｔｒｅｔｃｈｅｄ Ｓｅｑｕｅｎｃｅ）

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

【０２７２】

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：２７－５２を構築するために用いた配列を、ＢＬＡＳＴ及びＳｍｉｔｈ－Ｗａｔｅｒｍａｎアルゴリズムを用いて、Ｉｎｃｙｔｅ ＬＩＦＥＳＥＱデータベース及び公共のドメインデータベースの配列と比較した。ＳＥＱ ＩＤ ＮＯ：２７－５２と一致するこれらのデータベースの配列を、Ｐｈｒａｐ（表７）などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。スタンフォード・ヒトゲノムセンター（ＳＨＧＣ）、ホワイトヘッド・ゲノム研究所（ＷＩＧＲ）、Ｇｅｎｅｔｈｏｎ等の公的な情報源から入手可能な放射線ハイブリッド及び遺伝地図データを用いて、クラスタ化された配列が前もってマッピングされたかを決定した。マッピングされた配列がクラスタに含まれている場合は、個々の配列番号を含めてそのクラスタの全配列が地図上の位置に割り当てられた。

【０２７３】

地図上の位置は、ヒト染色体の範囲または間隔として表される。センチモルガン間隔の地図上の位置は、染色体のｐアームの末端に関連して測定する。（センチモルガン（ｃＭ）は、染色体マーカー間の組換え頻度に基づく計測単位である。平均して、１ｃＭは、ヒト中のＤＮＡの１メガベース（Ｍｂ）にほぼ等しい。 尤も、この値は、組換えのホットスポット及びコールドスポットに起因して広範囲に変化する。）ｃＭ距離は、配列が各クラスタ内に含まれるような放射線ハイブリッドマーカーに対して境界を提供するようなＧｅｎｅｔｈｏｎによってマッピングされた遺伝マーカーに基づく。ＮＣＢＩ「ＧｅｎｅＭａｐ９９」（ｈｔｔｐ：／／ｗｗｗ．ｎｃｂｉ．ｎｌｍ．ｎｉｈ．ｇｐｖ／ｇｅｎｅｍａｐ）などの一般個人が入手可能なヒト遺伝子マップおよびその他の情報源を用いて、上記した区間が既に同定されている疾患遺伝子マップ内若しくは近傍に位置するかを決定できる。

【０２７４】

このようにして、ＳＥＱ ＩＤ ＮＯ：３１は１３３．００から１３７．３０センチモルガン内で染色体１にマップされた。ＳＥＱ ＩＤ ＮＯ：３３ は、染色体１２でｑ末端から

１２０．５０以内、より具体的には１２６．１０～１４５．７０センチモルガン以内の区間にマッピングされる。

7　ポリヌクレオチド発現の分析

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

【０２７５】

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

【数１】

$$\frac{(\text{BLAST スコア} \times \text{配列一致率})}{5 \times (\text{長さ(配列1)}, \text{長さ(配列2)}) \text{の最小値}}$$

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

【０２７６】

或いは、ＴＲＩＣＨをコードするポリヌクレオチド配列は、由来する組織に対して分析する。例えば或る完全長配列は、Ｉｎｃｙｔｅ ｃＤＮＡ配列（実施例３を参照）と少なくとも一部は重畳するように構築される。各ｃＤＮＡ配列は、ヒト組織から作製されたｃＤＮＡライブラリに由来する。各ヒト組織は、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、または尿管などの生物／組織のカテゴリーの一つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。同様に、各ヒト組織は、以下の疾患／病状カテゴリー即ち癌、細胞系、発達、炎症、神経性、外傷、心血管、鬱血、その他の１つに分類される。 各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。得られるパーセンテージは、ＴＲＩＣＨをコードするｃＤＮＡの疾患特異的な発現を反映する。 ｃＤＮＡ配列およびｃＤＮＡライブラリ／組織の情報は、ＬＩＦＥＳＥＱ ＧＯＬＤ データベース（Ｉｎｃｙｔｅ Ｇｅｎｏｍｉｃｓ，Ｐａｌｏ Ａｌｔｏ ＣＡ）から得ることができる。

【０２７７】

8　ＴＲＩＣＨをコードするポリヌクレオチドの伸長

完全長のポリヌクレオチド配列もまた、完全長分子の適切な断片から設計したオリゴヌク

レオチドプライマーを用いて該断片を伸長させて生成した。一方のプライマーは既知の断片の５’伸長を開始するべく合成し、他方のプライマーは既知の断片の３’伸長を開始するべく合成した。開始プライマーは、長さが約２２～３０ヌクレオチド、ＧＣ含有率が約５０％以上となり、約６８～７２℃の温度で標的配列にアニーリングするように、ＯＬＩＧＯ ４．０６ソフトウェア（Ｎａｔｉｏｎａｌ Ｂｉｏｓｃｉｅｎｃｅｓ）或いは別の適切なプログラムを用いて、ｃＤＮＡから設計した。 ヘアピン構造及びプライマー－プライマー二量体を生ずるようなヌクレオチドの伸長は全て回避した。

【０２７８】

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

【０２７９】

高忠実度の増幅が、当業者によく知られている方法を利用したＰＣＲ法によって得られた。 ＰＣＲは、ＰＴＣ－２００ サーマルサイクラー（ＭＪ Ｒｅｓｅａｒｃｈ, Ｉｎｃ.）を用いて９６穴プレート内で行った。反応混合液は、鋳型ＤＮＡ及び２００ ｎｍｏｌの各プライマー、Mg^{2+}と$(NH_4)_2SO_4$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秒間
ステップ３　　６０℃で１分間
ステップ４　　６８℃で２分間
ステップ５　　ステップ２、３、及び４を２０回繰り返す
ステップ６　　６８℃で５分間
ステップ７　　４℃で保存

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

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

各ウェルのＤＮＡ濃度は、１Ｘ ＴＥ及び０．５μｌの希釈していないＰＣＲ産物に溶解した１００μｌのＰＩＣＯＧＲＥＥＮ定量試薬（０．２５（ｖ／ｖ） ＰＩＣＯＧＲＥＥＮ；Ｍｏｌｅｃｕｌａｒ Ｐｒｏｂｅｓ, Ｅｕｇｅｎｅ ＯＲ）を不透明な蛍光光度計プレート（Ｃｏｍｉｎｇ Ｃｏｓｔａｒ, Ａｃｔｏｎ ＭＡ）の各ウェルに分配してＤＮＡが試薬と結合できるようにして測定する。サンプルの蛍光を計測してＤＮＡの濃度を定量するべくプレートをＦｌｕｏｒｏｓｋａｎ ＩＩ（Ｌａｂｓｙｓｔｅｍｓ Ｏｙ, Ｈｅｌｓｉｎｋｉ, Ｆｉｎｌａｎｄ）でスキャンした。反応混合物のアリコート５～１０μｌを１％アガロースミニゲル上で電気泳動法によって解析し、どの反応が配列の伸長に成功したかを決定した。

【０２８０】

伸長させたヌクレオチドは、脱塩及び濃縮して３８４穴プレートに移し、ＣｖｉＪＩコレラウイルスエンドヌクレアーゼ（Ｍｏｌｅｃｕｌａｒ Ｂｉｏｌｏｇｙ Ｒｅｓｅａｒｃｈ, Ｍａｄｉｓｏｎ ＷＩ）を用いて消化し、ｐＵＣ １８ベクター（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）への再連結反応前に音波処理またはせん断した。ショットガン・シークエンシングのために、消化したヌクレオチドを低濃度（０．６～０．

８％）のアガロースゲル上で分離し、断片を切除し、寒天をＡｇａｒ ＡＣＥ（Ｐｒｏｍｅｇａ）で消化した。伸長させたクローンをＴ４リガーゼ（Ｎｅｗ Ｅｎｇｌａｎｄ Ｂｉｏｌａｂｓ, Ｂｅｖｅｒｌｙ ＭＡ）を用いてｐＵＣ １８ベクター（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）に再連結し、Ｐｆｕ ＤＮＡポリメラーゼ（Ｓｔｒａｔａｇｅｎｅ）で処理して制限部位のオーバーハングを満たし、大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りとってＬＢ／２Ｘカルベニシリン培養液の３８４ウェルプレートに３７℃で一晩培養した。

【０２８１】

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

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

【０２８２】

同様に、上記手順を用いて完全長ヌクレオチド配列を検証し、或いはそのような伸長のために設計されたオリゴヌクレオチド及び適切なゲノムライブラリを用いて５’調節配列を得る。

【０２８３】

９　個々のハイブリダイゼーションプローブの標識及び使用

ＳＥＱ ＩＤ ＮＯ：２７－５２から導き出されたハイブリダイゼーションプローブを用いて、ｃＤＮＡ、ｍＲＮＡ、またはゲノムＤＮＡをスクリーニングする。約２０塩基対からなるオリゴヌクレオチドの標識について特に記載するが、より大きなヌクレオチド断片に対しても事実上同一の手順が用いられる。オリゴヌクレオチドは、ＯＬＩＧＯ ４．０６ソフトウェア（Ｎａｔｉｏｎａｌ Ｂｉｏｓｃｉｅｎｃｅｓ）等の最新ソフトウェアを用いて設計し、各オリゴマー５０ｐｍｏｌと、［γ－^{32}Ｐ］アデノシン３リン酸（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）２５０μＣｉと、Ｔ４ポリヌクレオチドキナーゼ（ＤｕＰｏｎｔ ＮＥＮ, Ｂｏｓｔｏｎ ＭＡ）を混合することにより標識する。標識したオリゴヌクレオチドは、ＳＥＰＨＡＤＥＸ Ｇ－２５超細繊分子サイズ排除デキストラン ビードカラム（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）を用いて十分に精製する。Ａｓｅ Ｉ、Ｂｇｌ ＩＩ、Ｅｃｏ ＲＩ、Ｐｓｔ Ｉ、Ｘｂａ１またはＰｖｕ ＩＩ（ＤｕＰｏｎｔ ＮＥＮ）のいずれか１つのエンドヌクレアーゼで消化されたヒトゲノムＤＮＡの典型的な膜ベースのハイブリダイゼーション解析において、毎分10^{7}カウントの標識されたプローブを含むアリコットを用いる。

【０２８４】

各消化物から得たＤＮＡは、０．７％アガロースゲル上で分画してナイロン膜（Ｎｙｔｒａｎ Ｐｌｕｓ, Ｓｃｈｌｅｉｃｈｅｒ ＆ Ｓｃｈｕｅｌｌ, Ｄｕｒｈａｍ ＮＨ）に移

す。ハイブリダイゼーションは、４０℃で１６時間行う。 非特異的シグナルを除去するため、例えば０．１×クエン酸ナトリウム食塩水及び０．５％ドデシル硫酸ナトリウムに一致する条件下で、ブロットを室温で順次洗浄する。オートラジオグラフィーまたはそれに代わるイメージング手段を用いてハイブリダイゼーションパターンを視覚化し、比較する。

【０２８５】

１０　マイクロアレイ

マイクロアレイの表面上でアレイエレメントの結合または合成は、フォトリソグラフィ、圧電印刷（インクジェット印刷、前出のＢａｌｄｅｓｃｈｗｅｉｌｅｒ等を参照）、機械的マイクロスポッティング技術及びこれらから派生したものを用いて達成することが可能である。上記各技術において基板は、均一且つ非多孔性の固体とするべきである（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。或いは、ドットブロット法またはスロットブロット法に類似のアレイを利用して、熱的、紫外線的、化学的または機械的結合手順を用いて基板の表面にエレメントを配置及び結合させてもよい。通常のアレイは、手作業で、または利用可能な方法や機械を用いて作製でき、任意の適正数のエレメントを有し得る（Ｓｃｈｅｎａ，Ｍ．ら（１９９５）Ｓｃｉｅｎｃｅ２７０：４６７－４７０、Ｓｈａｌｏｎ．Ｄ．ら（１９９６）Ｇｅｎｏｍｅ Ｒｅｓ．６：６３９－６４５、Ｍａｒｓｈａｌｌ，Ａ．およびＪ．Ｈｏｄｇｓｏｎ（１９９８）Ｎａｔ．Ｂｉｏｔｅｃｈｎｏｌ．１６：２７－３１．を参照）。

【０２８６】

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

【０２８７】

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サンプルを、ＭＭＬＶ逆転写酵素、０．０５ｐｇ／μｌのオリゴ（ｄＴ）プライマー（２１ｍｅｒ）、１×第一鎖合成バッファー、０．０３ｕｎｉｔ／μｌのＲＮアーゼ阻害因子、５００μＭのｄＡＴＰ、５００μＭのｄＧＴＰ、５００μＭのｄＴＴＰ、４０μＭのｄＣＴＰ、４０μＭのｄＣＴＰ－Ｃｙ３（ＢＤＳ）またはｄＣＴＰ－Ｃｙ５（Ａｍｅｒｓｈａｍ Ｐｈａｒｍａｃｉａ Ｂｉｏｔｅｃｈ）を用いて逆転写する。逆転写反応は、ＧＥＭＢＲＩＧＨＴキット（Ｉｎｃｙｔｅ）を用いてポリ（Ａ）$^{+}$ ＲＮＡ含有の２５体積ｍｌ内で行う。特異的制御ポリ（Ａ）$^{+}$ ＲＮＡは、非コード酵母ゲノムＤＮＡからｉｎ ｖｉｔｒｏ転写により合成する。各反応サンプル（１つはＣｙ３、もう１つはＣｙ５標識）は、２．５ｍｌの０．５Ｍ水酸化ナトリウムで処理し、８５℃で２０分間インキュベートし、反応を停止させてＲＮＡを分解させる。サンプルは、２つの連続するＣＨＲＯＭＡ ＳＰＩＮ３０ゲル濾過スピンカラム（ＣＬＯＮＴＥＣＨ Ｌａｂｏｒａｔｏｒｉｅｓ，Ｉｎｃ．（ＣＬＯＮＴＥＣＨ），Ｐａｌｏ Ａｌｔｏ ＣＡ）を用いて精製する。 混合後、２つの反応サンプルは、１ｍｌのグリコーゲン（１ｍｇ／ｍｌ）、６０ｍｌの酢酸ナトリウム及

び３００ｍｌの１００％エタノールを用いてエタノール析出させる。サンプルは次に、ＳｐｅｅｄＶＡＣ（Ｓａｖａｎｔ Ｉｎｓｔｒｕｍｅｎｔｓ Ｉｎｃ．，Ｈｏｌｂｒｏｏｋ ＮＹ）を用いて乾燥して仕上げ、１４μｌの５×ＳＳＣ／０．２％ＳＤＳ中で再懸濁する。

【０２８８】

マイクロアレイの調製

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

【０２８９】

精製したアレイエレメントは、ポリマーコートされたスライドグラス上に固定する。顕微鏡スライドグラス（Ｃｏｒｎｉｎｇ）は、処理中及び処理後に０．１％のＳＤＳ及びアセトン中で超音波をかけ、蒸留水で非常に良く洗って洗浄する。スライドグラスは、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％エタノール中で０．０５％アミノプロピルシラン（Ｓｉｇｍａ）を用いてコーティングする。コーティングしたスライドガラスは、１１０℃のオブンで硬化させる。

【０２９０】

米国特許第５，８０７，５２２号で説明されている方法を用いて、コーティングしたガラス基板にアレイエレメントを付加する。 該特許は、引用を以って本明細書の一部となす。平均濃度が１００ｎｇ／μｌのアレイエレメントＤＮＡ１μｌを高速機械装置により開放型キャピラリープリンティングエレメント（ｏｐｅｎ ｃａｐｉｌｌａｒｙ ｐｒｉｎｔｉｎｇ ｅｌｅｍｅｎｔ）に充填する。装置はここで、スライド毎に約５ｎｌのアレイエレメントサンプルを加える。

【０２９１】

マイクロアレイには、ＳＴＲＡＴＡＬＩＮＫＥＲ ＵＶ架橋剤（Ｓｔｒａｔａｇｅｎｅ）を用いてＵＶ架橋する。マイクロアレイは、室温において０．２％ＳＤＳで１度洗浄し、蒸留水で３度洗浄する。リン酸緩衝生理食塩水 （ＰＢＳ）（Ｔｒｏｐｉｘ，Ｉｎｃ．，Ｂｅｄｆｏｒｄ ＭＡ）中の０．２％カゼイン中において６０℃で３０分間マイクロアレイをインキュベートした後、前に行ったように０．２％ＳＤＳ及び蒸留水で洗浄することにより、非特異結合部位をブロックする。

【０２９２】

ハイブリダイゼーション

ハイブリダイゼーション反応は、５×ＳＳＣ，０．２％ＳＤＳハイブリダイゼーション緩衝液中のＣｙ３及びＣｙ５標識したｃＤＮＡ合成生成物を各０．２μｇ含む９μｌのサンプル混合体を有する。サンプル混合体は、６５℃まで５分間加熱し、マイクロアレイ表面上で等分して１．８cm^2 のカバーガラスで覆う。アレイは、顕微鏡スライドより僅かに大きい空洞を有する防水チェンバーに移す。チェンバーのコーナーに１４０μｌの５×ＳＳＣを加えることにより、チェンバー内部を湿度１００％に保持する。アレイを含むチェンバーは、６０℃で約６．５時間インキュベートする。 アレイは、第１洗浄緩衝液中（１×ＳＳＣ，０．１％ＳＤＳ）において４５℃で１０分間洗浄し、第２洗浄緩衝液中（０．１×ＳＳＣ）において４５℃で１０分間各々３度洗浄して乾燥させる。

【０２９３】

検出

レポーター標識ハイブリダイゼーション複合体は、Ｃｙ３の励起のためには４８８ｎｍ、Ｃｙ５の励起のためには６３２ｎｍでスペクトル線を発生し得るＩｎｎｏｖａ ７０混合

ガス１０Ｗレーザ（Ｃｏｈｅｒｅｎｔ，Ｉｎｃ．，Ｓａｎｔａ Ｃｌａｒａ ＣＡ）を備えた顕微鏡で検出する。２０×顕微鏡対物レンズ（Ｎｉｋｏｎ，Ｉｎｃ．，Ｍｅｌｖｉｌｌｅ ＮＹ）を用いて、アレイ上に励起レーザ光の焦点を当てる。アレイを含むスライドを顕微鏡のコンピュータ制御のＸ－Ｙステージに置き、対物レンズを通過してラスタースキャンする。本実施例で用いた１．８ｃｍ×１．８ｃｍのアレイは、２０μｍの解像度でスキャンした。

【０２９４】

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

【０２９５】

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対照種により生成されるシグナル強度を用いて較正する。アレイ上の特定の位置には相補的ＤＮＡ配列が含まれ、その位置におけるシグナルの強度をハイブリダイジング種の重量比１：１００，０００に相関させる。 異なる源泉（例えば試験される細胞及び対照細胞など）からの２つのサンプルを、各々異なる蛍光色素で標識し、他と異なって発現した遺伝子を同定するために単一のアレイにハイブリダイズする場合には、その較正を、２つの蛍光色素で較正するｃＤＮＡのサンプルを標識し、ハイブリダイゼーション混合体に各々等量を加えることによって行う。

【０２９６】

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－ディジタル（Ａ／Ｄ）変換ボード（Ａｎａｌｏｇ Ｄｅｖｉｃｅｓ，Ｉｎｃ．，Ｎｏｒｗｏｏｄ ＭＡ）を用いてディジタル化される。ディジタル化されたデータは、青色（低シグナル）から赤色（高シグナル）までの擬似カラー範囲へのリニア２０色変換を用いてシグナル強度がマッピングされたようなイメージとして表示される。データは、定量的にも分析される。２つの異なる蛍光色素を同時に励起及び測定する場合には、各蛍光色素の発光スペクトルを用いて、データは先ず蛍光色素間の光学的クロストーク（発光スペクトルの重なりに起因する）を補正する。

【０２９７】

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

【０２９８】

１１　相補的ポリヌクレオチド

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

【０２９９】

１２　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をサブクローニングする。このようなプロモーターには、ｌａｃオペレーター調節エレメントに関連するＴ５またはＴ７バクテリオファージプロモーター及びｔｒｐ－ｌａｃ（ｔａｃ）ハイブリッドプロモーターが含まれるが、これらに限定するものではない。組換えベクターを、ＢＬ２１（ＤＥ３）等の好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピルβ－Ｄチオガラクトピラノシド（ＩＰＴＧ）で誘発されるとＴＲＩＣＨを発現する。真核細胞でのＴＲＩＣＨの発現は、昆虫細胞株または哺乳動物細胞株に一般にバキュロウイルスとして知られているＡｕｔｏｇｒａｐｈｉｃａ ｃａｌｉｆｏｒｎｉｃａ核多面性ウイルス（ＡｃＭＮＰＶ）を感染させて行う。バキュロウイルスの非必須ポリヘドリン遺伝子を、相同組換え或いは転移プラスミドの媒介を伴う細菌の媒介による遺伝子転移のどちらかによって、ＴＲＩＣＨをコードするｃＤＮＡと置換する。ウイルスの感染力は維持され、強い多角体プロモーターによって高いレベルのｃＤＮＡの転写が行われる。組換えバキュロウイルスは、多くの場合はＳｐｏｄｏｐｔｅｒａ ｆｒｕｇｉｐｅｒｄａ（Ｓｆ９）昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる（Ｅｎｇｅｌｈａｒｄ．Ｅ．Ｋ．ら（１９９４）Ｐｒｏｃ．Ｎａｔｌ．Ａｃａｄ．Ｓｃｉ．ＵＳＡ９１：３２２４－３２２７、Ｓａｎｄｉｇ，Ｖ．ら（１９９６）Ｈｕｍ．Ｇｅｎｅ Ｔｈｅｒ．７：１９３７－１９４５．等を参照）。

【０３００】

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

【０３０１】

１３　機能的アッセイ

ＴＲＩＣＨの機能は、哺乳動物細胞培養系において生理学的に高められたレベルでのＴＲＩＣＨをコードする配列の発現によって評価する。ｃＤＮＡを、ｃＤＮＡを高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにサブクローニングする。選り抜きのベクターには、ｐＣＭＶ ＳＰＯＲＴプラスミド（Ｌｉｆｅ Ｔｅｃｈｎｏｌｏｇｉｅｓ）及びｐＣＲ ３．１プラスミド（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有する。リポソーム製剤或いは電気穿孔法を用いて、５～１０μｇの組換えベクターをヒト細胞株、例えば内皮由来または造血由来の細胞株に一過的に形質移入する。更に、標識タンパク質をコードする配列を含む１～２μｇ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、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、ｃＤＮＡの組換えベクターからの発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質（ＧＦＰ；Ｃｌｏｎｔｅｃｈ）、ＣＤ６４またはＣＤ６４－ＧＦＰ融合タンパク質から選択できる。自動化された、レーザー光学に基づく技術であ

るフローサイトメトリー（ＦＣＭ）を用いて、ＧＦＰまたはＣＤ６４－ＧＦＰを発現する形質移入された細胞を同定し、その細胞のアポトーシス状態や他の細胞特性を評価する。ＦＣＭは、細胞死に先行するか或いは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、プロピジウムヨウ化物によるＤＮＡ染色によって計測される核ＤＮＡ内容物の変化、前方光散乱と９０°側方光散乱によって計測される細胞サイズと顆粒状性の変化、ブロモデオキシウリジンの取込量の低下によって計測されるＤＮＡ合成の下方調節、特異抗体との反応性によって計測される細胞表面及び細胞内におけるタンパンク質の発現の変化、及び蛍光複合アネキシンＶタンパク質の細胞表面への結合によって計測される原形質膜組成の変化とがある。フローサイトメトリー法については、Ｏｒｍｅｒｏｄ，Ｍ．Ｇ．（１９９４）Ｆｌｏｗ Ｃｙｔｏｍｅｔｒｙ Ｏｘｆｏｒｄ，Ｎｅｗ Ｙｏｒｋ，ＮＹ．に記述がある。

【０３０２】

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

【０３０３】

１４　ＴＲＩＣＨに特異的な抗体の作製

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

【０３０４】

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

【０３０５】

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

【０３０６】

１５　特異的抗体を用いる天然ＴＲＩＣＨの精製

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

洗浄する。

【０３０７】

ＴＲＩＣＨを含む培養液をイムノアフィニティーカラムに通し、ＴＲＩＣＨを優先的に吸着できる条件で（例えば、界面活性剤の存在下において高イオン強度のバッファーで）そのカラムを洗浄する。そのカラムを、抗体とＴＲＩＣＨとの結合を切るような条件で（例えば、ｐＨ２～３のバッファー、或いは高濃度の尿素またはチオシアン酸イオンのようなカオトロピックイオンで）溶出させ、ＴＲＩＣＨを回収する。

【０３０８】

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、Ｇβγタンパク質（Ｒｅｉｍａｎｎ既出）のような調節性タンパク質、又はＭＡＧＵＫｓ（Ｃｒａｖｅｎ既出）の様なＴＲＩＣＨの局在化若しくはクラスター形成に関与するタンパク質が含まれる。

ＴＲＩＣＨまたは生物学的に活性であるＴＲＩＣＨ断片を、^{125}Ｉボルトンハンター試薬で標識する（例えばＢｏｌｔｏｎ Ａ．Ｅ．およびＷ．Ｍ．Ｈｕｎｔｅｒ（１９７３）Ｂｉｏｃｈｅｍ．Ｊ．１３３：５２９－５３９を参照）。マルチウェルプレート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。

【０３０９】

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

【０３１０】

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

【０３１１】

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。

【０３１２】

１７　ＴＲＩＣＨ活性の実証

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

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

【０３１３】

β－ガラクトシダーゼを発現させる形質転換細胞は、適切な比色基質が本技術分野で既知の条件下で培地へ添加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。ＴＲＩＣＨを発現する細胞は、対照細胞に対して高いアニオンコンダクタンス、若しくはカチオンコンダクタンスを有する。ＴＲＩＣＨの伝導度への寄与は、ＴＲＩＣＨに特異的な抗体を用いて細胞をインキュベーションすることによって確認できる。抗体は、ＴＲＩＣＨの細胞外側面に結合し、それによってイオンチャネル内の孔とそれに伴うコンダクタンスをブロッキングする。

【０３１４】

別法では、ＴＲＩＣＨのイオンチャネル活性は、二電極電位クランプ技術（Ｉｓｈｉ ら既出、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を通る電流として測定される。ＴＲＩＣＨは、ｐＢＦのような適切なアフリカツメガエル卵母細胞発現ベクターへとサブクローニングされ、また０．５～５ｎｇのｍＲＮＡが成熟段階４の卵母細胞へ注入される。注入された卵母細胞は摂氏１８度で１－５日間インキュベートされる。インサイド‐アウトマクロパッチが、切り取られて、１１６ ｍＭのＫ－グルコン酸、４ ｍＭのＫＣ１、および１０ ｍＭのＨｅｐｅｓ（ｐＨ ７．２）を含む細胞内溶液へと移される。細胞内溶液には、適切な、ｃＡＭＰ、ｃＧＭＰ、またはCa^{+2}（$CaCl_2$の形態）のような様々な濃度のＴＲＩＣＨ媒介物質が補足される。電極の抵抗は２－５ＭΩでセットされ、電極は媒介物質のない細胞内溶液で満たされる。実験は、室温で０ ｍＶの保持電位で実行される。－１００ｍＶから１００ｍＶの電圧ランプが、サンプリング周波数５００Ｈｚで得られる。電流測定はアッセイに於けるＴＲＩＣＨの活性に比例する。特に、ＴＲＩＣＨ－２５の活性はＣｌ－ の伝導度として測定される。

【０３１５】

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（８２．５ｍＭのＮａＣｌ、２．５ ｍＭのＫＣ１、１ｍＭの$CaCl_2$、１ｍＭの$MgCl_2$、１ｍＭ Na_2HPO_4、５ ｍＭのＨｅｐｅｓ、３．８ ｍＭのＮａＯＨ、５０μｇ／ｍｌ ゲンタマイシン、ｐＨ ７．８）内で摂氏１８度で３日間インキュベートされ、ＴＲＩＣＨの発現を可能とする。卵母細胞は、次に標準取り込み培地（１００ｍＭのＮａＣｌ、２ ｍＭのＫＣ１、１ｍＭの$CaCl_2$、１ｍＭの$MgCl_2$、１０ ｍＭのＨｅｐｅｓ／Ｔｒｉｓ、ｐＨ ７．５）へと移される。様々な基質（例えばアミノ酸、糖、薬剤、イオン、及び神経伝達物質）の取り込みは、標識された基質（例えば^{3}Ｈで放射標識されていたり、ローダミンで蛍光標識されている）を添加することで開始される。３０分間のインキュベートの後、取り込みは、卵母細胞を３回、Na^+のない培養液で洗浄することで終了し、取り込まれた標識を測定して、対照と比較する。ＴＲＩＣＨ活性は、取り込まれた標識基質のレベルに比例する。特に、ＴＲＩＣＨ－１に対する試験基質としてはアミノ酸、ＴＲＩＣＨ－３に対してはキサンチンおよびウラシル、ＴＲＩＣＨ－１８に対してはメリビオス（ｍｅｌｉｂｉｏｓｅ）、ＴＲＩＣＨ－２０に対してはモノカルボン酸、ＴＲＩＣＨ－２２に対してはガンマアミノ酪酸などの神経伝達物質、ＴＲＩＣＨ－２３に対してはヌクレオシドが挙げられる。

【０３１６】

ＴＲＩＣＨに関連するＡＴＰａｓｅ活性は、放射性ラベルされたＡＴＰ－［γ－^{32}Ｐ］の加水分解、クロマトグラフィー法による加水分解生成物の分離、及びシンチレーションカウンターを用いる再生^{32}Ｐの定量化によって測定される。反応混合物は、ＡＴＰ－［γ－^{32}Ｐ］と、好適な時間をかけて摂氏３７度でインキュベートされた好適なバッファ

中の様々な量のＴＲＩＣＨとを有する。反応は、トリクロロ酢酸の沈殿によって終了し、塩基で中和され、反応混合物のアリコットは、反応混合物を分離するべく、膜若しくはペーパークロマトグラフィーにかけられる。遊離された^{32}Ｐの量は、シンチレーション計数器で測定される。回収した放射活性の量はアッセイ中のＴＲＩＣＨのＡＴＰａｓｅ活性に比例する。

【０３１７】

１８　ＴＲＩＣＨアゴニスト及びアンタゴニストの同定

ＴＲＩＣＨは、ＣＨＯ（チャイニーズハムスターの卵巣）またはＨＥＫ（ヒト胎児腎臓）２９３のような真核細胞株内で発現する。形質転換細胞のイオンチャネル活性は、候補アゴニスト又はアンタゴニストの存在若しくは不在条件で測定される。イオンチャネル活性は実施例１７に記載の本技術分野では既知のパッチクランプ法を用いて検定される。別法では、イオンチャネル活性が、細胞膜を通るイオンの流れを測定する蛍光技術を用い検定される（Ｖｅｌｉｃｅｌｅｂｉ，Ｇ．ら（１９９９）Ｍｅｔｈ．Ｅｎｚｙｍｏｌ．２９４：２０－４７；Ｗｅｓｔ，Ｍ．Ｒ．及びＣ．Ｒ．Ｍｏｌｌｏｙ（１９９６）Ａｎａｌ．Ｂｉｏｃｈｅｍ．２４１：５１－５８）。これらアッセイは、ミクロプレートを用いる高処理のスクリーニングに適合しうる。内部イオン濃度の変化は、Ca^{2+}指示薬Ｆｌｕｏ－４ＡＭ、ＳＢＦＩ及びナトリウムグリーンのようなナトリウム感受性色素、Ｃｌ指示薬ＭＱＡＥ（すべてＭｏｌｅｃｕｌａｒ Ｐｒｏｂｅｓ社より入手可能）のような蛍光色素を用い、ＦＬＩＰＲ蛍光定量プレートリーディングシステム（Ｍｏｌｅｃｕｌａｒ Ｄｅｖｉｃｅ社）と併せて測定される。アッセイのさらに一般的な別法では、原形質膜を通るイオンの流れによって生じる膜電位の変化は、$DiBAC_4$（Ｍｏｌｅｃｕｌａｒ Ｐｒｏｂｅｓ社）のようなｏｘｏｎｙｌ色素を用いて測定される。$DiBAC_4$は、細胞膜電位に従って細胞外溶液と細胞内との間で平衡となる。色素の蛍光強度は、疎水性の細胞内部位に結合すると２０倍以上となり、細胞内への$DiBAC_4$の流入が検出可能となる（Ｇｏｎｚａｌｅｚ，Ｊ．Ｅ．及びＰ．Ａ．Ｎｅｇｕｌｅｓｃｕ（１９９８）Ｃｕｒｒ．Ｏｐｉｎ．Ｂｉｏｔｅｃｈｎｏｌ．９：６２４－６３１）。候補アゴニスト若しくはアンタゴニストは、既知のイオンチャネルアゴニスト、若しくはアンタゴニスト、ペプチドライブラリ、若しくは組み合わせの化学ライブラリより選択されてもよい。

【０３１８】

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

【０３１９】

（表の簡単な説明）

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

【０３２０】

表２は、ＧｅｎＢａｎｋ識別番号及び本発明のポリペプチドに最も近いＧｅｎＢａｎｋ相同体の注釈を示す。各ポリペプチドとそのＧｅｎＢａｎｋ相同体が一致する確率スコアも併せて示す。

【０３２１】

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

【０３２２】

表４は、本発明のポリヌクレオチド配列を構築するために用いたｃＤＮＡやゲノムＤＮＡ

断片を、ポリヌクレオチド配列の選択した断片と共に示す。

【０３２３】

表５は、本発明のポリヌクレオチドの代表的なｃＤＮＡライブラリを示す。

【０３２４】

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

【０３２５】

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

【表１】

表1

Incyte プロジェクト ID	ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID
1687189	1	1687189CD1	27	1687189CB1
7078207	2	7078207CD1	28	7078207CB1
1560619	3	1560619CD1	29	1560619CB1
2614283	4	2614283CD1	30	2614283CB1
2667691	5	2667691CD1	31	2667691CB1
3211415	6	3211415CD1	32	3211415CB1
4739923	7	4739923CD1	33	4739923CB1
5503045	8	5503045 9CD1	34	5503045 9CB1
6113039	9	6113039CD1	35	6113039CB1
7101781	10	7101781CD1	36	7101781CB1
7473036	11	7473036CD1	37	7473036CB1
7476943	12	7476943CD1	38	7476943CB1
8003355	13	8003355CD1	39	8003355CB1
3116448	14	3116448CD1	40	3116448CB1
622868	15	622868CD1	41	622868CB1
7476494	16	7476494CD1	42	7476494CB1
7477260	17	7477260CD1	43	7477260CB1
1963058	18	1963058CD1	44	1963058CB1
2395967	19	2395967CD1	45	2395967CB1
3586648	20	3586648CD1	46	3586648CB1
7473396	21	7473396CD1	47	7473396CB1
7476283	22	7476283CD1	48	7476283CB1
7477105	23	7477105CD1	49	7477105CB1
7482079	24	7482079CD1	50	7482079CB1
55145506	25	55145506CD1	51	55145506CB1
5950519	26	5950519CD1	52	5950519CB1

【表２】

表2－1

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	確率スコア	GenBank 相同体
1	1687189CD1	g2116552	7.80E-275	[ラット] カチオンアミノ酸輸送体 3 (Hosokawa,H. 他 (1997) J. Biol. Chem. 272 (13), 8717-8722)
2	7078207CD1	g495259	0	[マウス] abc2 (Luciani,M.F. 他 (1994) Genomics 21 (1), 150-159)
3	1560619CD1	g1002424	8.60E-253	[マウス] YSPL-1 フォーム 1 (Guimaraes,M.J. 他 (1995) Development 121 (10), 3335-3346)
4	2614283CD1	g1256378	1.90E-152	[ラット] 亜鉛輸送体 ZnT-2 (Palmiter,R.D. 他 (1996) EMBO J. 15 (8), 1784-1791)
5	2667691CD1	g2506078	2.30E-259	[マウス] テトラサイクリン輸送体様タンパク質 (Matsuo,N. 他 (1997) Biochem. Biophys. Res. Commun. 238 (1), 126-129)
6	3211415CD1	g7243710	9.80E-197	[マウス] 亜鉛輸送体様 2
7	4739923CD1	g13785620	4.00E-96	[3' incom] [マウス] siderofiexin 5 (Fleming,M.D. 他 (2001) Genes Dev. 15 (6), 652-657)
8	5503045CD1	g4186073	9.40E-15	[マウス] カルシウムチャネルアルファ 2デルタ C サブユニット (Klugbauer,N. 他 (1999) J. Neurosci. 19(2), 684-691)
9	6113039CD1	g310183	5.00E-273	[ラット] ナトリウム依存性硫酸輸送体 (Markovich,D. 他 (1993) Proc. Natl. Acad. Sci. U.S.A. 90, 8073-8077)
10	7101781CD1	g13506808	0	[fil] [マウス] 胸腺間質共役輸送体 (Chen,C. 他 (2000) Biochim. Biophys. Acta 1493 (1-2), 159-169)
11	7473036CD1	g13249295	0	[fil]ヒト陰イオン交換体 AE4 (Palmiter,R.D. 他 (2001) Biochem. Biophys. Res. Commun. 282 (5), 1103-1109)
12	7476943CD1	g3047402	5.00E-67	[ヒト] モノカルボン酸輸送体 2

【表３】

表2－2

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	確率スコア	GenBank 相同体
13	8003355CD1	g825618	3.00E-273	[ヒト] ach_cds (Shibahara,S. 他 (1985) Eur. J. Biochem. 146 (1), 15-22)
14	3116448CD1	g10732815	0	[fl][ヒト] 濃縮性 Na+-スクレオシド共役輸送体 hCNT3 (Ritzel,M.W.L. 他 (2001) J. Biol. Chem. 276 (4), 2914-2927)
15	622868CD1	g5924012	9.50E-160	[ヒト] dJ261K5.1 (新規の有機カチオン輸送体 (BAC ORF RG331P03))
16	7476494CD1	g8979801	3.90E-147	[ヒト] dJ37C10.3 (新規の ATP 分解酵素)
17	7477260CD1	g13447747	0	[fl][ヒト] 重炭酸ナトリウム共役輸送体 NBC4a (Pushkin,A. 他 (2000) IUBMB Life 50 (1), 13-19)
18	1963058CD1	g1653342	6.80E-18	[ラン藻] メリビオス(melibiose)担体タンパク質 (Kaneko,T. 他 (1995) DNA Res. 2 (4), 153-166)
19	2395967CD1	g37643	3.20E-129	[ヒト] 液胞プロトン ATP 分解酵素 (van Hille,B. 他 (1993) Biochem. Biophys. Res. Commun. 197 (1), 15-21)
20	3586648CD1	g2198807	8.90E-49	[ニワトリ] モノカルボン酸輸送体 3
		g2463628	6.00E-43	[fl][ヒト] 推定上のモノカルボン酸輸送体
21	7473396CD1	g2618842	1.10E-139	[粘菌] excinuclease ABC サブユニット (Reizer,J. 他 (1998) Mol. Microbiol. 27 (6), 1157-1169)A
22	7476283CD1	g56176	4.40E-244	[ドブネズミ] GABA(A) 受容体ガンマ 1 サブユニット (Ymer,S. 他 (1990) EMBO J. 9 (10), 3261-3267)
23	7477105CD1	g3176684	2.20E-11	[シロイヌナズナ] ヒトの equilibrataive スクレオシド輸送体 gb|U81375 に類似のものを含む。ESTs gb|N65317, gb|T20785, gb|AA586285 およびgb|AA712578 はこの遺伝子に由来
		g12656639	3.00E-05	[fl][ヒト] equilibrative スクレオシド輸送体

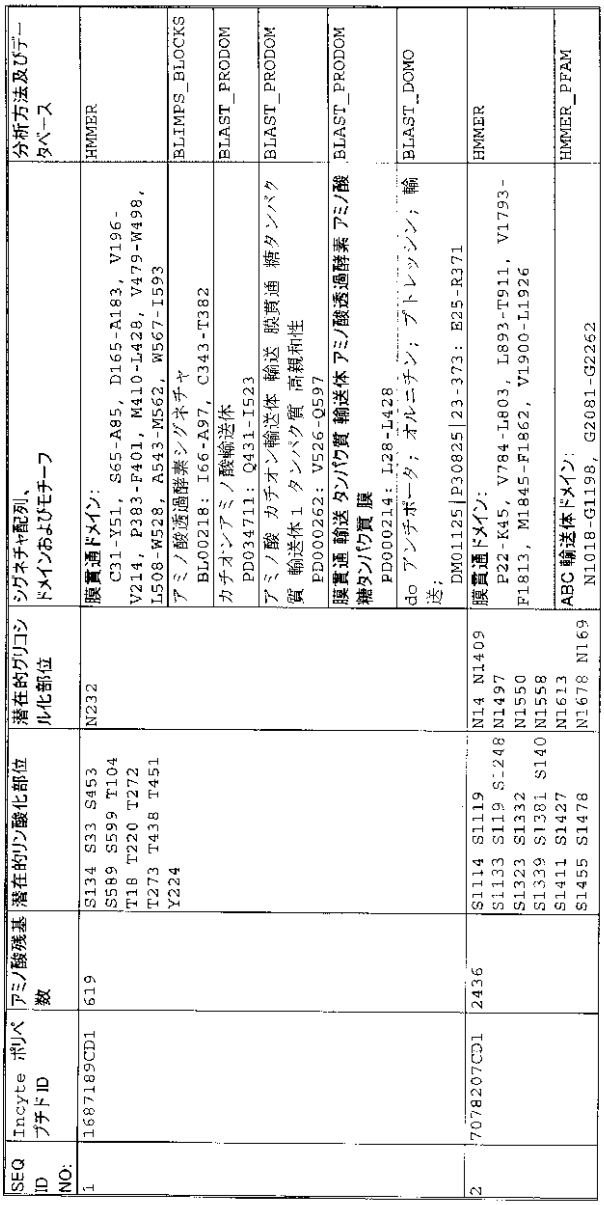
【表４】

表2－3

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	確率スコア	GenBank 相同体
24	7482079CD1	g2815899	9.60E-84	[ヒト] Shab 関連遅延整流 K+チャンネルアルファ (Shepard,A.R. 他 (1999) Am. J. Physiol. 277 (3), C412-C424)
25	5514550 6CD1	g289404	4.70E-105	[ウシ] クロライドチャネルタンパク質 (Landry,D. 他 (1993) J. Biol. Chem. 268, 14948-14955)
26	5950519CD1	g2352427	6.40E-156	[ウサギ] ベルオキシソーム Ca 依存性溶質担体 (Weber,F.E. 他 (1997) Proc. Natl. Acad. Sci. U.S.A. 94 (16), 8509-8514)

【表５】

表3－1



SEQ ID NO:	Incyte ポリペプチドID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
1	1687189CD1	619	S134 S33 S453 S589 S599 T104 T18 T220 T272 T273 T438 T451 Y224	N232	膜貫通ドメイン: C31-Y51, S65-A85, D165-A183, V196-V214, P383-F401, M410-L428, V479-W498, L508-W528, A543-M562, W567-I593	HMMER
					アミノ酸透過酵素シグネチャ BL00218: I66-A97, C343-T382	BLIMPS_BLOCKS
					カチオンアミノ酸輸送体 PD034711: Q431-I523	BLAST_PRODOM
					アミノ酸 カチオン輸送体 輸送 脱貫通 膵タンパク質 輸送体1 タンパク質 高親和性 PD000262: V526-Q597	BLAST_PRODOM
					膜貫通 輸送 タンパク質 輸送体 アミノ酸透過酵素 アミノ酸 糖タンパク質 膜 PD000214: I28-L428	BLAST_PRODOM
					do アンチポータ; オルニチン; プトレッシン; 輸送; DM01125|P30825|23-373: E25-R371	BLAST_DOMO
2	7078207CD1	2436	S1114 S1119 S1133 S119 S-248 S1323 S1332 S1339 S1381 S140 S1411 S1427 S1455 S1478	N14 N1409 N1497 N1550 N1558 N1613 N1678 N169	膜貫通ドメイン: P22-K45, V784-L803, L893-T911, V1793-F1813, M1845-F1862, V1900-L1926	HMMER
					ABC 輸送体ドメイン: N1018-G1198, G2081-G2262	HMMER_PFAM

【表６】

表3－2

SEQ ID NO:	Incyte ポリペプチドID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
			S1560 S1604 S1687 S1819 S1982 S199 S2024 S2036 S2062 S21 S2159 S2196 S2245 S2292 S2333 S2366 S2420 S256 S281 S467 S50 S502 S533 S631 S884 S940 S959 S971 T1058 T1081 T1212 T1271 T1313 T1314 T1532 T16 T2097 T2102 T2108 T2-44 T2215 T2235 T2284 T2352 T2413 T252 T353 T382 T440 T48 T612 T633 T696 T844 T955 Y1390	N174 N1776 N2055 N306 N369 N380 N421 N433 N477 N485 N495 N531 N545 N591 N601 N629 N90	ABC 輸送体ファミリシグネチャ: D1105-D1155, V2167-D2218	PROFILESCAN
2					ATP 結合 輸送体カセット ABC 輸送タンパク質 膜貫通 リム ABCR PD005939: L1787-Y1971 PD006867: I663-S809 PD006285: A811-K1001 PD010138: G1957-K2061	BLAST_PRODOM

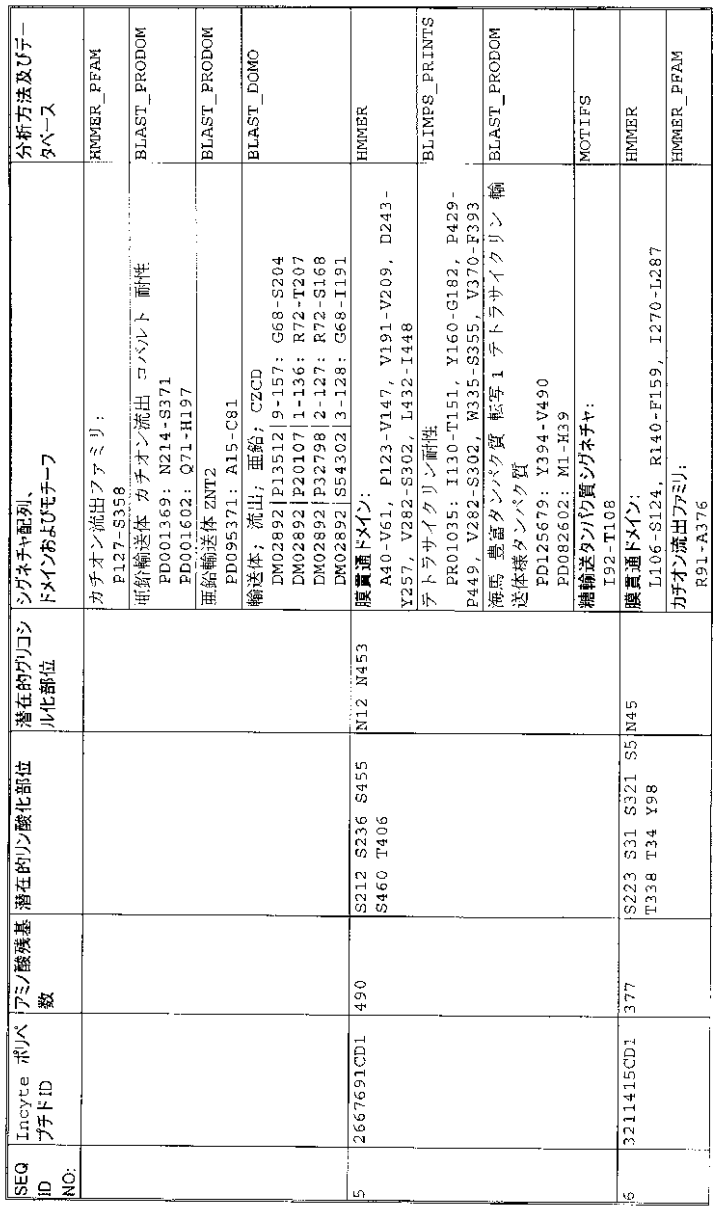
【表 7 】

表3－3

SEQ ID NO:	Incyte ポリペプチドID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					ABC 輸送体ファミリ DM00008|P41233|839-1045: V991-H1197, V2051-M2259	BLAST_DOMO
					ABC 輸送体モチーフ: L1124-F1138	MOTIFS
					ATP/GTP 結合部位(P-ループ): G1025-T1032, G2088-T2095	MOTIFS
					Lipocalin モチーフ: G1424-V1437, G1426-V1437	MOTIFS
3	15606l9CD1	610	S127 S161 S251 S409 S450 S483 S582 S601 S608 T313 T514 T529	N139 N159	**膜貫通ドメイン:** F215-C233, L264-I286	HMMER
					キサンチン/ウラシル透過酵素ファミリドメイン: G46-E473	HMMER_PFAM
					キサンチン/ウラシル透過酵素シグネチャ BL01126: G407-F443	BLIMPS_BLOCKS
					卵黄嚢透過酵素様 YSPL1 FORM1 **卵黄嚢**透過酵素様 YSPL1 FORM4 **卵黄嚢**透過酵素様 YSPL1 FORM2 **卵黄嚢**透過酵素様 YSPL1 FORM3 卵黄嚢透過酵素様 PD019501: G429-Q609 PD137940: Q29-P83	BLAST_PRODOM
					タンパク質輸送 硫酸輸送体 膜貫通 透過酵素 遺伝子間領域 親和性 糖タンパク質 PD001255: L174-L467	BLAST_PRODOM
4	2614283CD1	372	S124 S216 S338 S61 T281		キサンチン/ウラシル透過酵素ファミリ DM01485|S33349|7-188: G355-L465	BLAST_DOMO
					膜貫通ドメイン: I141-V159	HMMER

【表８】

表3－4



SEQ ID NO:	Incyte ポリペプチドID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					カチオン流出ファミリ; P127-S358	HMMER_PFAM
					亜鉛輸送体 カチオン流出 コバルト 耐性 PD001369: N214-S371 PD001602: Q71-H197	BLAST_PRODOM
					亜鉛輸送体 ZNT2 PD095371: A15-C81	BLAST_PRODOM
					輸送体; 流出; 亜鉛; CZCD DM02892|P13512|9-157: G68-S204 DM02892|P20107|1-136: R72-T207 DM02892|P32798|2-127: R72-S168 DM02892|S54302|3-128: G68-I191	BLAST_DOMO
5	2667691CD1	490	S212 S236 S455 S460 T406	N12 N453	**膜貫通ドメイン**: A40-V61, P123-V147, V191-V209, D243-Y257, V282-S302, L432-I448	HMMER
					テトラサイクリン耐性 PR01035: I130-T151, Y160-G182, P429-P449, V282-S302, W335-S355, V370-F393	BLIMPS_PRINTS
					海馬 豊富タンパク質 転写 1 テトラサイクリン 輸送体様タンパク質 PD125679: Y394-V490 PD082602: M1-H39	BLAST_PRODOM
					糖輸送タンパク質シグネチャ: I92-T108	MOTIFS
6	3211415CD1	377	S223 S31 S321 S5 T338 T34 Y98	N45	**膜貫通ドメイン**: L106-S124, R140-F159, I270-L287	HMMER
					カチオン流出ファミリ; R91-A376	HMMER_PFAM

【表９】

表3－5

SEQ ID NO:	Incyte ポリペプチドID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
6					亜鉛 輸送体 カチオン 流出 コバルト 耐性 PD001602:L38-V158	BLAST_PRODOM
					輸送体; 流出; 亜鉛; CZCD DM02892|S61568|396-545: D32-H166 DM02892|P20107|1-136: I29-S170 DM02892|P13512|9-157: R36-G162 DM02892|P32798|2-127: L42-I154	BLAST_DOMO
7	4739923CD1	340	S240 S272 S314 S319 S330 T56 T74	N127 N140 N153	染色体 推定 輸送体 C17G6.15C 輸送体 XV リーディングフレーム PD006986: S20-L274	BLAST_PRODOM
8	55030459CD1	1274	S1025 S1138 S1142 S1155 S1189 S1201 S1242 S134 S190 S238 S256 S298 S303 S353 S354 S40 S405 S430 S624 S664 S670 S700 S746 S79 S892 S894 S952 T1006 T1050 T1191 T221 T268 T272 T293 T349 T361 T581 T674 T717 T75 T755 T813 T852 T868 T987 Y1056 Y114	N145 N329 N373 N568 N587 N905 N940 N985	signal_cleavage: M1-A35 膜貫通ドメイン: V1096-R1118	SPSCAN HMMER
9	6113039CD1	595	S213 S214 S483 S74 T209 T230 T236 T240 T423	N140 N174 N207 N591	膜貫通ドメイン: Y10-L30, P287-W305, V349-D365, G556-M576	HMMER

【表１０】

表3－6

SEQ ID NO:	Incyte ポリペプチドID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
			T97 Y39		ナトリウム:硫酸シンポータ BL01271: T131-I150, T240-I264, P432-G453, A505-I559 共役輸送体のナトリウムシンポート PD000549: B331-W572, L242-K402, V16-A167, F13-V154 ナトリウム/硫酸 共役輸送体 NA+/硫酸 輸送 膜貫通 ナトリウム シンポート PD084897: A161-K238	BLIMPS_BLOCKS BLAST_PRODOM BLAST_PRODOM
10	7101781CD1	475	S100 S108 S170 S34 S61 T20 T252 T390	N55	do 腎臓; 結合; PRO-SER-ALA; NA; DM02914|A47714|28-576: I28-F577 DM02914|S43561|28-507: L242-I569, E34-A161 DM02914|P46556|1-520: K198-F577, E34-I159 DM02914|P32739|25-517: K238-F577, E34-V154 膜貫通ドメイン: I283-V308, Y322-V340, M350-E370, S440-V459	BLAST_DOMO HMMER
11	7473036CD1	927	S149 S163 S217 S23 S260 S265 S325 S51 S65 S733 S738 S874 S891 S904 S906 T292 T324 T344 T567 T594 T629 T802 T99	N493 N520 N544 N923	膜貫通ドメイン: V444-Y466, V761-P780, I779-M810	HMMER
11					HCO3- 輸送体ファミリ: K108-I835	HMMER_PFAM

【表１１】

表3－7

SEQ ID NO:	Incyte ポリペプチドID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					アニオン交換体ファミリ BL00219: V360-D383, W659-L700, G744-L789, Y790-T833, G89-H120, Q180-L223	BLIMPS_BLOCKS
					アニオン交換体ファミリシグネチャ: A457-G509	PROFILESCAN
					アニオン交換体シグネチャ PR00165: Q355-G375, V388-G407, L442-S461, G474-L492, D570-L589, W657-M676	BLIMPS_PRINTS
					アニオン交換体 糖タンパク質 パルミチン酸 重炭酸共役輸送体 PD001455: S346-L784, S505-I835, S156-E348, L109-V154	BLAST_PRODOM
					重炭酸 共役輸送体 起電性 NA+ 膵臓 HCO3 P52B5.1 PD018437: Q836-N927	BLAST_PRODOM
					バンド 3 アニオン輸送タンパク質 DM02294|P04920_602-1237: G558-E894, S346-P529 DM02294|P48751|601-1229: S537-G896, S346-I543 DM02294|A42497|403-1027: S537-G896, S346-I500 DM02294|P02730|311-908: P560-D882, S346-G519	BLAST_DOMO
12	7476943CD1	516	S11 S137 S169 S202 S253 S41 S92 T228 T234 T244 T30 T340	N10 N333 N487	膜貫通ドメイン: I118-T144, S181-W203, A206-M224, Y275-M293	HMMER
12					モノカルボン酸輸送体: C55-D499	HMMER_PFAM

【表１２】

表3－8

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					PEST; 輸送体; 関連: DM05037|P53988|1-465: P42-Q470 DM05037|Q03064|1-475: S41-D479 DM05037|P36021|155-612: A37-L258, V285-E477	BLAST_DOMO
13	8003355CD1	514	S174 S183 S330 S427 S453 S54 S64 T381 T382 Y94	N163 N328 N373 N52	signal_cleavage: M1-G22	SPSCAN
					シグナルペプチド: M1-G22	HMMER
					膜貫通ドメイン: P241-F264, C274-V291, Y308-N328, V472-M491	HMMER
					神経伝達物質依存性イオンチャネル: E26-F489	HMMER_PFAM
					神経伝達物質依存性イオンチャネルタンパク質 BL00236: V107-N116, D135-Y173, H228-A269, V53-D90	BLIMPS_BLOCKS
					神経伝達物質依存性イオンチャネルシグネチャ: V130-Q184	PROFILESCAN
					神経伝達物質依存性イオンチャネルファミリシグネチャ: PR00252: T73-R89, M106-N117, C150-C164, L235-N247	BLIMPS_PRINTS
					ニコチン性 アセチルコリン受容体シグネチャ PR00254: T60-V76, Y94-W108, I112-V124, V130-S148	BLIMPS_PRINTS
13					チャネル イオン性 糖タンパク質 シナプス後受容体 シグナル タンパク質 PD000153: N24-S393, A432-F489	BLAST_PRODOM

【表１３】

表3－9

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					神経伝達物質依存性イオンチャネル: DM00195|P13536|7-501: P7-V497 DM00195|P02713|5-498: L8-R496 DM00195|P05376|2-493: L10-R496 DM00195|P02714|1-491: L8-V497	BLAST_DOMO
					神経伝達物質依存性イオンチャネルシグネチャ: C150-C164	MOTIFS
14	3116448CD1	691	S326 S36 S549 S582 S63 S669 T100 T193 T262 T356 T411 T417 T50 T615 T637 Y87	N30 N34 N630 N636 N664	膜貫通ドメイン: I104-N124, W178-L207, L289-M308, I444-L461	HMMER
					Na+依存性ヌクレオシド輸送体 Nucleoside_tra2: Q198-S613	HMMER_PFAM
					銅輸送 ATP 分解酵素 L131-D145	BLIMPS_PRINTS
					NA+/ヌクレオシド 膜内輸送 PD003768: R223-I611, PD008773: F93-F215	BLAST_PRODOM
					ヌクレオシド; 輸送体; Na 依存性 DM01857|A54892|234-589: L256-L612 DM01857|A57532|230-585: L256-L612 DM01857|P44742|60-409: V260-I611 DM01857|P33021|60-412: V260-G610	BLAST_DOMO
15	622868CD1	342	S102 S309 S315 S325 S84 T121 T174 T286 T299 T300 T334	N110 N117 N311 N323	膜貫通ドメイン: Y205-Y227	HMMER
					ペリフェリン (RDS)/ROM-1 F PR00218: V9-V29, L207-L228	BLIMPS_PRINTS
					糖タンパク質シグネチャ PR00171: A231-V242	BLIMPS_PRINTS
					DM00135|P39932|141-478: W33-K295	BLAST_DOMO

【表１４】

表3－10

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
16	7476494CD1	791	S103 S110 S199 S289 S514 S659 S66 S688 S734 S782 T122 T191 T287 T314 T326 T507 T710 T747 T78 Y293 Y742	N697 N768	Atpase_E1_E2 D437-T443 膜貫通ドメイン: A177-Y193, D348-Y366 E1-E2 ATP 分解酵素 ATPase E1-E2_ATPase: C217-T443, P551-R680 E1-E2 ATP 分解酵素リン酸化部位 atpase_e1_e2.prf: I417-A471 E1-E2 ATP 分解酵素リン酸化部位 BL00154: V393-G429, L431-L449, K575-C585, N644-M684 P 型カチオン輸送 ATP 分解酵素スーパファミリシグネチャ PR00119: D260-T274, C435-L449, A660-D670 ナトリウム/カリウム輸送体 ATP 分解酵素シグネチャ PR00121: C428-L449, A572-V590 E1-E2 ATP 分解酵素 リン酸化部位 DM00115|P22189 49-801: L547-V685 DM00115|P37278 58-755: Q224-I692 DM00115|A42764 65-737: E141-T699 DM00115|P37367 60-746: L226-V691 ATP 結合 カルシウム マグネシウム 輸送 ポンプ PD000132: I180-D445, A612-Q689, M559-C585	MOTIFS HMMER HMMER_PFAM PROFILESCAN BLIMPS_BLOCKS BLIMPS_PRINTS BLIMPS_PRINTS BLAST_DOMO BLAST_PRODOM
17	7477260CD1	1108	S1012 S1061 S1063 S1088 S124	N399 N653 N658 N668	遺伝子制御モチーフ Leucine_Zippe L125-L146, L677-L698	MOTIFS

【表１５】

表3－11

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
			S14 S190 S218 S240 S314 S319 S388 S391 S434 S435 S686 S701 S870 S95 T1030 T1056 T1065 T1093 T102 T106 T183 T201 T454 T639 T678 T725 T766 T778 T78 Y1090	N676	アニオン交換体ファミリシグネチャ anion_exchanger1.prf: D438-F490 anion_exchanger2.prf: A585-T639 膜貫通ドメイン: I488-L506, L837-W856, I898-P917, V920-F938, I982-V1002 HCO3- 輸送体ファミリHCO3_cotransp: K104-V972 アニオン交換体ファミリ BL00219: H85-H116, K259-V302, T304-K342, A343-K378, G448-A487, I488-D511, L541-Q579, L581-I628, P706-D759, V796-L837, D838-E876, G881-L926, Y927-T970, V972-S1011 アニオン交換体シグネチャ PR00165: F458-F480, Q483-G503, V516-G535, I539-S558, L570-S589, G602-I620, D707-L726, L742-F762, W794-M813 バンド3アニオン輸送タンパク質 DM02294|P48751|601-1229: P706-N1020, E449-P634, I353-E367 タンパク質 アニオン 交換 膜貫通 バンド 糖タンパク質 リポタンパク質 バルミチン酸 重炭酸 共役輸送体 PD001455: H445-V972, V105-E394	PROFILESCAN HMMER HMMER_PFAM BLIMPS_BLOCKS BLIMPS_PRINTS BLAST_DOMO BLAST_PRODOM

【表１６】

表3－12

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					重炭酸 共役輸送体 ナトリウム 起電性 NA+ 膵臓 PD018437: Q973-M1078 PD018439: A53-E103	BLAST_PRODOM

【表１７】

表3－13

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
18	1963058CD1	480	S13 S194 S195 S204 S409 S49 S54 T286	N178 N219 N292 N341	膜貫通ドメイン: L349-Y369 ナトリウム:カラクトシンポータ ファミリー: DM01084|Q02581|1-462:L17-S195 (P 値 = 8.2e-10)	HMMER BLAST-DOMO
19	2395967CD1	381	S119 S170 S202 S211 S269 S327 S349 S378 S74 T102 T144 T147 T164 T246 T26 T328 T62	N192 N285 N30	液胞 ATP 分解酵素 C サブユニット PD014267: E3-D376 液胞 ATP シンターゼ: DM04365 P21282|1-381: M1-D381 DM04365 P54648|1-368: E3-L342 DM04365 P31412|1-392: I7-L377	BLAST-PRODOM BLAST-DOMO
20	3586648CD1	484	S236 S4 T21 T258 T290 T3 T312 Y301	N345 N389	膜貫通ドメイン: F42-W61, V75-I94, F311-Y327, I361-W382, W382-M402 モノカルボン酸輸送体: S40-L478 輸送体 DM05037|P53988|1-465: P16-N217 DM05037|Q03064|1-475: D29-Q263 DM05037|P36021|155-612: D29-L229	HMMER HMMER-PFAM BLAST-DOMO
21	7473396CD1	736	S236 S440 S462 S472 S501 S52 S579 S626 T10 T166 T191 T239 T316 T324 T345 T386 T491 T587 T607 T715 T89	N3 N367	シグナルペプチド: M1-G53 ABC 輸送体: G24-G210, G429-G700 ABC 輸送体: L396-V410, L625-L639 ATP/GTP 結合部位: G31-S38, G436-S443	SPScan HMMER-PFAM MOTIFS MOTIFS

【表１８】

表3－14

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					ABC 輸送体ファミリシグネチャ: Q606-H659, L378-D427	ProfileScan
					ABC 輸送体ファミリ: BL00211: L396-D427, L29-L40	BLIMPS-BLOCKS
					UVRA タンパク質 DM02034|P13567|759-959: F503-G704, I135-D202 DM02034|P07671|708-908: F503-G704, I135-D202 DM02034|S49424|2-201: D504-G704, K110-K198 DM02034|P47660|610-810: F503-G704, I135-D202	BLAST-DOMO
					Excinuclease ABC サブユニット A PD001646: D504-T624	BLAST-PRODOM
					Excinuclease ABC サブユニット A PD184930: C538-T713, R133-L297, V434-V502	BLAST-PRODOM
					Excinuclease ABC サブユニット A PD003881: M447-F503	BLAST-PRODOM
					リボース/ガラクトース ABC 輸送体: PD035715: K241-K311, M1-I55	BLAST-PRODOM
22	7476283CD1	465	S153 S224 S267 S426 T118 T129 T179 T331 T88	N127 N245 N393 N50	シグナルペプチド: M1-C35	SPScan
					膜貫通ドメイン: M270-I294	HMMER
					神経伝達物質依存性イオンチャネル: I64-W459	HMMER-PFAM

【表１９】

表3－15

SEQ ID NO:	Incyte ポリペプチドID	アミノ酸残基数	潜在的N-酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					神経伝達物質依存性イオンチャネルシグネチャ: L168-K222	ProfileScan
					神経伝達物質依存性イオンチャネル: C188-C202	MOTIFS
					神経伝達物質依存性イオンチャネル BL00236: I90-N127, I143-N152, D173-Y211, Y257-A298	BLIMPS-BLOCKS
					神経伝達物質依存性イオンチャネル: PR00252: T110-F126, K142-S153, C188-C202, F264-Q276	BLIMPS-PRINTS
					ガンマアミノ酪酸受容体: PR00253: F273-W293, V299-A320, M333-L354, Y442-Y462	BLIMPS-PRINTS
					ガンマアミノ酪酸受容体: PR01079: G62-Q73, D82-I99, F125-N138, W233-G255, K326-V339, I432-R444, V457-L465	BLIMPS-PRINTS
					神経伝達物質依存性イオンチャネル: DM00560|P23574|26-465: L26-L465 DM00560|P20237|20-556: L26-V396, A437-L463 DM00560 P16305|4-443: D63-S377, A437-L463 DM00560|P08219|14-456: T65-L463	BLAST-DOMO
					イオンチャネル/シナプス後膜受容体 PD000153: N127-Y356, Q66-V286	BLAST-PRODOM
22					イオンチャネル/シナプス後膜受容体 PD000604: G403-L463	BLAST-PRODOM

【表２０】

表3－16

SEQ ID NO:	Incyte ポリペプチドID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
23	7477105CD1	235	S151 S6 T218 T56 T57 T90		膜貫通ドメイン: S103-N123, I136-R162 スクレオンド輸送体, equilibrative: PD006749: P63-L157 (P-value = 1.0e-07)	HMMER BLAST-PRODOM
24	7482079CD1	662	S10 S12 S137 S211 S323 S5 S564 T130 T19 T195 T281 T403 T499 T627 T657 T83 Y187	N17 N440 N517	膜貫通ドメイン: G412-Y430 K+ チャネル四量体化ドメイン: S97-F203 イオン輸送タンパク質: G263-L609 カリウムチャネルシグネチャ PR00169: Q410-E433, F441-L463, G587-F613, E148-S167, P253-T281, H304-L327, F330-L350, L381-C407 カリウムチャネル CDRK: DM00436|JH0595|144-307: K215-L390 DM00436|P15387|136-299: R206-L381 DM00436|P17970|386-549: I216-L390 DM00490|P17970|268-384: A94-R200 電圧依存性カリウムチャネル: PD000141: P330-S469, V570-K619, I572-I645	HMMER HMMER-PFAM HMMER-PFAM BLIMPS-PRINTS BLAST-DOMO BLAST-PRODOM
25	55145506CD1	371	S113 S158 S194 S330 T151 T211 T323 T341 T63 T8	N127 N209	タンパク質 チャネル イオン性 イオン 輸送 電位依存性 P64 クロライド 細胞質内塩素 PD017366: A169-K355 塩素 チャネル タンパク質 P64 イオン性 イオン 輸送 電位依存性 膜貫通 リン酸化 PD118116: M1-Q125	BLAST_PRODOM BLAST_PRODOM

【表２１】

表3－17

SEQ ID NO:	Incyte ポリペプチドID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
26	5950519CD1	468	S105 S176 S23 S4 S56 T161 T170 T220 T308 T358 T466		ミトコンドリア担体タンパク質ドメイン: M184-T276, H278-H369, G375-R468	HMMER_PFAM
					EF ハンド: R13-L41, R81-L109, Q117-H145	HMMER_PFAM
					ミトコンドリアエネルギー伝達タンパク質シグネチャ BL00215: V190-Q214, I425-G437	BLIMPS_BLOCKS
					ミトコンドリアエネルギー伝達タンパク質シグネチャ: K187-L241, V279-P331, I376-Q428	PROFILESCAN
					ミトコンドリア 担体 タンパク質シグネチャ PR00926: Q188-T201, T201-V215, G244-E264, T292-R310, Y335-L353, G383-Q405	BLIMPS_PRINTS
					グレーブス病抗体タンパク質シグネチャ PR00928: P205-I225	BLIMPS_PRINTS
					タンパク質輸送 膜貫通 リピート ミトコンドリア 担体 膜内 ミトコンドリア ADP/ATP PD000117: Q273-L463, K187-A293	BLAST_PRODOM
					ミトコンドリアエネルギー伝達タンパク質 DM00026|S57544|26-107: V190-I270 DM00026|P29518|233-310: V284-K360 DM00026|S54495|534-620: F283-N361 DM00026|Q01888|126-214: H278-N361	BLAST_DOMO
					EF ハンドモチーフ: D22-L34, D90-I102	MOTIFS
					ミトコンドリア担体タンパク質モチーフ: P299-L307	MOTIFS

【表２２】

表4－1

ポリヌクレオチド SEQ ID NO:	Incyte ポリヌクレオチド ID	ポリ	配列長	選択された断片	配列断片	5'位置	3'位置
27	1687189CB1		2229	2190-2229, 759-1660	7056238V1	1269	1824
					7251266F7 (PROSTMY01)	1	693
					7030002J3D1	2059	2229
					7749453F8 (NOSEDIN01)	410	1006
					7056521SV1	1741	2207
					2416733F6 (HNT3AZT01)	1147	1688
					7711767J1 (TESTTUE02)	756	1205
28	7078207CB1		7610	1-5580	7070225H1 (BRAUTDR02)	6275	6895
					6911060J1 (PITUDIR01)	3289	3865
					6772031H1 (BRAUNOR01)	1	696
					710631B3V1	2885	3400
					6253219H1 (LUNFTUT02)	1442	2064
					6893301H1 (BRAITDR03)	5879	6157
					7106586OV1	2228	2898
					6763740H1 (BRAUNOR01)	3901	4568
					6765621H1 (BRAUNOR01)	3843	4506
					7467144H1 (LUNGNOE02)	3406	3914
					3767813H1 (BRSTNOT24)	5937	6250
					6953905H1 (BRAITDR02)	5169	5868
					6977243H1 (BRAHTDR04)	4508	5140
					8016696J1 (BMARTXE01)	2071	2870
					5964168H1 (BRATNOT05)	6042	6690
					6762808J1 (BRAUNOR01)	2849	3387
					6950389H1 (BRAITDR02)	725	1478
					4098906F8 (BRAITUT26)	6463	7085
					7757265J1 (SPLNTUE01)	660	1225
					5098681F8 (EPIMNON05)	7071	7610
					7179893H1 (BRAXDIC01)	6909	7418
					7196965JV1	1561	2109
					6908865J1 (PITUDIR01)	4613	5241
					6893778J1 (BRAITDR03)	5243	5900
29	1560619CB1		2219	1-1659	6452362F8 (COLNDIC01)	1	553
					7159747AV1	1539	2219
					7159478AV1	1331	2107

【表２３】

表4－2

ポリヌクレオチドSEQ ID NO:	Incyte ヌクレオチドID	ポリ	配列長	選択された断片	配列断片	5'位置	3'位置
					70683177V1	1272	1783
					70680523V1	738	1346
					71596281V1	364	920
30	2614283CB1		1280	415-559	60202200D1	385	843
					809735 2H1 (EYERNOA01)	850	1280
					7432729H1 (PANCDIR02)	425	1006
30					7987760H1 (UTRSTUC01)	1	439
31	2667691CB1		2727	1-330	5753102H1 (LUNGNOT35)	1332	1994
					71103888V1	654	1359
					GBI.g8081479.smoosh	1	266
					7312933H1 (SINTNON02)	1968	2591
					70233893V1	266	752
					GBI.g9988362.smoosh	168	326
					8093950H1 (EYERNOA01)	741	1387
					7925641H2 (COLNTUS02)	2052	2727
					7346378H1 (SYNODIN02)	1395	2014
32	3211415CB1		1631	1-43, 1303-1631	70062244V1	704	1145
					5313165F8 (KIDETXS02)	1	719
					70059213V1	1170	1631
					70057909V1	1010	1547
33	4739923CB1		2673	1483-1785, 1-37	7198215 0V1	1187	1830
					7198685 6V1	1513	2125
					4567241F7 (HELATXT01)	2337	2673
					7198344 7V1	1401	1841
					7260030H1 (BRAWNOC01)	1889	2499
					7997955H1 (BRAITUC02)	143	745
					6265341H1 (MCLDTXN03)	1	212
					3767715T6 (BRSTNOT24)	613	1253
34	55030459CB1		3958	1-274, 837-1623, 3909-3958	5503021 9H1	1368	1996
					7199252 9V1	2796	3508
					7199098 2V1	2748	3435
					6343107T8 (LUNGDIS03)	1167	1585
					GNN.g7960408_000016_02.edit	1	198
					5503049 1J1	300	785

【表２４】

表4－3

ポリヌクレオチド SEQ ID NO:	Incyte ヌクレオチド ID	ポリ	配列長	選択された断片	配列断片	5'位置	3'位置
					5503045 9H1	809	1368
					7199288 6V1	2086	2783
					7198959 5V1	3248	3958
					7199032 6V1	1875	2578
					5510963 7H1	495	1288
					g4018506	48	528
35	6113039CB1		2000	856-1096	7172164 5V1	1296	2000
					6782480 F9 (SINITMC01)	1	653
					7172271 9V1	411	1050
					7171983 4V1	1000	1641
36	7101781CB1		1997	1568-1778, 644-1026	FL7101781_g7939384_0 0014 g8131858	174	1543
					7092535 6V1	1321	1579
					3748173 F6 (UTRSNOT18)	1	726
					7099050 2V1	1424	1997
37	7473036CB1		3069	1-1362, 2182-2313, 1436-1763	FL7473036_g9255974_0 0002 g2198815	1	2839
					5050192 F6 (BRSTNOT33)	2484	3069
38	7476943CB1		2241	1-168, 1540-2241, 282-745	6392952 F8 (PANCNON03)	1562	2241
					FL7476943_g7739804_0 0008 g3047402	660	1765
					5514001 4J1	1	879
39	8003355CB1		1593	1-38, 1173-1315	8003355 H1 (MUSCTDC01)	28	620
					GNN.g7651721_000004_0 04	49	1593
					3292859 H1 (BONRFET01)	1	248
40	3116448CB1		2121	358-692, 1-22	2378367 F6 (ISLTNOT01)	1258	1771
					5513620 6J2	1	779
					5723184 F6 (SEMVNOT05)	1487	2121
					7076906 1V1	1069	1669
					5513620 6H1	4	858
					7169977 H1 (MCLRNOC01)	771	1092
41	6228668CB1		1225	1-87	7050176 8V1	708	1225
					1851960 F6 (LUNGFET03)	1	529
					7050213 4V1	624	1196

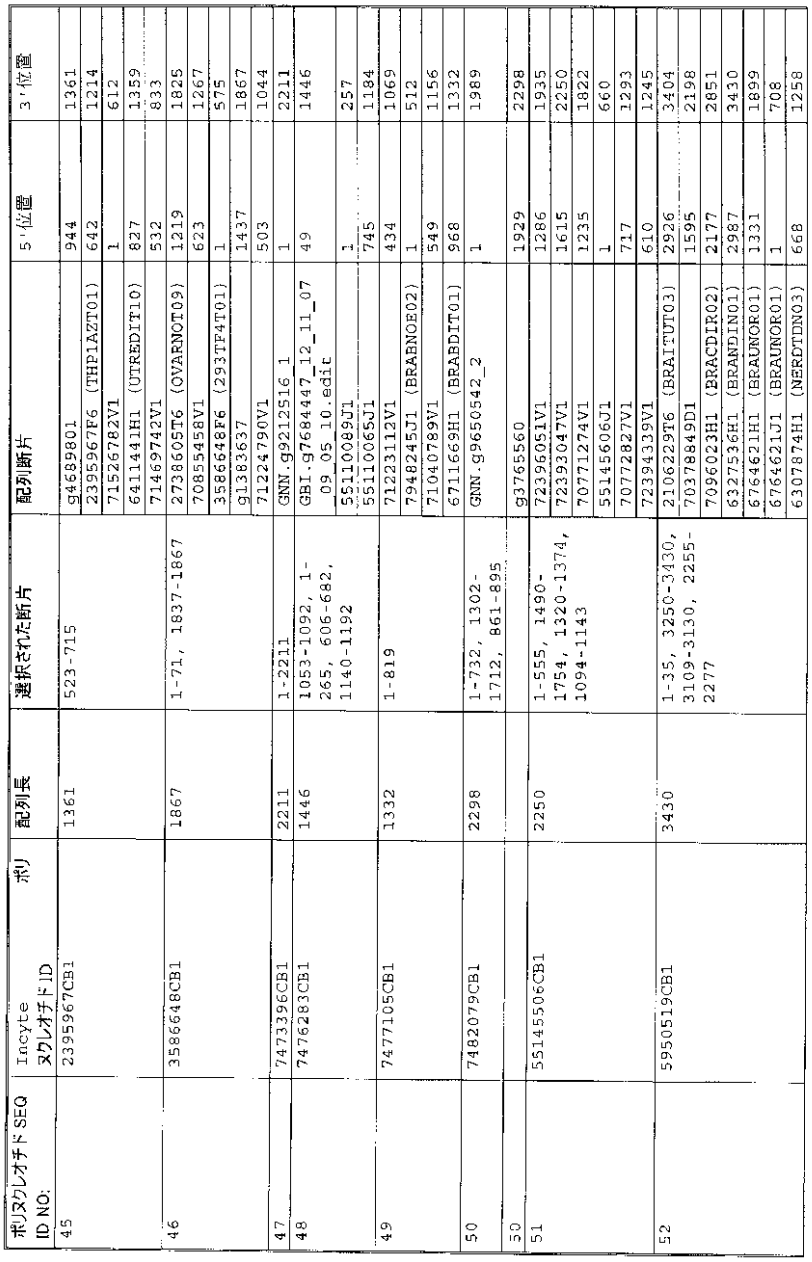
【表２５】

表4－4

ポリヌクレオチドSEQ ID NO:	Incyte ポリヌクレオチドID	ポリ	配列長	選択された断片	配列断片	5'位置	3'位置
					7050118 2V1	470	1114
42	7476494CB1		2693	1-1295, 2361-2451	1382551F6 (BRAITUT08)	1	504
					FL7476494_g9438678_000004_g7688148_1_5-6	2035	2271
					7175426H1 (BRSTTMC01)	1537	2100
					GNN.g9438678_000004_002	1707	2642
					FL7476494_g9438678_000004_g7688148_1_6-7	2147	2435
					5511634TJ1 (SPLNTUB01)	401	1094
					7757711H1 (SPLNTUE01)	762	1236
					FL7476494_g9438678_000004_g7688148_1_8-9	2436	2693
					7757711J1 (SPLNTUB01)	1204	1863
43	7477260CB1		3569	1-2249, 3310-3569	GNN.g8468993_000014_002.edit	3130	3569
43					5546177F8 (TESTNOC01)	2197	3093
					5313313F8 (KIDETXS02)	995	1522
					8011222H1 (NOSEDIC02)	1	767
					5508984 3H1 (PROTDNV21)	668	1007
					7227359H1 (BRAXTDR15)	2847	3383
					5512043 8J1	1497	2198
44	1963058CB1		3920	1-648, 2120-3920	6769623H1 (BRAUNOR01)	1745	2338
					7611869J1 (KIDCTME01)	3090	3825
					7696528H1 (KIDPTDE01)	2104	2636
					7001668H1 (HEALDIR01)	511	1014
					1963058R6 (BRSTNOT04)	3337	3920
					8174904H1 (FETANOA01)	964	1575
					6770724H1 (BRAUNOR01)	1587	2140
					3164103H1 (TLYMTXT04)	1329	1623
					7314310H1 (UTREDME02)	2753	3339
					2659167H1 (LUNGTUT09)	2421	2664
					7727654H1 (UTRCDIE01)	1	560
					7412125H1 (BONMTUB02)	631	1242

【表２６】

表4－5



ポリヌクレオチドSEQ ID NO:	Incyte ポリヌクレオチドID	配列長	選択された断片	配列断片	5'位置	3'位置
45	2395967CB1	1361	523-715	g4689801	944	1361
				2395967F6 (THP1AZT01)	642	1214
				7152678ZV1	1	612
				6411441H1 (UTREDIT10)	827	1359
				7146974ZV1	532	833
46	3586648CB1	1867	1-71, 1837-1867	2738605T6 (OVARNOT09)	1219	1825
				7085545BV1	623	1267
				3586648F6 (293TF4T01)	1	575
				g1383637	1437	1867
				7122479OV1	503	1044
47	7473396CB1	2211	1-2211	GNN.g9212516_1	1	2211
48	7476283CB1	1446	1053-1092, 1-265, 606-682, 1140-1192	GBI.g7684447_12_11_07_09_05_10.edit	49	1446
				5511008QJ1	1	257
				5511006SJ1	745	1184
49	7477105CB1	1332	1-819	7122311ZV1	434	1069
				7948245J1 (BRABNOE02)	1	512
				7104078QV1	549	1156
				6711663H1 (BRABDIT01)	968	1332
50	7482079CB1	2298	1-732, 1302-1712, 861-895	GNN.g9650542_2	1	1989
50				g3765560	1929	2298
51	5514550GCB1	2250	1-555, 1490-1754, 1320-1374, 1094-1143	7239605IV1	1286	1935
				7239304TV1	1615	2250
				7071274V1	1235	1822
				5514560GJ1	1	660
				7077282TV1	717	1293
				7239433QV1	610	1245
52	5950519CB1	3430	1-35, 3250-3430, 3109-3130, 2255-2277	2106229T6 (BRAITUT03)	2926	3404
				7037884QD1	1595	2198
				7096023H1 (BRACDIR02)	2177	2851
				6327536H1 (BRANDIN01)	2987	3430
				6764621H1 (BRAUNOR01)	1331	1899
				6764621J1 (BRAUNOR01)	1	708
				6307874H1 (NERDTDN03)	668	1258

【表２７】

表4－6

ポリヌクレオチドSEQ ID NO:	Incyte ヌクレオチドID	ポリ	配列長	選択された断片	配列断片	5'位置	3'位置
					6980581H1 (BRAHTDR04)	1177	1527
					6121921H1 (BRAKNON05)	2426	2986

【表２８】

表 5

ポリヌクレオチド SEQ ID NO:	Incyte プロジェクト ID	代表的ライブラリ
27	1687189CB1	PROSTMY01
28	7078207CB1	BRAUNOR01
29	1560619CB1	LUNGNOT37
30	2614283CB1	PROSTUT09
31	2667691CB1	STOMFET01
32	3211415CB1	BLADNOT08
33	4739923CB1	BRAITUT03
34	55030459CB1	BRAYDIN03
35	6113039CB1	SINITMC01
36	7101781CB1	LUNGNOT34
37	7473036CB1	BRSTNOT33
38	7476943CB1	PANCNON03
39	8003355CB1	BONRFET01
40	3116448CB1	SEMVNOT05
41	622868CB1	PGANNOT01
42	7476494CB1	SPLNTUB01
43	7477260CB1	TESTNOC01
44	1963058CB1	BRAUNOR01
45	239596 7CB1	THP1AZT01
46	3586648CB1	OVARNOT09
49	7477105CB1	COLNNOT11
51	55145506CB1	SINITMR01
52	5950519CB1	BRAUNOR01

【表２９】

表6－1

ライブラリ	ベクター	ライブラリの説明
BLADNOT08	pINCY	ライブラリは、弾丸による負傷で死亡した 11 才黒人男性の膀胱組織から単離された RNA を用いて作製された。
BONRFET01	pINCY	ライブラリは、20 週間の妊娠中にパトーシンドローム (trisomy 13) で死亡した白人の胎児(男)から採取した肋骨組織から単離された RNA を使用して作製された。
BRAITUT03	PSPORT1	ライブラリは、17 才の白人女性の脳髄膜病変の切除時に、左前頭葉から採取した脳腫瘍組織から単離した RNA を用いて作製した。病理学検査では、グレード 4 の線維性巨大および小星状細胞腫が見られた。家族歴には良性の高血圧および脳血管系疾患がある。
BRAUNOR01	pINCY	このランダムにプライムしたライブラリは、アテローム性動脈硬化症により胸部大動脈破裂による出血で死亡の 81 才白人女性から得た線条体、淡蒼球および被殻組織から単離した RNA を用いて作製した。病理学検査では、双方の内部頸動脈に中程度のアテローム性動脈硬化、前頭皮質および海馬に微視的な梗塞、および散在性広汎性のアミロイド斑、および神経原線維変性が見られたが、これは年齢と予盾しない。肉眼で見て、軟膜は上矢状静脈洞に沿った軽度の硝子質化と厚化が見られるのみであった。残りの軟膜は薄く、充血した血管がいくらか見られた。主に前頭極および前頭葉、及び側頭葉に軽度の萎縮が見られた。微視的には、アルツハイマータイプ II 星状細胞のペアが新皮質の深層にあった。中央前頭皮質の灰白質の奥深くのニューロンの周りでサテリトーシス(satellitosis)が増大していた。扁桃はまれな広汎性斑と神経原線維変性を含んでいた。後部海馬は反応性グリオーシスに囲まれたヘモジデリン含有マクロファージを伴う微視的な嚢胞性の空洞化領域を含んでいた。患者の病歴は、敗血症、胆管炎、術後無気肺、肺炎罹動脈疾患(CAD)、左心室肥大に起因する心肥大、巨脾症、細動脈性腎硬化、結節性コロイド甲状腺腫、気腫、鬱血心不全(CHF)、甲状腺機能低下症および抹消血管疾患を含んでいた。
BRAYDIN03	pINCY	この規格化ライブラリは、脳組織ライブラリの 670 万個の独立性クローンから作製した。開始 RNA は脳血管事故から死亡した 57 才の白人男子から採取された罹病視床下部組織から単離した RNA から作製した。患者の病歴にはハンチントン舞踏病および肺気腫がある。このライブラリは、極めて長時間(48 時間/一回の再アニーリングによるハイブリダイゼーションを用いたことを除いては、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 を改作した条件を用いて 2 回にわたり標準化した。このライブラリはクローンを含む挿入断片用に選択するために直線化し、再度環状にした。
BRSTNOT33	pINCY	ライブラリは 46 才の白人女性の片側性拡大単純乳房切除術時に採取した右乳房組織から単離した RNA を用いて作製した。関連する腫瘍組織の病理はアポクリン様像を有する侵襲度 3 の腺癌、管型、核グレード 3 で外四分円部分に境の形成を示した。50%以上の管内構成成分があった。患者の病歴には乳癌がある。
COLNNOT11	PSPORT1	ライブラリは 60 才の白人男性の左結腸切除時に採取した大腸組織から単離した RNA を用いて作製した。
LUNGNOT34	pINCY	ライブラリは 12 才の白人男子から採取した肺組織から単離した RNA を用いて作製した。
LUNGNOT37	pINCY	ライブラリは閉鎖型頭傷害で死亡した 15 才の白人男子から採取した肺組織から単離した RNA を用いて作製した。血清学は

【表３０】

表6－2

ライブラリ	ベクター	ライブラリの説明
		サイトメガロウイルスにて陽性であった。
OVARNOT09	pINCY	ライブラリは、28才白人女子の腹式子宮摘出術、並びに卵管・卵巣摘出時に摘出された卵巣組織から単離されたRNAを用いて作製された。病理学的検査は、右および左の卵巣内にサイズが0.4から5cmの多数の卵胞嚢胞、慢性子宮頸管炎および子宮頸管の扁平上皮異形成および弱い増殖相にある子宮内膜を示した。家族歴には、良性高血圧症、高脂血症およびアテローム硬化型冠動脈疾患が含まれている。
PANCNOND3	pINCY	規格化ライブラリは、膵臓組織ライブラリの1200万個の独立性クローンから作製した。開始RNAは脳血管事故から死亡した17才の白人女性から採取された膵臓組織から単離したRNAから作製した。血清学はサイトメガロウイルスに陽性であり、残りの血清学検査項目は陰生であった。患者が服用している薬剤はなかった。。このライブラリは、極めて長時間(48時間/一同)の再アニーリングによるハイブリダイゼーションを用いたことを除いては、Soares他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 を適応した条件を用いて2回にわたりノーマライズした。
PGANNOT01	PSPORT1	ライブラリは46才の白人男性の診査開腹時に採取した内腹部から採取した傍神経節から単離したRNAを用いて作製した。病理学検査では良性の傍神経節腫が見られ、これはグレード2の腎臓細胞癌（褪明細胞タイプ、被膜を浸透していない）と関連していた。外科的縁では腫瘍について陰性であった。
PROSTMY01	pINCY	この大型分画cDNAおよび規格化ライブラリは、55才の白人男性から閉鎖型前立腺生検、前立腺全摘出術および局所リンパ節切除時に摘出した病変前立腺組織から単離したRNAを用いて作製した。病理学検査では腺線維腫増殖症が見られた。一致する腫瘍組織の病理検査によると、腺癌（Gleason グレード4）が左側周辺に主要な塊を形成し、右後ろ上皮膜に拡張していた。腫瘍は被膜を侵襲し、被膜を突き抜けて左後ろ上腺域の前立腺周辺組織に影響していた。左下後ろおよび左上後ろの外科的縁は陽性だった。左骨盤のリンパ腺の一つが転移の影響を受けていた。患者の病歴には、腎臓の結石が見られた。家族の病歴としては、肺癌および乳癌が含まれていた。このサイズ選択されたライブラリは、ーラウンドで規格化された。条件は Soares 他. PNAS (1994) 91:9228-9232 および Bonaldo 他. Genome Research (1996) 6:791 を修正して用いた。
PROSTUT09	pINCY	ライブラリは、根治的前立腺除去、根治的摘脱切除、及び尿道回中の66才白人男性から取り除かれた前立腺腫瘍組織から単離された RNA を用いて作製された。病理学検査では、グレード3の転移性細胞癌を示した。患者は前立腺炎症疾患があった。患者の病歴には肺腫瘍と良性の高血圧症がある。家族歴には、悪性乳房腫瘍、結核、脳血管疾患、アテローム硬化型冠状動脈疾患、および肺癌が含まれる。
SEMVNOT05	pINCY	ライブラリは67才の白人男性の根本的前立腺切除時に採取した精嚢組織から単離したRNAを用いて作製した。関連する腫瘍組織の病理学検査では、腺癌(Gleasonグレード3+3)が見られた。

【表３１】

表6－3

ライブラリ	ベクター	ライブラリの説明
SINITMC01	pINCY	この大型分画ライブラリは二人の供与者からのプールされたcDNAを用いて作製した。cDNAは30才白人女性(供与者 A)の部分結腸切除術、切開肝生検、および永久的結腸瘻造設術時切除した回腸組織および70才白人女性(供与者 B)の右半結腸切除、切開肝生検、S状結腸鏡検査、結腸鏡検査および永久的結腸瘻造設術時に摘出された回腸組織から単離されたmRNAを用いて作製した。一致する腫瘍組織の病理(供与者 A)は回腸末端部に生じたカルチノイド腫瘍(グレード1の神経内分泌腺癌)を示した。この腫瘍は回腸壁から腸間膜脂肪を貫通して、接着した盲腸にまで広がっており、そこで腫瘍は腸壁を貫通して粘膜表面まで広がっている。複数のリンパ節が腫瘍に陽性であった。さらに2個のリンパ節にもまた直接の腫瘍拡大が見られた。供与者 B の病理は回腸の縁は非腫瘍性であった。一致する腫瘍の病理(供与者 B)は侵襲度2の腺癌、回盲弁近位部位に潰瘍性境の形成を示した。腫瘍の侵襲は固有筋を通過して漿膜脂肪組織にまで及んでいた。一つの限局性リンパ節は転移性腺癌のマイクロフォーカスに陽性であった。供与者 A は顔面紅潮を示し、非特定の上腹部/背痙症状を示した。患者の病歴は子宮内膜症とタバコ中毒、アルコール中毒があった。供与者 B の病歴には、悪性乳房腫瘍、II型糖尿病、高脂血症、ウイルス性肝炎、非特定甲状腺疾患、骨関節炎、悪性皮膚腫瘍が含まれている。供与者 B の服用薬剤はタモキシフェンである。
SINITMF01	PCDNA2.1	このランダムプライムされたライブラリは70才白人女性の右半結腸切除、閉腹式肝生検、S状結腸鏡検査、結腸鏡検査および永久的結腸瘻造設術時に摘出された回腸組織から単離された mRNA を用いて作製した。一致する腫瘍の病理は侵襲度2の腺癌、回盲弁の2cm遠位部位に潰瘍性塊の形成を示した。患者の病歴には、悪性乳房腫瘍、II型糖尿病、高脂血症、ウイルス性肝炎、非特定甲状腺疾患、骨関節炎、悪性皮膚腫瘍、欠乏性貧血、および正常分娩が含まれている。家族の病歴には、乳癌、アテローム硬化性冠動脈疾病、良性の高血圧、脳血管疾患、卵巣癌、および高脂質血症が見られた。
SPLNTUB01	PCDNA2.1	この5'バイアスのランダムプライムされたライブラリは 28 歳白人男性の脾臓全摘出術時に切除された腎臓腫瘍組織から単離した RNA を用いて作製した。病理は悪性リンパ腫、びまん性大細胞タイプ、B細胞表現型を示し、ほぼ45個の小結節形成がある脾臓、肝臓および多数のリンパ節にわたって多くの反応性T細胞および著しい肉芽腫反応が示された。
STOMFET01	pINCY	ライブラリは妊娠 20 週で死亡の白人女性胎児から採取した胃組織から単離した RNA を用いて作製された。
TESTNOC01	PBLUESCRIPT	この大型分画ライブラリは10～61歳の11人の白人男性から除去された精巣から単離された RNA を用いて作製された。
THP1AZT01	pINCY	ライブラリは、0．8マイクロモルの5アザ2'デオキシシチジンで3日間処理された THP-1プロモノサイト細胞から単離された RNA を用いて作製された。THP-1 (ATCC TIB 202)は急性の単球性白血病の1歳の白人男子の抹消血から得られたヒトの前単球系である (Int. J. Cancer (1980) 26:171)。

【表３２】

表7－1

プログラム	説明	引用文献	パラメータ閾値
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、注釈付けに有用である。	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 はアミノ酸およびヌクレオチド配列類似性検索に有用であり、BLASTには5つの機能がある：blastp、blastn、blastx、tblastnおよびtblastx。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) (1997) Nucleic Acids Res. 25:3389-3402.	*ESTs*:確率値=1.0E-8以下 完全長配列:確率値= 1.0E-10以下
FASTA	問合せ配列と同様の配列群配列との類似性を検索するPearsonおよびLipman アルゴリズム。FASTAは最低5つの機能からなる：fasta、tfasta、fastx、tfastxおよびssearch。	Pearson, W.R.及びD.J. Lipman (1988) Proc. Natl. Acad Sci.USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol.183:63-98; およびSmith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs*:fasta E 値=1.06E-6 *構築された ESTs*:fasta 同一性=95%以上 一致した長さ=200塩基以上 fastx E 値=1.0E-8 以下 *完全長配列*: fastx スコア=100以上
BLIMPS	配列をBLOCKS、PRINTS、DOMO、PRODOMおよびPFAMデータベースの配列と対応させて遺伝子ファミリ、配列相同性および構造的フィンガープリント領域を検索するBLocks iMProved サーチャー。	Henikoff, S.及びJ.G. Henikoff (1991) Nucleic Acids Res.19:6565-6572; Henikoff, J.G.及びS. Henikoff (1996) Methods Enzymol.266:88-105, Attwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	確率値= 1.0E-3以下
HMMER	PFAMのようなタンパク質ファミリ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 (1988) Nucleic Acids Res.26:320-322; Durbin, R. 他。(1998) Our World View, in a Nutshell. Cambridge Univ.Press. 1-350ページ。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３３】

表7－2

プログラム	説明	引用文献	パラメータ閾値
ProfileScan	Prositeで定義された配列パターンと一致するタンパク質配列の構造的モチーフおよび配列モチーフを検索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4:61-66; Gribskov, M. 他 (1989) Methods Enzymol.183:146-159; Bairoch, A. 他. (1997) Nucleic Acids Res. 25:217-221.	標準化された質のスコアとその特定のProsite モチーフに対するGCG指定" HIGH" 値通常、スコア=1.4-2.1.
Phred	高い感度と信頼で自動配列決定機のトレースを調べる塩基誘出しアルゴリズム。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B.及び 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、配列相同性の検索やDNA配列の構築に有用である。	Smith, T.F. および M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. および M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	スコア=120以上一致した長さ=56以上
Consed	Phrap で構築したもの表示、編集用グラフィカルツール。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキャンして、分泌シグナルペプチドの存在を調べる重み付けマトリックス解析プログラム。	Nielson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J.M. 及び S. Audic (1997) CABIOS 12:431-439.	スコア=3.5以上
TMAP	加重マトリックスを用いて蛋白配列での膜貫通セグメントの描写および配向を決定するプログラム。	Persson, B. およびP. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. およびP. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	隠れMarkov モデル(HMM) を使って蛋白配列の膜貫通セグメントを描写し、配向を決定するプログラム。	Sonnhammer, E.L. 他 (1998) Proc.Sixth Intl. Conf. on Intelligent Systems for Mol.Biol., Glasgow 他, eds., The Am.Assoc. for Artificial Intelligence Press, Menlo Park, CA. 175-182ページ。175-182.	
Motifs	Prositeで定義された配列と一致したパターンのアミノ酸配列を検索するプログラム。	Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, 第9版、M51-59頁, Genetics Computer Group, Madison, WI	

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 March 2002 (21.03.2002)

PCT

(10) International Publication Number
WO 02/22684 A2

(51) International Patent Classification⁷: C07K 14/705

(21) International Application Number: PCT/US01/28938

(22) International Filing Date:
14 September 2001 (14.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/232,685	15 September 2000 (15.09.2000)	US
60/234,842	22 September 2000 (22.09.2000)	US
60/236,882	29 September 2000 (29.09.2000)	US
60/239,057	5 October 2000 (05.10.2000)	US
60/240,540	13 October 2000 (13.10.2000)	US
60/241,700	18 October 2000 (18.10.2000)	US

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

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LEE, Ernestine, A.** [US/US]; 624 Kains Street, Albany, CA 94706 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **LAL, Preeti, G.** [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **WALIA, Narinder, K.** [US/US]; 890 Davis Street #205, San Leandro, CA 94577 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **WARREN, Bridget, A.** [US/US]; 10130 Parkwood Drive #2, Cupertino, CA 95014 (US). **LEE, Sally** [US/US]; 825 East Evelyn #425, Sunnyvale, CA 94086 (US). **SANJANWALA, Madhu, S.** [US/US]; 210 Sylvia Court, Los Altos, CA 94024 (US). **YAO, Monique, G.** [US/US]; 1189 Woodgate Drive, Carmel, IN 46033 (US). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **THORNTON, Michael** [US/US]; 9 Medway Road, Woodside, CA 94062-2612 (US). **GANDHI, Ameena, R.** [US/US]; 837 Roble Avenue #1, Menlo Park, CA 94025 (US). **POLICKY, Jennifer, L.** [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **ARVIZU, Chandra** [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). **RAUMANN, Brigitte, E.** [US/US]; 5801 South Dorchester Avenue #3B, Chicago, IL 60637 (US). **BRUNS, Christopher, M.** [US/US]; 575 S. Rengstorff Avenue #126, Mountain View, CA 94040 (US). **NAINI, Amir** [US/US]; 51 Spyglass Hill, Oakland, CA 94618 (US). **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **NGUYEN, Danniel, B.** [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). **LU, Dyung, Aina, M.** [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). **ISON, Craig, H.** [US/US]; 1242 Weathersfield Way, San Jose, CA 95118 (US). **GRIFFIN, Jennifer, A.** [US/US]; 33691 Mello Way, Fremont, CA 94555 (US). **REDDY, Roopa, M.** [IN/US]; 1233 W. McKinley Avenue #3, Sunnyvale, CA 94086 (US). **BURFORD, Neil** [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US).

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

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

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

Published:
— *without international search report and to be republished upon receipt of that 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 02/22684 A2

(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.

TRANSPORTERS AND ION CHANNELS

TECHNICAL FIELD

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

BACKGROUND OF THE INVENTION

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

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

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

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 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 Martinelle, K. and I. Haggstrom (1993) J. Biotechnol. 30:339-350).

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

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

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

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

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

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

Ion Channels

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

Ion Transporters

Ion transporters generate and maintain the resting electrical potential of a cell. Utilizing the

energy derived from ATP hydrolysis, they transport ions against the ion's concentration gradient. These transmembrane ATPases are divided into three families. The phosphorylated (P) class ion transporters, including Na^+-K^+ ATPase, Ca^{2+}-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 Ca^{2+} are low and cytosolic concentration of K^+ is high. The vacuolar (V) class of ion transporters includes H^+ pumps on intracellular organelles, such as lysosomes and Golgi. V-class ion transporters are responsible for generating the low pH within the lumen of these organelles that is required for function. The coupling factor (F) class consists of H^+ pumps in the mitochondria. F-class ion transporters utilize a proton gradient to generate ATP from ADP and inorganic phosphate (P_i).

The P-ATPases are hexamers of a 100 kD subunit with ten transmembrane domains and several large cytoplasmic regions that may play a role in ion binding (Scarborough, G.A. (1999) Curr. Opin. Cell Biol. 11:517-522). The V-ATPases are composed of two functional domains: the V_1 domain, a peripheral complex responsible for ATP hydrolysis; and the V_0 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 F_0 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 V_0 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 Ca^{2+} out of the cell with transport of Na^+ into the cell (antiport).

Gated Ion Channels

Gated ion channels control ion flow by regulating the opening and closing of pores. The ability to control ion flux through various gating mechanisms allows ion channels to mediate such diverse signaling and homeostatic functions as neuronal and endocrine signaling, muscle contraction, fertilization, and regulation of ion and pH balance. Gated ion channels are categorized according to the manner of regulating the gating function. Mechanically-gated channels open their pores in

response to mechanical stress; voltage-gated channels (e.g., Na^+, K^+, Ca^{2+}, and Cl^- channels) open their pores in response to changes in membrane potential; and ligand-gated channels (e.g., acetylcholine-, serotonin-, and glutamate-gated cation channels, and GABA- and glycine-gated chloride channels) open their pores in the presence of a specific ion, nucleotide, or neurotransmitter. The gating properties of a particular ion channel (i.e., its threshold for and duration of opening and closing) are sometimes modulated by association with auxiliary channel proteins and/or post translational modifications, such as phosphorylation.

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

The pore-forming subunits of the voltage-gated cation channels form a superfamily of ion channel proteins. The characteristic domain of these channel proteins comprises six transmembrane domains (S1-S6), a pore-forming region (P) located between S5 and S6, and intracellular amino and carboxy termini. In the Na^+ and Ca^{2+} subfamilies, this domain is repeated four times, while in the K^+ channel subfamily, each channel is formed from a tetramer of either identical or dissimilar subunits. The P region contains information specifying the ion selectivity for the channel. In the case of K^+ channels, a GYG tripeptide is involved in this selectivity (Ishii, T.M. et al. (1997) Proc. Natl. Acad. Sci. USA 94:11651-11656).

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

Voltage-gated Na^+ channels are heterotrimeric complexes composed of a 260 kDa pore-forming α subunit that associates with two smaller auxiliary subunits, β1 and β2. The β2 subunit is a integral membrane glycoprotein that contains an extracellular Ig domain, and its association with α and β1 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; Eglen, R.M. et al. (1999) Trends Pharmacol. Sci. 20:337-342).

K^+ channels are located in all cell types, and may be regulated by voltage, ATP concentration, or second messengers such as Ca^{2+} and cAMP. In non-excitable tissue, K^+ channels are involved in protein synthesis, control of endocrine secretions, and the maintenance of osmotic equilibrium across membranes. In neurons and other excitable cells, in addition to regulating action potentials and repolarizing membranes, K^+ channels are responsible for setting resting membrane potential. The cytosol contains non-diffusible anions and, to balance this net negative charge, the cell contains a Na^+-K^+ pump and ion channels that provide the redistribution of Na^+, K^+, and Cl^-. The pump actively transports Na^+ out of the cell and K^+ into the cell in a 3:2 ratio. Ion channels in the plasma membrane allow K^+ and Cl^- to flow by passive diffusion. Because of the high negative charge within the cytosol, Cl^- flows out of the cell. The flow of K^+ is balanced by an electromotive force pulling K^+ into the cell, and a K^+ concentration gradient pushing K^+ out of the cell. Thus, the resting membrane potential is primarily regulated by K^+ flow (Salkoff, L. and T. Jegla (1995) Neuron 15:489-492).

Potassium channel subunits of the Shaker-like superfamily all have the characteristic six transmembrane/1 pore domain structure. Four subunits combine as homo- or heterotetramers to form

functional K channels. These pore-forming subunits also associate with various cytoplasmic β subunits that alter channel inactivation kinetics. The Shaker-like channel family includes the voltage-gated K^+ channels as well as the delayed rectifier type channels such as the human ether-a-go-go related gene (HERG) associated with long QT, a cardiac dysrythmia syndrome (Curran, M.E. (1998) Curr. Opin. Biotechnol. 9:565-572; Kaczorowski, G.J. and M.L. Garcia (1999) Curr. Opin. Chem. Biol. 3:448-458).

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

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

The voltage-gated Ca^{2+} channels have been classified into several subtypes based upon their electrophysiological and pharmacological characteristics. L-type Ca^{2+} channels are predominantly expressed in heart and skeletal muscle where they play an essential role in excitation-contraction coupling. T-type channels are important for cardiac pacemaker activity, while N-type and P/Q-type channels are involved in the control of neurotransmitter release in the central and peripheral nervous system. The L-type and N-type voltage-gated Ca^{2+} 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. These subunits are encoded by at least six α_1, one $\alpha_2\delta$, and four β genes. A fourth subunit, γ, has been identified in skeletal muscle (Walker, D. et al. (1998) J. Biol. Chem. 273:2361-2367; McCleskey, E.W. (1994) Curr. Opin. Neurobiol. 4:304-312).

The transient receptor family (Trp) of calcium ion channels are thought to mediate capacitative calcium entry (CCE). CCE is the Ca^{2+} influx into cells to resupply Ca^{2+} stores depleted by the action of inositol triphosphate (IP3) and other agents in response to numerous hormones and growth factors. Trp and Trp-like were first cloned from Drosophila and have similarity to voltage

gated Ca2+ channels in the S3 through S6 regions. This suggests that Trp and/or related proteins may form mammalian CCC entry channels (Zhu, X. et al. (1996) Cell 85:661-671; Boulay, G. et al. (1997) J. Biol. Chem. 272:29672-29680). Melastatin is a gene isolated in both the mouse and human, and whose expression in melanoma cells is inversely correlated with melanoma aggressiveness in vivo. The human cDNA transcript corresponds to a 1533-amino acid protein having homology to members of the Trp family. It has been proposed that the combined use of malastatin mRNA expression status and tumor thickness might allow for the determination of subgroups of patients at both low and high risk for developing metastatic disease (Duncan, L.M. et al (2001) J. Clin. Oncol. 19:568-576).

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

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

Ligand-gated channels open their pores when an extracellular or intracellular mediator binds to the channel. Neurotransmitter-gated channels are channels that open when a neurotransmitter binds to their extracellular domain. These channels exist in the postsynaptic membrane of nerve or muscle cells. There are two types of neurotransmitter-gated channels. Sodium channels open in response to excitatory neurotransmitters, such as acetylcholine, glutamate, and serotonin. This opening causes an influx of Na^+ and produces the initial localized depolarization that activates the voltage-gated channels and starts the action potential. Chloride channels open in response to inhibitory neurotransmitters,

such as γ-aminobutyric acid (GABA) and glycine, leading to hyperpolarization of the membrane and the subsequent generation of an action potential. Neurotransmitter-gated ion channels have four transmembrane domains and probably function as pentamers (Jentsch, supra). Amino acids in the second transmembrane domain appear to be important in determining channel permeation and selectivity (Sather, W.A. et al. (1994) Curr. Opin. Neurobiol. 4:313-323).

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

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

The activity of other types of ion channel proteins may also be modulated by a variety of intracellular signalling proteins. Many channels have sites for phosphorylation by one or more protein kinases including protein kinase A, protein kinase C, tyrosine kinase, and casein kinase II, all of which regulate ion channel activity in cells. Kir channels are activated by the binding of the Gβγ 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.E. and D.S. Bredt (1998) Cell 93:495-498).

Disease Correlation

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

Human diseases caused by mutations in ion channel genes include disorders of skeletal muscle, cardiac muscle, and the central nervous system. Mutations in the pore-forming subunits of sodium and chloride channels cause myotonia, a muscle disorder in which relaxation after voluntary contraction is delayed. Sodium channel myotonias have been treated with channel blockers. Mutations in muscle sodium and calcium channels cause forms of periodic paralysis, while mutations in the sarcoplasmic calcium release channel, T-tubule calcium channel, and muscle sodium channel cause malignant hyperthermia. Cardiac arrythmia 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).

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

Ion channels in the immune system have recently been suggested as targets for

immunomodulation. T-cell activation depends upon calcium signaling, and a diverse set of T-cell specific ion channels has been characterized that affect this signaling process. Channel blocking agents can inhibit secretion of lymphokines, cell proliferation, and killing of target cells. A peptide antagonist of the T-cell potassium channel Kv1.3 was found to suppress delayed-type hypersensitivity and allogenic responses in pigs, validating the idea of channel blockers as safe and efficacious immunosuppressants (Cahalan, M.D. and K.G. Chandy (1997) Curr. Opin. Biotechnol. 8:749-756).

The discovery of new transporters and ion channels, and the polynucleotides encoding them, satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of transport, neurological, muscle, immunological, and cell proliferative disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of transporters and ion channels.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, transporters and ion channels, referred to collectively as "TRICH" and individually as "TRICH-1," "TRICH-2," "TRICH-3," "TRICH-4," "TRICH-5," "TRICH-6," "TRICH-7," "TRICH-8," "TRICH-9," "TRICH-10," "TRICH-11," "TRICH-12," "TRICH-13," "TRICH-14," "TRICH-15," "TRICH-16," "TRICH-17," "TRICH-18," "TRICH-19," "TRICH-20," "TRICH-21," "TRICH-22," "TRICH-23," "TRICH-24," "TRICH-25," and "TRICH-26." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-26.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-26. In

another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:27-52.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to

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

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. The invention additionally

provides a method of treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional TRICH, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an

immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, ii) a

polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

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

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

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

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

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

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

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

DESCRIPTION OF THE INVENTION

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

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

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

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

"Altered" nucleic acid sequences encoding TRICH include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as TRICH or a polypeptide with at least one functional characteristic of TRICH. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding TRICH, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding TRICH.

The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent TRICH. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of TRICH is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. 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 TRICH. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of TRICH either by directly interacting with TRICH or by acting on components of the biological pathway in which TRICH participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind TRICH polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies

which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an in vitro evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH_2), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker. (See, e.g., Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13.)

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

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

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

The term "biologically active" refers to a protein having structural, regulatory, or biochemical

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

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

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

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

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

Original Residue	Conservative Substitution
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

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

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

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

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

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

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

A "fragment" is a unique portion of TRICH or the polynucleotide encoding TRICH which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment

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

A fragment of SEQ ID NO:27-52 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:27-52, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:27-52 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:27-52 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:27-52 and the region of SEQ ID NO:27-52 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:1-26 is encoded by a fragment of SEQ ID NO:27-52. A fragment of SEQ ID NO:1-26 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-26. For example, a fragment of SEQ ID NO:1-26 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-26. The precise length of a fragment of SEQ ID NO:1-26 and the region of SEQ ID NO:1-26 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

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

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

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

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

sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

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

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

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

length over which percentage identity may be measured.

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

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

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

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

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

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

used to describe a length over which percentage identity may be measured.

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

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

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

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

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

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

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

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

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

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

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

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

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

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

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding

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

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

"Post-translational modification" of an TRICH may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of TRICH.

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

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

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

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

purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

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

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

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions

(UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

THE INVENTION

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

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

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

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

polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are transporters and ion channels. For example, SEQ ID NO:2 is 94% identical from amino acids 965 through 2436 to mouse abc2 transporter (GenBank ID g495259) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:2 also contains two ABC transporter domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:3 is an ABC transporter. In an alternate example, SEQ ID NO:13 is 97% identical to human gamma subunit precursor of muscle acetylcholine receptor (GenBank ID g825618) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 3.0e-273, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:13 also contains a neurotransmitter-gated ion-channel domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:13 is a neurotransmitter-gated ion-channel protein. In an alternate example, SEQ ID NO:19 is 62% identical to human vacuolar proton-ATPase (GenBank ID g37643) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 3.2e-129, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. Data from BLAST analyses provide further corroborative evidence that SEQ ID NO:19 is a vacuolar ATP synthase. In an alternate exampple, SEQ ID NO:22 is 94% identical to rat GABA(A) receptor gamma-1 subunit (GenBank ID g56176) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 4.4e-244, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:22 also contains a neurotransmitter-gated ion channel domain as determined by searching for statistically significant

matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:22 is a neurotransmitter-gated ion channel. In an alternate example, SEQ ID NO:26 is 61% identical to rabbit peroxisomal Ca-dependent solute carrier (GenBank ID g2352427) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 6.4e-156, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:26 also contains three mitochondrial carrier protein domains, as well as three EF hand domains, as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:26 is a calcium dependent carrier protein. In an alternate examplpe, SEQ ID NO:17 is 69% identical to *Ambystoma tigrinum* electrogenic $NaHCO_3$ cotransporter (GenBank ID g2198815) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:17 also contains an HCO_3 transporter family domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:17 is an anion transporter. SEQ ID NO:1, SEQ ID NO:3-12, SEQ ID NO:14-16, SEQ ID NO:18, and SEQ ID NO:20-25 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-26 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:27-52 or that distinguish between SEQ ID NO:27-52 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5')

and stop (3') positions of the cDNA and/or genomic sequences in column 5 relative to their respective full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 7251266F7 is the identification number of an Incyte cDNA sequence, and PROSTMY01 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., 70564238V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g4689801) which contributed to the assembly of the full length polynucleotide sequences. In addition, the identification numbers in column 5 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the identification numbers in column 5 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, FL_*XXXXXX*_N_1_N_2_*YYYYY*_N_3_N_4 represents a "stitched" sequence in which *XXXXXX* is the identification number of the cluster of sequences to which the algorithm was applied, and *YYYYY* is the number of the prediction generated by the algorithm, and $N_{1,2,3...}$, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the identification numbers in column 5 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, FL*XXXXXX*_g*AAAAA*_g*BBBBB*_1_*N* is the identification number of a "stretched" sequence, with *XXXXXX* being the Incyte project identification number, g*AAAAA* being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, g*BBBBB* being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and *N* referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, g*BBBBB*).

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

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

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

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

The invention also encompasses TRICH variants. A preferred TRICH variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the TRICH amino acid sequence, and which contains at least one functional or structural characteristic of TRICH.

The invention also encompasses polynucleotides which encode TRICH. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:27-52, which encodes TRICH. The polynucleotide sequences of SEQ ID NO:27-52, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding TRICH. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding TRICH. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:27-

52 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:27-52. 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.

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

The invention also encompasses production of DNA sequences which encode TRICH and TRICH derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding TRICH or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:27-52 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment

of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding 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.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a 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, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

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

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

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

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 (Applied Biosystems). Additionally, the amino acid sequence of TRICH, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

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

In order to express a biologically active TRICH, the nucleotide sequences encoding TRICH or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding TRICH. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding TRICH. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding TRICH and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided

by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding TRICH and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

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., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending 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, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of TRICH are needed, e.g. for the production of antibodies, vectors which direct high level expression of TRICH may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of TRICH. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

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

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding TRICH may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses TRICH in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers,

or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

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

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *apr* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding TRICH is inserted within a marker gene sequence, transformed cells containing sequences encoding TRICH can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding TRICH under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding TRICH and that express TRICH may be identified by a variety of procedures known to those of skill in the art. These

procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of TRICH using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on TRICH is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

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.

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 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 WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding TRICH may be ligated to a heterologous sequence resulting in translation of a 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 immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the TRICH encoding sequence and the heterologous protein sequence, so that TRICH may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

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

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

In one embodiment, the compound thus identified is closely related to the natural ligand of TRICH, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a

natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which TRICH binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express TRICH, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing TRICH or cell membrane fractions which contain TRICH are then contacted with a test compound and binding, stimulation, or inhibition of activity of either TRICH or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with TRICH, either in solution or affixed to a solid support, and detecting the binding of TRICH to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

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

In another embodiment, polynucleotides encoding TRICH or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science

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

Polynucleotides encoding TRICH may also be manipulated in vitro in ES cells derived from 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 (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding TRICH is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress TRICH, e.g., by secreting TRICH in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of TRICH and transporters and ion channels. In addition, the expression of TRICH is closely associated with brain, lung, prostate, bladder, bone, hypothalamus, breast, ileum, stomach, pancreas, and gastrointestinal tissues and tumors of the brain and prostrate. Therefore, TRICH appears to play a role in transport, neurological, muscle, immunological, and cell proliferative disorders. In the treatment of disorders associated with increased TRICH expression or activity, it is desirable to decrease the expression or activity of TRICH. In the treatment of disorders associated with decreased TRICH expression or activity, it is desirable to increase the expression or activity of TRICH.

Therefore, in one embodiment, TRICH or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH. Examples of such disorders include, but are not limited to, a transport disorder

such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer, cardiac disorders associated with transport, e.g., angina, bradyarrythmia, tachyarrythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis, neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia, and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, 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 disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a muscle disorder

such as cardiomyopathy, myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, ethanol myopathy, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, pheochromocytoma, and myopathies including encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, ophthalmoplegia, and acid maltase deficiency (AMD, also known as Pompe's disease); an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus; systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing TRICH or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH including, but not limited to, those described above.

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

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

In a further embodiment, an antagonist of TRICH may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TRICH. Examples of such disorders include, but are not limited to, those transport, neurological, muscle, immunological, and cell proliferative disorders described above. In one aspect, an antibody which specifically binds TRICH may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express TRICH.

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

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

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

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

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to TRICH have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of TRICH amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

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

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. 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 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 retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) Blood 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87(11):1308-1315; and Morris, M.C. et al. (1997) Nucleic Acids Res. 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding TRICH may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in TRICH expression or regulation causes disease, the expression of TRICH from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in TRICH are treated by constructing mammalian expression vectors encoding TRICH and introducing these vectors by mechanical means into TRICH-deficient cells. Mechanical transfer technologies for

use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Récipon (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, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). TRICH may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and 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. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for

receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent No. 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver 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 described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding TRICH to target cells which have one or more genetic abnormalities with respect to the expression of TRICH. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing TRICH to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and

ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu, H. et al. (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding TRICH to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for TRICH into the alphavirus genome in place of the capsid-coding region results in the production of a large number of TRICH-coding RNAs and the synthesis of high levels of TRICH in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy 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 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 library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding TRICH is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding TRICH are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding TRICH. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved

using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

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

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

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

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

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

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

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

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

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

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

DIAGNOSTICS

In another embodiment, antibodies which specifically bind TRICH may be used for the diagnosis of disorders characterized by expression of TRICH, or in assays to monitor patients being treated with TRICH or agonists, antagonists, or inhibitors of TRICH. Antibodies useful for diagnostic

purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for TRICH include methods which utilize the antibody and a label to detect TRICH in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

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

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

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

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the TRICH encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:27-52 or from genomic sequences including promoters, enhancers, and introns of the TRICH gene.

Means for producing specific hybridization probes for DNAs encoding TRICH include the cloning of polynucleotide sequences encoding TRICH or TRICH derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA

polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding TRICH may be used for the diagnosis of disorders associated with expression of TRICH. Examples of such disorders include, but are not limited to, a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer, cardiac disorders associated with transport, e.g., angina, bradyarrythmia, tachyarrythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis, neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia, and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, 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 disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular

disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a muscle disorder such as cardiomyopathy, myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, ethanol myopathy, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, pheochromocytoma, and myopathies including encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, ophthalmoplegia, and acid maltase deficiency (AMD, also known as Pompe's disease); an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding TRICH may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in

dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered TRICH expression. Such qualitative or quantitative methods are well known in the art.

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

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

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

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

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

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding TRICH may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding TRICH are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of 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; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray

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

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

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

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

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a

signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at http://www.niehs.nih.gov/oc/news/toxchip.htm.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

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

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome 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; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

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

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

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared

with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding TRICH may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

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

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery

techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

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

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

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

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

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 and including U.S. Ser. No. 60/232,685, U.S. Ser. No. 60/234,842, U.S. Ser. No. 60/236,882, U.S. Ser.

No. 60/239,057, U.S. Ser. No. 60/240,540, and U.S. Ser. No. 60/241,700 are expressly incorporated by reference herein.

EXAMPLES

I. Construction of cDNA Libraries

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

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

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA (Invitrogen), PCMV-ICIS (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

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

databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

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

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:27-52. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative transporters and ion channels were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is

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

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

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

along their parent sequences to generate the longest possible sequence, as well as sequence variants. 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 gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

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

VI. Chromosomal Mapping of TRICH Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:27-52 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:27-52 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in

humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

In this manner, SEQ ID NO:31 was mapped to chromosome 1 within the interval from 133.00 to 137.30 centiMorgans. SEQ ID NO:33 was mapped to chromosome 12 within the interval from 120.50 to the q terminal, or more specifically, within the interval from 126.10 to 145.70 centiMorgans.

VII. Analysis of Polynucleotide Expression

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

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

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

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the

other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

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

VIII. Extension of TRICH Encoding Polynucleotides

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

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

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min;

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

The concentration of DNA in each well was determined by dispensing 100 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 CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

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

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

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:27-52 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

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

X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra.), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the

biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

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

Microarray Preparation

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

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

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

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

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

Hybridization

Hybridization reactions contain 9 μl of sample mixture consisting of 0.2 μg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65° C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 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.

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).

XI. Complementary Polynucleotides

Sequences complementary to the TRICH-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring TRICH. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of TRICH. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the TRICH-encoding transcript.

XII. Expression of TRICH

Expression and purification of TRICH is achieved using bacterial or virus-based expression systems. For expression of TRICH in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA

transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express TRICH upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of TRICH in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding TRICH by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

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

XIII. Functional Assays

TRICH function is assessed by expressing the sequences encoding TRICH at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg 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 μg of an additional plasmid containing sequences encoding a

marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

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

XIV. Production of TRICH Specific Antibodies

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

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

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to

increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-TRICH activity by, for example, binding the peptide or TRICH to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring TRICH Using Specific Antibodies

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

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

XVI. Identification of Molecules Which Interact with TRICH

Molecules which interact with TRICH may include transporter substrates, agonists or antagonists, modulatory proteins such as Gβγ 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 ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled TRICH, washed, and any wells with labeled TRICH complex are assayed. Data obtained using different concentrations of TRICH are used to calculate values for the number, affinity, and association of TRICH with the candidate molecules.

Alternatively, proteins that interact with TRICH are isolated using the yeast 2-hybrid system (Fields, S. and O. Song (1989) Nature 340:245-246). TRICH, or fragments thereof, are expressed as fusion proteins with the DNA binding domain of Gal4 or lexA, and potential interacting proteins are expressed as fusion proteins with an activation domain. Interactions between the TRICH fusion protein and the TRICH interacting proteins (fusion proteins with an activation domain) reconstitute a transactivation function that is observed by expression of a reporter gene. Yeast 2-hybrid systems are commercially available, and methods for use of the yeast 2-hybrid system with ion channel proteins are discussed in Niethammer, M. and M. Sheng (1998, Meth. Enzymol. 293:104-122).

TRICH may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Potential TRICH agonists or antagonists may be tested for activation or inhibition of TRICH ion channel activity using the assays described in section XVIII.

XVII. Demonstration of TRICH Activity

Ion channel activity of TRICH is demonstrated using an electrophysiological assay for ion conductance. TRICH can be expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector encoding TRICH. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. A second plasmid which expresses any one of a number of marker genes, such as ß-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 ß-galactosidase.

Transformed cells expressing ß-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 ß-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 MΩ 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. In particular, the activity of TRICH-25 is measured as Cl- conductance.

Transport activity of TRICH is assayed by measuring uptake of labeled substrates into Xenopus laevis oocytes. Oocytes at stages V and VI are injected with TRICH mRNA (10 ng per oocyte) and incubated for 3 days at 18°C in OR2 medium (82.5mM NaCl, 2.5 mM KCl, 1mM $CaCl_2$, 1mM $MgCl_2$, 1mM Na_2HPO_4, 5 mM Hepes, 3.8 mM NaOH , 50μg/ml gentamycin, pH 7.8) to allow expression of TRICH. Oocytes are then transferred to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM $CaCl_2$, 1mM $MgCl_2$, 10 mM Hepes/Tris pH 7.5). Uptake of various substrates (e.g., amino acids, sugars, drugs, ions, and neurotransmitters) is initiated by adding labeled substrate (e.g. radiolabeled with ^3H, fluorescently labeled with rhodamine, etc.) to the oocytes. After incubating for 30 minutes, uptake is terminated by washing the oocytes three times in Na^+-free medium, measuring the incorporated label, and comparing with controls. TRICH activity is proportional to the level of internalized labeled substrate. In particular, test substrates include amino acids for TRICH-1, xanthine and uracil for TRICH-3, melibiose for TRICH-18, monocarboxylate for TRICH-20, neurotransmitters such as gamma-aminobutyric acid (GABA) for TRICH-22, and nucleosides for TRICH-23.

ATPase activity associated with TRICH can be measured by hydrolysis of radiolabeled ATP-[γ-^{32}P], separation of the hydrolysis products by chromatographic methods, and quantitation of the recovered ^{32}P using a scintillation counter. The reaction mixture contains ATP-[γ-^{32}P] 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 ^{32}P liberated is counted in a scintillation counter. The amount of radioactivity recovered is proportional to the ATPase activity of TRICH in the assay.

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 (Velicelebi, G. et al. (1999) Meth. Enzymol. 294:20-47; West, M.R. and C.R. Molloy (1996) Anal. Biochem. 241:51-58). These assays may be adapted for high-throughput screening using microplates. Changes in internal ion concentration are measured using fluorescent dyes such as the

Ca^{2+} indicator Fluo-4 AM, sodium-sensitive dyes such as SBFI and sodium green, or the Cl^- indicator MQAE (all available from Molecular Probes) in combination with the FLIPR fluorimetric plate reading system (Molecular Devices). In a more generic version of this assay, changes in membrane potential caused by ionic flux across the plasma membrane are measured using oxonyl dyes such as $DiBAC_4$ (Molecular Probes). $DiBAC_4$ 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 $DiBAC_4$ entry into the cell (Gonzalez, J.E. and P.A. Negulescu (1998) Curr. Opin. Biotechnol. 9:624-631). Candidate agonists or antagonists may be selected from known ion channel agonists or antagonists, peptide libraries, or combinatorial chemical libraries.

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

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
1687189	1	1687189CD1	27	1687189CB1
7078207	2	7078207CD1	28	7078207CB1
1560619	3	1560619CD1	29	1560619CB1
2614283	4	2614283CD1	30	2614283CB1
2667691	5	2667691CD1	31	2667691CB1
3211415	6	3211415CD1	32	3211415CB1
4739923	7	4739923CD1	33	4739923CB1
55030459	8	55030459CD1	34	55030459CB1
6113039	9	6113039CD1	35	6113039CB1
7101781	10	7101781CD1	36	7101781CB1
7473036	11	7473036CD1	37	7473036CB1
7476943	12	7476943CD1	38	7476943CB1
8003355	13	8003355CD1	39	8003355CB1
3116448	14	3116448CD1	40	3116448CB1
622868	15	622868CD1	41	622868CB1
7476494	16	7476494CD1	42	7476494CB1
7477260	17	7477260CD1	43	7477260CB1
1963058	18	1963058CD1	44	1963058CB1
2395967	19	2395967CD1	45	2395967CB1
3586648	20	3586648CD1	46	3586648CB1
7473396	21	7473396CD1	47	7473396CB1
7476283	22	7476283CD1	48	7476283CB1
7477105	23	7477105CD1	49	7477105CB1
7482079	24	7482079CD1	50	7482079CB1
55145506	25	55145506CD1	51	55145506CB1
5950519	26	5950519CD1	52	5950519CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
1	1687189CD1	g2116552	7.80E-275	[Rattus norvegicus] cationic amino acid transporter 3 (Hosokawa,H. et al. (1997) J. Biol. Chem. 272 (13), 8717-8722)
2	7078207CD1	g495259	0	[Mus musculus] abc2 (Luciani,M.F. et al. (1994) Genomics 21 (1), 150-159)
3	1560619CD1	g1002424	8.60E-253	[Mus musculus] YSPL-1 form 1 (Guimaraes,M.J. et al. (1995) Development 121 (10), 3335-3346)
4	2614283CD1	g1256378	1.90E-152	[Rattus norvegicus] zinc transporter ZnT-2 (Palmiter,R.D. et al. (1996) EMBO J. 15 (8), 1784-1791)
5	2667691CD1	g2506078	2.30E-259	[Mus musculus] tetracycline transporter-like protein (Matsuo,N. et al. (1997) Biochem. Biophys. Res. Commun. 238 (1), 126-129)
6	3211415CD1	g7243710	9.80E-197	[Mus musculus] zinc transporter like 2
7	4739923CD1	g13785620	4.00E-96	[3' incom][Mus musculus] sideroflexin 5 (Fleming,M.D. et al. (2001) Genes Dev. 15 (6), 652-657)
8	5503045 9CD1	g4186073	9.40E-15	[Mus musculus] calcium channel alpha-2-delta-C subunit (Klugbauer,N. et al. (1999) J. Neurosci. 19(2), 684-691)
9	6113039CD1	g310183	5.00E-273	[Rattus norvegicus] sodium dependent sulfate transporter (Markovich,D. et al. (1993) Proc. Natl. Acad. Sci. U.S.A. 90, 8073-8077)
10	7101781CD1	g13506808	0	[fl][Mus musculus] thymic stromal co-transporter (Chen,C. et al. (2000) Biochim. Biophys. Acta 1493 (1-2), 159-169)
11	7473036CD1	g13249295	0	[fl][Homo sapiens] anion exchanger AE4 (Parker,M.D. et al. (2001) Biochem. Biophys. Res. Commun. 282 (5), 1103-1109)
12	7476943CD1	g3047402	5.00E-67	[Homo sapiens] acn_cds
13	8003355CD1	g8256618	3.00E-273	[Homo sapiens] acn_cds (Shibahara,S. et al. (1985) Eur. J. Biochem. 146 (1), 15-22)

Table 2 (cont.)

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
14	3116448CD1	gi|7732815	0	[fl][Homo sapiens] concentrative Na+-nucleoside cotransporter hCNT3 (Ritzel,M.W.L. et al. (2001) J. Biol. Chem. 276 (4), 2914-2927)
15	622868CD1	g5924012	9.50E-160	[Homo sapiens] dJ261K5.1 (novel organic cation transporter (BAC ORF RG331P03))
16	7476494CD1	g8979801	3.90E-147	[Homo sapiens] dJ37C10.3 (novel ATPase)
17	7477260CD1	g13447747	0	[fl][Homo sapiens] sodium bicarbonate cotransporter NBC4a (Pushkin,A. et al. (2000) IUBMB Life 50 (1), 13-19)
18	1963058CD1	g1653342	6.80E-18	[Synechocystis sp.] melibiose carrier protein (Kaneko,T. et al. (1995) DNA Res. 2 (4), 153-166)
19	2395967CD1	g37643	3.20E-129	[Homo sapiens] vacuolar proton-ATPase (van Hille,B. et al. (1993) Biochem. Biophys. Res. Commun. 197 (1), 15-21)
20	3586648CD1	g2198807	8.90E-49	[Gallus gallus] monocarboxylate transporter 3
		g2463628	6.00E-43	[fl][Homo sapiens] putative monocarboxylate transporter
21	7473396CD1	g2618842	1.10E-139	[Bacillus subtilis] exonuclease ABC subunit (Reizer,J. et al. (1998) Mol. Microbiol. 27 (6), 1157-1169)A
22	7476283CD1	g56176	4.40E-244	[Ratus norvegicus] GABA(A) receptor gamma-1 subunit (Ymer,S. et al. (1990) EMBO J. 9 (10), 3261-3267)
23	7477105CD1	g3176684	2.20E-11	[Arabidopsis thaliana] Contains similarity to equilibrative nucleoside transporter 1 gb|U81375 from Homo sapiens. ESTs gb|N65317, gb|r20785, gb|AA586285 and gb|AA711578 come from this gene
		g12656639	3.00E-05	[fl][Homo sapiens] equilibrative nucleoside transporter 3
24	7482079CD1	g2815899	9.60E-84	[Homo sapiens] Shab-related delayed-rectifier K+ channel, alpha (Shepard,A.R. et al. (1999) Am. J. Physiol. 277 (3), C412-C424)
25	5514550GCD1	g289404	4.70E-105	[Bos taurus] chloride channel protein (Landry,D. et al. (1993) J. Biol. Chem. 268, 14948-14955)

Table 2 (cont.)

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
26	5950519CD1	g2352427	6.40E-156	[Oryctolagus cuniculus] peroxisomal Ca-dependent solute carrier (Weber,F.E. et al. (1997) Proc. Natl. Acad. Sci. U.S.A. 94 (16), 8509-8514)

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	1687189CD1	619	S134 S33 S453 S589 S599 T104 T18 T220 T272 T273 T438 T451 Y224	N232	Transmembrane domains: C31-Y51, S65-A85, D165-A183, V196-V214, P383-F401, M410-L428, V479-W498, L508-W528, A543-M562, W567-I593 Amino acid permeases signature BL00218: I66-A97, C343-T382 CATIONIC AMINO ACID TRANSPORTER PD034711: Q431-I523 AMINO ACID CATIONIC TRANSPORTER TRANSPORT TRANSMEMBRANE GLYCOPROTEIN TRANSPORTER1 PROTEIN HIGHAFFINITY PD000262: V526-Q597 TRANSMEMBRANE TRANSPORT PROTEIN TRANSPORTER AMINOACID PERMEASE AMINO ACID GLYCOPROTEIN MEMBRANE PD000214: L28-L428 do ANTTPORTER; ORNITHINE; PUTRESCINE; TRANSPORT; DM01125|P30825|23-373: E25-R371	HMMER BLIMPS_BLOCKS BLAST_PRODOM BLAST_PRODOM BLAST_PRODOM BLAST_DOMO

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
2	7078207CD1	2436	S1114 S1119 S1133 S1191 S1248 S1323 S1332 S1339 S1381 S140 S1411 S1427 S1455 S1478 S1560 S1604 S1687 S1819 S1982 S199 S2024 S2036 S2062 S21 S2159 S2196 S2245 S2292 S2333 S2366 S2420 S256 S281 S467 S50 S502 S533 S631 S884 S940 S959 S971 T1058 T1081 T1212 T1271 T1313 T1314 T1532 T16 T2097 T2102 T2108 T2144 T2215 T2235 T2284 T2352 T2413 T252 T353 T382 T440 T48 T612 T633 T696 T844 T955 Y1390	N14 N1409 N1497 N1550 N1558 N1613 N1678 N169 N174 N1776 N2055 N306 N369 N380 N421 N433 N477 N485 N495 N531 N545 N591 N601 N629 N90	Transmembrane domains: P22-K45, V784-L803, L893-T911, V1793-P1813, M1845-F1862, V1900-L1926 ABC transporter domains: N1018-G1198, G2081-G2262 ABC transporters family signature: D1105-D1155, V2167-D2218	HMMER HMMER_PFAM PROFILESCAN

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
2					ATPBINDING TRANSPORTER CASSETTE ABC TRANSPORT PROTEIN GLYCOPROTEIN TRANSMEMBRANE RIM ABCR PD005939: L1787-Y1971 PD006867: I663-S809 PD006385: A811-K1001 PD010138: G1957-K2061 ABC TRANSPORTERS FAMILY DM00008|P41133|839-1045: V991-H1197, V2051-M2259 ABC transporter motif: L1124-F1138 ATP/GTP binding site (P-loop): G1025-T1032, G2088-T2095 Lipocalin motif: G1424-V1437, G1426-V1437	BLAST_PRODOM BLAST_DOMO MOTIFS MOTIFS MOTIFS
3	1560619CD1	610	S127 S161 S251 S409 S450 S483 S582 S601 S608 T313 T514 T529	N139 N159	Transmembrane domains: F215-C233, L264-I286 Xanthine/uracil permeases family domain: G46-E473 Xanthine/uracil permease signature BL01116: G407-F443 YOLK SAC PERMEASELIKE YSPL1 FORM 1 YOLK SAC PERMEASELIKE YSPL1 FORM 4 YOLK SAC PERMEASELIKE YSPL1 FORM 3 YOLK SAC PERMEASELIKE YSPL1 FORM 2 PD019501: G429-Q609 PD137940: Q29-P83 PROTEIN TRANSPORT SULFATE TRANSPORTER TRANSMEMBRANE PERMEASE INTERGENIC REGION AFFINITY GLYCOPROT PD001255: L174-I467 XANTHINE/URACIL PERMEASES FAMILY DM01485|S33349|7-188: G355-L465	HMMER HMMER_PFAM BLIMPS_BLOCKS BLAST_PRODOM BLAST_PRODOM BLAST_DOMO

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
4	2614283CD1	372	S124 S216 S338 S61 T281		Transmembrane domain: I141-V159 Cation efflux family: P127-S358 ZINC TRANSPORTER CATION EFFLUX COBALT RESISTANCE PD001369: N214-S371 PD001602: Q71-H197 ZINC TRANSPORTER ZNT2 PD095371: A15-C81 TRANSPORTER; EFFLUX; ZINC; CXCD DM02892|P13512|9-157: G68-S204 DM02892|P20107|1-136: R72-T207 DM02892|P32798|2-127: R72-S168 DM02892|S54302|3-128: G68-I191	HMMER HMMER_PFAM BLAST_PRODOM BLAST_PRODOM BLAST_DOMO
5	2667691CD1	490	S212 S236 S455 S460 T406	N12 N453	Transmembrane domains: A40-V61, P123-V147, V191-V209, D243-Y257, V282-S302, L432-I448 TETRACYCLINE RESISTANCE PR01035: I130-T151, Y160-G182, P429-P449, V282-S302, W335-S355, V370-F393 HIPPOCAMPUS ABUNDANT PROTEIN TRANSCRIPT 1 TETRACYCLINE TRANSPORTER LIKE PROTEIN PD125679: Y394-V490 PD082602: M1-H39 Sugar transport proteins signatures: I92-T108	HMMER BLIMPS_PRINTS BLAST_PRODOM MOTIFS
6	3211415CD1	377	S223 S31 S321 S5 T338 T34 Y98	N45	Transmembrane domains: L106-S124, R140-F159, I270-L287 Cation efflux family: R91-A376	HMMER HMMER_PFAM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosyla-tion Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
6					ZINC TRANSPORTER CATION EFFLUX COBALT RESISTANCE PD001602:L38-Y158 TRANSPORTER; EFFLUX; ZINC; CXCD DM02892|S61568|396-545: D32-H166 DM02892|P20107|1-136: I29-S170 DM02892|P13512|9-157: R36-G162 DM02892|P32798|2-127: L42-I154	BLAST_PRODOM BLAST_DOMO
7	4739923CD1	340	S240 S272 S314 S319 S330 T56 T74	N127 N140 N153	CHROMOSOME PUTATIVE TRANSPORTER C17G6.15C TRANSPORT XV READING FRAME PD006986: S20-L274	BLAST_PRODOM
8	55030459CD1	1274	S1025 S1138 S1142 S1155 S1189 S1201 S1242 S134 S190 S238 S256 S298 S303 S353 S354 S40 S405 S430 S624 S664 S670 S700 S746 S79 S892 S894 S952 T1006 T1050 T1191 T221 T268 T272 T293 T349 T361 T581 T674 T717 T75 T755 T813 T852 T868 T987 Y1056 Y114	N145 N329 N373 N568 N587 N905 N940 N985	signal_cleavage: M1-A35 Transmembrane domain: V1096-R1118	SPSCAN HMMER

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
9	6113039CD1	595	S213 S214 S483 S74 T209 T230 T236 T240 T423 T97 Y39	N140 N174 N207 N591	Transmembrane domains: Y10-L30, F287-W305, V349-D365, G556-M576 Sodium:sulfate symporter BL01271: T131-I150, T240-I264, P432-G453, A505-I559 SODIUM SYMPORT OF COTRANSPORTER PD000549: R331-W572, L242-K402, V16-A167, F13-V154 SODIUM/SULFATE COTRANSPORTER NA+/SULFATE TRANSPORT TRANSMEMBRANE SODIUM SYMPORT PD084897: A161-K238	HMMER BLIMPS_BLOCKS BLAST_PRODOM BLAST_PRODOM
					co RENAL; BOUND; PRO-SER-ALA; NA; DM02914|A47714|28-576: I28-F577 DM02914|S43561|28-507: L242-I569, E34-A161 DM02914|P46556|1-520: K198-F577, E34-I159 DM02914|P32739|25-517: K238-F577, E34-V154	BLAST_DOMO
10	7101781CD1	475	S100 S108 S170 S34 S61 T20 T252 T390	N55	Transmembrane domains: I283-V308, Y322-V340, M350-E370, S440-V459	HMMER
11	7473036CD1	927	S149 S163 S217 S23 S260 S265 S325 S51 S65 S733 S738 S874 S891 S904 S906 T292 T324 T344 T567 T594 T629 T802 T99	N493 N520 N544 N923	Transmembrane domains: V444-Y466, V761-P780, I779-M810	HMMER

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
11					HCO3- transporter family: K108-I835 Anion exchangers family BL00219: V360-D383, W659-L700, G744-L789, Y790-T833, G89-H120, Q180-L223 Anion exchangers family signatures: A457-G509 ANION EXCHANGER SIGNATURE PR00165: Q355-G375, V388-G407, L442-S461, G474-L492, D570-L589, W657-M676 ANION EXCHANGE GLYCOPROTEIN PALMITATE BICARBONATE COTRANSPORTER PD001455: S346-L784, S505-I835, S156-F348, L109-V154 BICARBONATE COTRANSPORTER ELECTRO-GENIC NA+ PANCREAS HCO3 F52B5.1 PD018437: Q836-N927 BAND 3 ANION TRANSPORT PROTEIN DM02294|P04920|602-1237: G558-E894, S346-P529 DM02294|P48751|601-1229: S537-G896, S346-I543 DM02294|A42497|403-1027: S537-G896, S346-I500 DM02294|P02730|311-908: P560-D882, S346-G519	HMMER_PFAM BLIMPS_BLOCKS PROFILESCAN BLIMPS_PRINTS BLAST_PRODOM BLAST_PRODOM BLAST_DOMO
12	7476943CD1	516	S11 S137 S169 S202 S253 S41 S92 T228 T234 T244 T30 T340	N10 N333 N487	Transmembrane domains: I118-T144, S181-W203, A206-M224, Y275-M293	HMMER

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
12					Monocarboxylate transporter: C55-D499	HMMER_PFAM
					PEST; TRANSPORTER; LINKED; DM05037|P53988|1-465: P42-Q470 DM05037|Q03064|1-475: S41-D479 DM05037|P36021|155-612: A37-D258, V285-E477	BLAST_DOMO
13	8003355CD1	514	S174 S183 S330 S427 S453 S54 S64 T381 T382 Y94	M163 N328 N373 N52	signal_cleavage: M1-G22	SPSCAN
					signal peptide: M1-G22	HMMER
					Transmembrane domains: P241-F264, C274-V291, Y308-N328, V472-M491	HMMER
					Neurotransmitter-gated ion-channel: E26-F489	HMMER_PFAM
					Neurotransmitter-gated ion-channels proteins BL00236: V107-N116, D135-Y173, H228-A269, V53-D90	BLIMPS_BLOCKS
					Neurotransmitter-gated ion-channels signature: V130-Q184	PROFILESCAN
					Neurotransmitter-gated ion channel family signature PR00252: T73-R89, M106-N117, C150-C164, L235-N247	BLIMPS_PRINTS
					Nicotinic acetylcholine receptor sig. PR00254: T60-V76, Y94-W108, I112-V124, V130-S148	BLIMPS_PRINTS

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
13					CHANNEL IONIC GLYCOPROTEIN POSTSYNAPTIC RECEPTOR SIGNAL PROTEIN PD000153: N24-S393, A432-F489	BLAST_PRODOM
					NEUROTRANSMITTER-GATED ION-CHANNELS DM00195|P13536|7-501: P7-V497 DM00195|P02713|5-498: L8-R496 DM00195|P05376|2-493: L10-R496 DM00195|P02714|1-491: L8-V497	BLAST_DOMO
					Neurotransmitter-gated ion-channels signature: C150-C164	MOTIFS
14	3116448CD1	691	S326 S36 S549 S582 S63 S669 T100 T193 T262 T356 T411 T417 T50 T615 T637 Y87	N30 N34 N630 N636 N664	Transmembrane domain: I104-N124, W178-L207, L289-M308, I444-L461	HMMER
					Na+ dependent nucleoside transporter Nucleoside_tra2: Q198-S613	HMMER_PFAM
					Copper-transporting ATPase L131-D145	BLIMPS_PRINTS
					NA+/NUCLEOSIDE INNER MEMBRANE TRANSPORT PD003768: R223-I611, PD008773: F93-F215	BLAST_PRODOM
					NUCLEOSIDE; TRANSPORT; NaDEPENDENT DM01857|A54892|234-589: L256-L612 DM01857|A57532|230-585: L256-L612 DM01857|P44742|60-409: V260-I611 DM01857|P33021|60-412: V260-G610	BLAST_DOMO

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	622868CD1	342	S102 S309 S315 S325 S84 T121 T174 T286 T299 T300 T334	N110 N117 N311 N323	Transmembrane domain: Y205-Y227 PERIPHERIN (RDS)/ROM-1 F PR00218: V9-V29, L207-L228 SUGAR TRANSPORTER SIGNATURE PR00171: A231-V242 DM00135|P39932|141-478: W33-K295	HMMER BLIMPS_PRINTS BLIMPS_PRINTS BLAST_DOMO
16	7476494CD1	791	S103 S110 S199 S289 S514 S659 S66 S688 S734 S782 T122 T191 T287 T314 T326 T507 T710 T747 T78 Y293 Y742	N697 N768	Atpase_E1_E2 D437-T443 transmembrane domain: A177-Y193, D348-Y366 E1-E2 ATPase E1-E2_ATPase: C217-T443, P551-R680 E1-E2 ATPases phosphorylation site atpase_e1_e2.prf: I417-A471 E1-E2 ATPases phosphorylation site BL00154: V393-G429, L431-L449, K575-C585, N644-M684 P-type cation-transporting ATPase superfamily signature PR00119: D260-P274, C435-L449, A660-D670 Sodium/potassium-transporting ATPase signature PR00121: C428-L449, A572-V590 E1-E2 ATPASES PHOSPHORYLATION SITE DM00115|P22189|49-801: L547-V685 DM00115|P37278|58-755: Q224-I692 DM00115|A42764|65-737: E141-T699 DM00115|P37367|60-746: L226-V691 ATPBINDING CALCIUM MAGNESIUM TRANSPORT PUMP PD000132: I180-D445, A612-Q689, M559-C585	MOTIFS HMMER HMMER_PFAM PROFILESCAN BLIMPS_BLOCKS BLIMPS_PRINTS BLIMPS_PRINTS BLAST_DOMO BLAST_PRODOM

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
17	7477360CD1	1108	S1011 S1061 S1063 S1088 S124 S14 S190 S218 S240 S314 S319 S388 S391 S434 S435 S686 S701 S870 S95 T1030 T1056 T1065 T1093 T1102 T16 T183 T201 T454 T639 T678 T725 T766 T778 T78 Y1090	N399 N653 N658 N668 N676	Gene regulatory motif Leucine_Zippe L125-L146, L677-L698	MOTIFS
					Anion exchangers family signatures anion_exchanger1.prf: D438-P490 anion_exchanger2.prf: A585-T639	PROFILESCAN
					Transmembrane domain: I488-L506, L837-W856, I898-P917, V920-F938, I982-V1002	HMMER
					HCO3- transporter family HCO3_cotransp: K104-Y972	HMMER_PFAM
					Anion exchangers family BL00219: H85-H116, K259-V302, T304-K342, A343-K378, G448-A487, I488-D511, L541-Q579, L581-I628, F706-D759, V796-L837, D838-E876, G881-L926, Y927-T970, V972-S1011	BLIMPS_BLOCKS
					ANION EXCHANGER SIGNATURE PR00165: F458-F480, Q483-G503, V516-G535, I539-S558, L570-S589, G602-I620, D707-L726, L742-F762, W794-M813	BLIMPS_PRINTS
					BAND 3 ANION TRANSPORT PROTEIN DM02294|P48751|601-1229: P706-N1020, E449-P634, I353-E367	BLAST_DOMO
					PROTEIN ANION EXCHANGE TRANSMEMBRANE BAND GLYCOPROTEIN LIPOPROTEIN PALMITATE BICARBONATE COTRANSPORTER PD001455: H445-V972, V105-E394	BLAST_PRODOM
					BICARBONATE COTRANSPORTER SODIUM ELECTROGENIC NA+ PANCREAS PD018437: Q973-M1078 PD018439: A53-E103	BLAST_PRODOM

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
18	1963058CD1	480	S13 S194 S195 S204 S409 S49 S54 T286	N178 N219 N292 N341	Transmembrane domain: L349-Y369 Sodium:galactoside symporter family: DM01084|Q02581|1-462:L17-S195 (P-value = 8.2e-10)	HMMER BLAST-DOMO
19	2395367CD1	381	S119 S170 S202 S211 S269 S327 S349 S378 S74 T102 T144 T147 T164 T246 T26 T328 T62	N192 N285 N30	Vacuolar ATPase C subunit PD014267: E3-D376 Vacuolar ATP synthase: DM04365|P21282|1-381: M1-D381 DM04365|P54648|1-368: E3-L342 DM04365|P31412|1-392: I7-L377	BLAST-PRODOM BLAST-DOMO
20	3586648CD1	484	S236 S4 T21 T258 T290 T3 T312 Y301	N345 N389	Transmembrane domains: F42-W61, V75-I94, F311-Y327, I361-W382, W382-N402 Monocarboxylate transporter: S40-L478 Transporter DM05037|P53988|1-465: P16-N217 DM05037|Q03064|1-475: D29-Q263 DM05037|P36021|155-612: D29-L229	HMMER HMMER-PFAM BLAST-DOMO

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
21	7473396CD1	736	S236 S440 S462 S472 S501 S52 S579 S626 T10 T166 T191 T239 T316 T324 T345 T386 T491 T587 T607 T715 T89	N3 N367	Signal peptide: M1-G53	SPScan
					ABC transporter: G24-G210, G429-G700	HMMER-PFAM
					ABC transporter: L396-V410, L625-L639	MOTIFS
					ATP/GTP binding sites: G31-S38, G436-S443	MOTIFS
					ABC transporters family signatures: Q606-H659, L378-D427	ProfileScan
					ABC transporters family BL00211: L396-D427, L29-L40	BLIMPS-BLOCKS
					UVRA protein DM02034|P13567|759-959: F503-G704, I135-D202 DM02034|P07671|708-908: F503-G704, I135-D202 DM02034|S49424|2-201: D504-G704, K110-K198 DM02034|P47660|610-810: F503-G704, I135-D202	BLAST-DOMO
					Exclnuclease ABC subunit A PD001646: D504-T624	BLAST-PRODOM
					Exclnuclease ABC subunit A PD184930: C538-V713, R133-L297, V434-V502	BLAST-PRODOM
					Exclnuclease ABC subunit A PD003881: N447-P503	BLAST-PRODOM
					Ribose/galactose ABC transporter: PD035715: K241-K311, M1-I55	BLAST-PRODOM

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
22	7476283CD1	465	S153 S224 S267 S426 T118 T129 T179 T331 T88	N127 N245 N393 N50	Signal peptide: M1-C35	SPScan
					Transmembrane domain: M270-I294	HMMER
					Neurotransmitter-gated ion channel: I64-W459	HMMER-PFAM
					Neurotransmitter-gated ion-channels signature: L168-K222	ProfileScan
					Neurotransmitter-gated ion channel: C188-C202	MOTIFS
					Neurotransmitter-gated ion channel BL00236: I90-N127, I143-N152, D173-Y211, Y257-A298	BLIMPS-BLOCKS
					Neurotransmitter-gated ion channel: PR00252: T110-F126, K142-S153, C188-C202, F264-Q276	BLIMPS-PRINTS
					Gamma-aminobutyric acid receptor: PR00253: F273-W293, V299-A320, M333-L354, Y442-Y462	BLIMPS-PRINTS
					Gamma-aminobutyric acid receptor: PR01079: G62-Q73, D82-I99, F125-N138, W233-G255, K326-V339, I432-R444, V457-L465	BLIMPS-PRINTS
					Neurotransmitter-gated ion channel: DM00560|P23574|26-465: L26-L465 DM00560|P20237|20-556: L26-V396, A437-L463 DM00560|P16305|4-443: D63-S377, A437-L463 DM00560|P08219|14-456: T65-L463	BLAST-DOMO
					Ion channel/postsynaptic membrane receptor PD000153: N127-Y356, Q66-Y285	BLAST-PRODOM

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
22					Ion channel/postsynaptic membrane receptor PD000604: G403-L463	BLAST-PRODOM
23	7477105CD1	235	S151 S6 T218 T56 T57 T90		Transmembrane domains: S103-N123, I136-R162	HMMER
					Nucleoside transporter, equilibrative: PD006749: P63-L157 (P-value = 1.0e-07)	BLAST-PRODOM
24	7482079CD1	662	S10 S12 S137 S211 S323 S5 S564 T130 T19 T195 T281 T403 T499 T627 T657 T83 Y187	N17 N440 N517	Transmembrane domain: G412-Y430	HMMER
					K+ channel tetramerisation domain: S97-F203	HMMER-PFAM
					Ion transport protein: G263-L509	HMMER-PFAM
					Potassium channel signature PR00169: Q410-R433, F441-L463, G587-F613, E148-S167, P253-T281, H304-L327, F330-L350, L381-C407	BLIMPS-PRINTS
					Potassium channel CDRK: DM00436|JH0595|144-307: K215-L390 DM00436|P15387|136-299: R206-L381 DM00436|P17970|386-549: I216-L390 DM00490|P17970|268-384: A94-R200	BLAST-DOMO
					Voltage-gated potassium channel: PD000141: F330-S469, V570-K619, I572-I645	BLAST-PRODOM
25	55145506CD1	371	S113 S158 S194 S330 T151 T211 T323 T341 T63 T8	N127 N209	PROTEIN CHANNEL IONIC ION TRANSPORT VOLTAGEGATED P64 CHLORIDE INFRACELLULAR CHLORINE PD017366: A169-K355	BLAST_PRODOM
					CHLORINE CHANNEL PROTEIN P64 IONIC ION TRANSPORT VOLTAGEGATED TRANSMEMBRANE PHOSPHORYLATION PD118116: M1-Q125	BLAST_PRODOM

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
26	5950519CD1	468	S105 S176 S23 S4 S56 T161 T170 T220 T308 T358 T466		Mitochondrial carrier proteins domain: M184-Y276, H278-H369, G375-R468	HMMER_PFAM
					EF hand: R13-L41, R81-L109, Q117-H145	HMMER_PFAM
					Mitochondrial energy transfer proteins signature BL00215: V190-Q214, I425-G437	BLIMPS_BLOCKS
					Mitochondrial energy transfer proteins signature: K187-I241, V279-P331, I376-Q428	PROFILESCAN
					Mitochondrial carrier proteins signature PR00926: Q188-T201, T201-V215, G244-E264, T292-R310, Y335-L353, G383-Q405	BLIMPS_PRINTS
					Grave's disease carrier protein signature PR00928: P205-I225	BLIMPS_PRINTS
					PROTEIN TRANSPORT TRANSMEMBRANE RSPEAT MITOCHONDRION CARRIER MEMBRANE INNER MITOCHONDRIAL ADP/ATP PD000117: Q273-L463, K187-A293	BLAST_PRODOM
					MITOCHONDRIAL ENERGY TRANSFER PROTEINS DM00026|S57544|26-107: V190-I270 DM00026|P29518|233-310: V284-K360 DM00026|S54495|534-620: F283-N361 DM00026|Q01888|126-214: H278-N361	BLAST_DOMO
					EF hand motifs: D22-L34, D90-I102	MOTIFS
					Mitochondrial carrier proteins motif: P299-L307	MOTIFS

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
27	1687189CB1	2229	2190-2229, 759-1660	70564238V1	1269	1824
				7251266F7 (PROSTMY01)	1	693
				7030002301 (NOSEDIN01)	2059	2229
				7749453F8	410	1006
				70565215V1	1741	2207
				2416733F6 (HNT3AZT01)	1147	1688
				7711767U1 (TESTTUE02)	756	1205
28	7078207CB1	7610	1-5580	7070225H1 (BRAUTDR02)	6275	6895
				6911060U1 (PITUDIR01)	3289	3865
				6772031H1 (BRAINOR01)	1	696
				7106318J1 (LUNPTUT02)	2885	3400
				6253219H1 (LUNPTUT02)	1442	2064
				6893301H1 (BRAITDR03)	5879	6157
				7106586F9V1	2228	2898
				6763740H1 (BRAUNOR01)	3901	4568
				6765621H1 (BRAUNOR01)	3843	4506
				7467144H1 (LUNGNOS02)	3406	3914
				3767813H1 (BRSTNOT24)	5937	6250
				6953905H1 (BRAITDR02)	5169	5868
				6977243H1 (BRAPTDR04)	4508	5140
				8016696J1 (BMARTXB01)	2071	2870
				5964168H1 (BRATNOT05)	6042	6690
				6762808J1 (BRAUNOR01)	2849	3387
				6950389H1 (BRAITDR02)	725	1478
				4098906F8 (BRAITTP26)	6463	7085
				7757265J1 (SPLNTUB01)	660	1225
				5098681F8 (BPTMNON05)	7071	7610
				7179893H1 (BRAXDIC01)	6909	7418
				7196965S3V1	1561	2109
				6908865J1 (PITUDIR01)	4613	5241
				6893778J1 (BRAITDR03)	5243	5900
29	1560619CB1	2219	1-1659	6452362F8 (COLNDIC01)	1	553
				7159747AV1	1539	2219
				7159478AV1	1331	2107
				7068317J1	1272	1783
				7068052J1	738	1346
				7159628J1	364	920
30	2614283CB1	1280	415-559	6020220D1	385	843
				8097352H1 (PITBNOA01)	850	1280
				7432729H1 (PANCTUR02)	425	1006

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
30				7987760H1 (UTRSTDC01)	1	439
31	2667691CB1	2727	1-330	5753102H1 (LUNGNOT35)	1332	1994
				7110038V1 (SINTNOR02)	654	1359
				GBI_g8081479_smcosh	1	266
				7312933H1 (SINTNOR02)	1968	2591
				7023389V1 (SINTNOR02)	266	752
				GBI_g9988362_smcosh	168	326
				8095950H1 (KIDNOB01)	741	1387
				7925641H2 (COLNTUS02)	2052	2727
				7346378H1 (SINODIN02)	1395	2014
32	3211415CB1	1631	1-43, 1303-1631	7006244V1 (KIDPTXS02)	704	1145
				5313185F8 (KIDPTXS02)	1	719
				7005921V1	1170	1631
				7005790V1	1010	1547
33	4739923CB1	2673	1483-1785, 1-37	7198215OV1	1187	1830
				7198685GV1	1513	2125
				4567241F7 (HELATXT01)	2337	2673
				7198344V1	1401	1841
				7260030H1 (BRAWNOC01)	1889	2499
				7997955H1 (BRAITUC02)	143	745
				6265341H1 (MCLDTXN03)	1	212
				3767715T6 (BRSTNOT24)	613	1253
34	55030459CB1	3958	1-274, 837-1623, 3909-3958	5503021911	1368	1996
				7199252SV1	2796	3508
				7199098ZV1	2748	3435
				6343107T8 (LUNGDIS03)	1167	1585
				GNN_g7960408_0000016_002_edit	1	198
				5503049L71	300	785
				5503045911	809	1368
				7199288GV1	2086	2783
				7199595S1	3248	3958
				7199032GV1	1875	2578
				5510963711	495	1288
				9401850G	48	528
35	61130390B1	2000	856-1096	7172164SV1	1296	2000
				6782480P9 (SINTNOC01)	1	653
				7172271GV1	411	1050
				7171983GV1	1000	1641

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
36	7101781CB1	1997	1568-1778, 644-1026	FL7101781_g7939384_000014_g8131858	174	1543
				7092535V1	1321	1579
				3748173F6 (UTHSNOT18)	1	726
				7099050201	1424	1997
37	7473036CB1	3069	1-1362, 2182-2313, 1436-1763	FL7473036_g9255974_000002_g2198815	1	2839
				5050192F6 (BRSTNOT33)	2484	3069
38	7476943CB1	2241	1-168, 1540-2241, 282-745	6392952F6 (PANCNON03)	1562	2241
				FL7476943_g7539804_000008_g3107402	660	1765
				5514401V1	1	879
39	8003355CB1	1593	1-38, 1173-1315	8003355H1 (MUSCTDC01)	28	620
				GNN.g765171_000004_004	49	1593
				3292859H1 (BONFEPT01)	1	248
40	3116446CB1	2121	358-692, 1-22	2378367F6 (ISLTNOT01)	1258	1771
				5513620612	1	779
				5723184F6 (SEMVNOT05)	1487	2121
				7076906101	1069	1669
				5513620HI	4	858
				7169977H1 (MCLRNOC01)	771	1092
41	623868CB1	1225	1-87	7050176801	708	1225
				1851960F6 (LUNGFET03)	1	529
				7050213401	624	1196
				7050118201	470	1114
42	7476494CB1	2693	1-1295, 2361-2451	1382551F6 (BRATUT08)	1	504
				FL7476494_g9438678_000004_g7688148_1_5-6	2035	2271
				7175426H1 (BRSTTMC01)	1537	2100
				GNN.g9438678_000004_002	1707	2642
				FL7476494_g9438678_000004_g7688148_1_6-7	2147	2435
				5511634T21	401	1094
				7757711H1 (SPLNTUB01)	762	1236
				FL7476494_g9438678_000004_g7688148_1_8-9	2436	2693
				7757711T1 (SPLNTUB01)	1204	1863
43	7477260CB1	3569	1-2249, 3310-3569	GNN.g9468993_000014_02.edit	3130	3569

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
43				5546177F8 (TESTNOC01)	2197	3093
				5313313F8 (KIDETXS02)	995	1522
				8011222H1 (NOSEDIC02)	1	767
				55089843H1 (PROTDNV21)	668	1007
				7227359H1 (BRAXTDR15)	2847	3383
				55120438J1	1497	2198
44	1963058CB1	3920	1-648, 2120-3920	6769623H1 (BRAUNOR01)	1745	2338
				7611869J1 (KIDCTME01)	3090	3825
				7696528H1 (KIDPTDE01)	2104	2636
				7001668H1 (HEALDIR01)	511	1014
				1963058R6 (BRSTNOT04)	3337	3920
				8174904H1 (FETANOA01)	964	1575
				6770724H1 (BRAUNOR01)	1587	2140
				3164103H1 (TLYMTXT04)	1329	1623
				7314310H1 (UTREDME02)	2753	3339
				2659167H1 (LUNGTUT09)	2421	2664
				7727654H1 (UTRCDIE01)	1	560
				7412125H1 (BONMTUE02)	631	1242
45	2395967CB1	1361	523-715	g4689801	944	1361
				2395967F6 (THP1AZT01)	642	1214
				71526782V1	1	612
				6411441H1 (UTREDIT10)	827	1359
				71469742V1	532	833
46	3586648CB1	1867	1-71, 1837-1867	2738605T6 (OVARNOT09)	1219	1825
				70855458V1	623	1267
				3586648F6 (293TF4T01)	1	575
				g1383637	1437	1867
				71224790V1	503	1044
47	7473396CB1	2211	1-2211	GNN.g9212516_1	1	2211
48	7476283CB1	1446	1053-1092, 1-265, 606-682, 1140-1192	GBI.g7684447_12_11_07_09_05_10.edit	49	1446
				55110089J1	1	257
				55110065J1	745	1184
49	7477105CB1	1332	1-819	71223112V1	434	1069
				7948245J1 (BRABNOE02)	1	512
				71040789V1	549	1156
				6711669H1 (BRABDIT01)	968	1332
50	7482079CB1	2298	1-732, 1302-1712, 861-895	GNN.g9650542_2	1	1989

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
50				g3765560	1929	2298
51	55145506CB1	2250	1-555, 1490-1754, 1320-1374, 1094-1143	7239605IV1	1286	1935
				7239304?V1	1615	2250
				7077127?V1	1235	1822
				55145606d1	1	660
				7077288?V1	717	1293
				7239433?V1	610	1245
52	5950519CB1	3430	1-35, 3250-3430, 3109-3130, 2255-2277	2106229T6 (BRATTDT03)	2926	3404
				7037884?D1	1595	2198
				7096023H1 (BRACDIR02)	2177	2851
				6327536H1 (BRANDIN01)	2987	3430
				6764621H1 (BRAUNOR01)	1331	1899
				6764621J1 (BRAUNOR01)	1	708
				6307874H1 (NEROTMN03)	668	1258
				6980581H1 (BRAHTDR04)	1177	1527
				6121971H1 (BRAHNON05)	2426	2986

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
27	1687189CB1	PROSTMY01
28	7078207CB1	BRAUNOR01
29	1560619CB1	LUNGNOT37
30	2614283CB1	PROSTUT09
31	2667691CB1	STOMFET01
32	3211415CB1	BLADNOT08
33	4739923CB1	BRAITUT03
34	5503045?CB1	BRAYDIN03
35	6113039CB1	SINTTMC01
36	7101781CB1	LUNGNOT34
37	7473036CB1	BRSTNOT33
38	7476943CB1	PANCNOM03
39	8003355CB1	BONRFET01
40	3116440CB1	SDMVNOT05
41	622868CB1	PGANNOT01
42	7476494CB1	SPLNFUB01
43	7477260CB1	TESTNOC01
44	1963058CB1	BRAUNOR01
45	2395967CB1	THP1AZT01
46	3586648CB1	OVARNOT09
49	7477105CB1	COLNNOT11
51	5514550GCB1	SINTTMS01
52	5950519CB1	BRAUNOR01

Table 6

Library	Vector	Library Description
BLADNOT08	pINCY	Library was constructed using RNA isolated from the bladder tissue of an 11-year-old black male, who died from a gunshot wound.
BONRFET01	pINCY	Library was constructed using RNA isolated from rib bone tissue removed from a Caucasian male fetus, who died from Patau's syndrome (trisomy 13) at 20-weeks' gestation.
BRATTUT03	PSPORT1	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 meningeal lesion. Pathology indicated a grade 4 fibrillary giant and small-cell astrocytoma. Family history included benign hypertension and cerebrovascular disease.
BRAUNOR01	pINCY	This random primed library was constructed using RNA isolated from striatum, globus pallidus and posterior putamen tissue removed from an 81-year-old Caucasian female who died from a hemorrhage and ruptured thoracic aorta due to atherosclerosis. Pathology indicated moderate atherosclerosis involving the internal carotids, bilaterally; microscopic infarcts of the frontal cortex and hippocampus; and scattered diffuse amyloid plaques and neurofibrillary tangles, consistent with age. Grossly, the leptomeninges showed only mild thickening and hyalinization along the superior sagittal sinus. The remainder of the leptomeninges was thin and contained some congested blood vessels. Mild atrophy was found mostly in the frontal poles and lobes, and temporal lobes, bilaterally. Microscopically, there were pairs of Alzheimer type II astrocytes within the deep layers of the neocortex. There was increased satellitosis around neurons in the deep gray matter in the middle frontal cortex. The amygdala contained rare diffuse plaques and neurofibrillary tangles. The posterior hippocampus contained a microscopic area of cystic cavitation with hemosiderin-laden macrophages surrounded by reactive gliosis. Patient history included sepsis, cholangitis, post-operative atelectasis, pneumonia CAD, cardiomegaly due to left ventricular hypertrophy, splenomegaly, arteriolonephrosclerosis, nodular colloidal goiter, emphysema, CHF, hypothyroidism, and peripheral vascular disease.
BRAYDIN03	pINCY	This normalized library was constructed from 6.7 million independent clones from a brain tissue library. Starting RNA was made from RNA isolated from diseased hypothalamus tissue removed from a 57-year-old Caucasian male who died from a cerebrovascular accident. Patient history included Huntington's disease and emphysema. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 -hours/round) reannealing hybridization was used. The library was linearized and recircularized to select

Table 6 (cont.)

Library	Vector	Library Description
		for insert containing clones.
BRSTNOT33	pINCY	Library was constructed using RNA isolated from right breast tissue removed from a 46-year-old Caucasian female during unilateral extended simple mastectomy with breast reconstruction. Pathology for the associated tumor tissue indicated invasive grade 3 adenocarcinoma, ductal type, with apocrine features, nuclear grade 3 forming a mass in the outer quadrant. There was greater than 50% intraductal component. Patient history included breast cancer.
COLANNOT11	PSPORT1	Library was constructed using RNA isolated from colon tissue removed from a 60-year-old Caucasian male during a left hemicolectomy.
LUNGNOT34	pINCY	Library was constructed using RNA isolated from lung tissue removed from a 12-year-old Caucasian male.
LUNGNOT37	pINCY	Library was constructed using RNA isolated from lung tissue removed from a 15-year-old Caucasian female who died from a closed head injury. Serology was positive for cytomegalovirus.
OVARNOT09	pINCY	Library was constructed using RNA isolated from ovarian tissue removed from a 28-year-old Caucasian female during a vaginal hysterectomy and removal of the fallopian tubes and ovaries. Pathology indicated multiple follicular cysts ranging in size from 0.4 to 1.5 cm in the right and left ovaries, chronic cervicitis and squamous metaplasia of the cervix, and endometrium in weakly proliferative phase. Family history included benign hypertension, hyperlipidemia, and atherosclerotic coronary artery disease.
PANCNOT03	pINCY	This normalized pancreas tissue library was constructed from 12 million independent clones from a pancreas library. Starting RNA was made from RNA isolated from pancreas tissue removed from a 17-year-old Caucasian female who died from head trauma. Serology was positive for cytomegalovirus and remaining serologies were negative. The patient was not taking any medications. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
PGANNOT01	PSPORT1	Library was constructed using RNA isolated from paraganglionic tumor tissue removed from the intra-abdominal region of a 46-year-old Caucasian male during exploratory laparotomy. Pathology indicated a benign paraganglioma and was associated with a grade 2 renal cell carcinoma, clear cell type, which did not penetrate the capsule. Surgical margins were negative for tumor.
PROSTMAY01	pINCY	This large size-fractionated cDNA and normalized library was constructed using RNA isolated from diseased prostate tissue removed from a 55-year-old Caucasian male during closed prostatic biopsy, radical prostatectomy, and regional lymph node

Table 6 (cont.)

Library	Vector	Library Description
		excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the matched tumor tissue indicated adenocarcinoma Gleason grade 4 forming a predominant mass involving the left side peripherally with extension into the right posterior superior region. The tumor invaded the capsule and perforated the capsule to involve periprostatic tissue in the left posterior superior region. The left inferior posterior and left superior posterior surgical margins are positive. One left pelvic lymph node is metastatically involved. Patient history included calculus of the kidney. Family history included lung cancer and breast cancer. The size-selected library was normalized in 1 round using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791.
PROSTUT09	pINCY	Library was constructed using RNA isolated from prostate tumor tissue removed from a 66-year-old Caucasian male during a radical prostatectomy, radical cystectomy, and urinary diversion. Pathology indicated grade 3 transitional cell carcinoma. The patient presented with prostatic inflammatory disease. Patient history included lung neoplasm, and benign hypertension. Family history included a malignant breast neoplasm, tuberculosis, cerebrovascular disease, atherosclerotic coronary artery disease and lung cancer.
SEMVNOT05	pINCY	Library was constructed using RNA isolated from seminal vesicle tissue removed from a 67-year-old Caucasian male during radical prostatectomy. Pathology for the associated tumor tissue indicated adenocarcinoma, Gleason grade 3 + 3.
SINITMC01	pINCY	This large size-fractionated library was constructed using pooled cDNA from two donors. cDNA was generated using mRNA isolated from ileum tissue removed from a 30-year-old Caucasian female (donor A) during partial colectomy, open liver biopsy, and permanent colostomy, and from ileum tissue removed from a 70-year-old Caucasian female (donor B) during right hemicolectomy, open liver biopsy, sigmoidoscopy, colonoscopy, and permanent colostomy. Pathology for the matched tumor tissue (donor A) indicated carcinoid tumor (grade 1 neuroendocrine carcinoma) arising in the terminal ileum. The tumor permeated through the ileal wall into the mesenteric fat and extended into the adherent cecum, where tumor extended through the bowel wall up to the mucosal surface. Multiple lymph nodes were positive for tumor. Additional (2) lymph nodes were also involved by direct tumor extension. Pathology for donor B indicated a non-tumorous margin of ileum. Pathology for the matched tumor (donor B) indicated invasive grade 2 adenocarcinoma forming an ulcerated mass, situated distal to the ileocecal valve. The tumor invaded through the muscularis propria just into the serosal adipose tissue. One regional lymph node was positive for a microfocus of metastatic

Table 6 (cont.)

Library	Vector	Library Description
		adenocarcinoma. Donor A presented with flushing and unspecified abdominal/pelvic symptoms. Patient history included endometriosis, and tobacco and alcohol abuse. Donor B's history included a malignant breast neoplasm, type II diabetes, hyperlipidemia, viral hepatitis, an unspecified thyroid disorder, osteoarthritis, and a malignant skin neoplasm. Donor B's medication included tamoxifen.
SINTTMR01	PCDNA2.1	This random primed library was constructed using RNA isolated from ileum tissue removed from a 70-year-old Caucasian female during right hemicolectomy, open liver biopsy, flexible sigmoidoscopy, colonoscopy, and permanent colostomy. Pathology for the matched tumor tissue indicated invasive grade 2 adenocarcinoma forming an ulcerated mass, situated 2 cm distal to the ileocecal valve. Patient history included a malignant breast neoplasm, type II diabetes, hyperlipidemia, viral hepatitis, an unspecified thyroid disorder, osteoarthritis, a malignant skin neoplasm, deficiency anemia, and normal delivery. Family history included breast cancer, atherosclerotic coronary artery disease, benign hypertension, cerebrovascular disease, ovarian cancer, and hyperlipidemia.
SPLNTUE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from spleen tumor tissue removed from a 28-year-old male during total splenectomy. Pathology indicated malignant lymphoma, diffuse large cell type, B-cell phenotype with abundant reactive T-cells and marked granulomatous response involving the spleen, where it formed approximately 45 nodules, liver, and multiple lymph nodes.
STOMFET01	pINCY	Library was constructed using RNA isolated from the stomach tissue of a Caucasian female fetus, who died at 20 weeks' gestation.
TESTNOC01	PBLUESCRIPT	This large size fractionated library was constructed using RNA isolated from testicular tissue removed from a pool of eleven, 10 to 61-year-old Caucasian males.
THP1AZT01	pINCY	Library was constructed using RNA isolated from THP-1 promonocyte cells treated for three days with 0.8 micromolar 5-aza-2'-deoxycytidine. THP-1 (ATCC TIB 202) is a human promonocyte line derived from peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (Int. J. Cancer (1980) 26:171).

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL PDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences:* Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs:* fasta Identity= 95% or greater and Match length=200 bases or greater; fastx 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 Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	*PFAM hits:* Probability value= 1.0E-3 or less *Signal peptide hits:* Score= 0 or greater

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Pfrerl	A base-calling algorithm that examines autocorrelated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; 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
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth 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.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
   a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26,
   b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26,
   c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and
   d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-26.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:27-52.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method of producing a polypeptide of claim 1, the method comprising:
   a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. A method of claim 9, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.

11. An isolated antibody which specifically binds to a polypeptide of claim 1.

12. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52,

b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52,

c) a polynucleotide complementary to a polynucleotide of a),

d) a polynucleotide complementary to a polynucleotide of b), and

e) an RNA equivalent of a)-d).

13. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 12.

14. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

15. A method of claim 14, wherein the probe comprises at least 60 contiguous nucleotides.

16. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 12, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

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

18. A composition of claim 17, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.

19. A method for treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment the composition of claim 17.

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

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

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

22. A method for treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment a composition of claim 21.

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

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

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

25. 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 24.

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

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

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

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

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

a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,

b) detecting altered expression of the target polynucleotide, and

c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

29. A method of assessing toxicity of a test compound, the method comprising:

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

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,

c) quantifying the amount of hybridization complex, and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

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

a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and

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

31. The antibody of claim 11, wherein the antibody is:

a) a chimeric antibody,

b) a single chain antibody,

c) a Fab fragment,

d) a $F(ab')_2$ fragment, or

e) a humanized antibody.

32. A composition comprising an antibody of claim 11 and an acceptable excipient.

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

34. A composition of claim 32, wherein the antibody is labeled.

35. A method of diagnosing a condition or disease associated with the expression of TRICH in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibodies from said animal, and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.

37. A polyclonal antibody produced by a method of claim 36.

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

39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, or an immunogenic fragment thereof, under conditions to elicit an antibody response,

b) isolating antibody producing cells from the animal,

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,

d) culturing the hybridoma cells, and

e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.

40. A monoclonal antibody produced by a method of claim 39.

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

42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.

43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.

44. A method of detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26 in a sample, the method comprising:

a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26 in the sample.

45. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26 from a sample, the method comprising:

a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and

b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.

46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.

47. A method of generating a transcript image of a sample which contains polynucleotides, the method comprising:

a) labeling the polynucleotides of the sample,

b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides of the sample under conditions suitable for the formation of a hybridization complex, and

c) quantifying the expression of the polynucleotides in the sample.

48. An array comprising different nucleotide molecules affixed in distinct physical locations on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide

or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

53. An array of claim 48, further comprising said target polynucleotide hybridized to a nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to said solid substrate.

55. An array of claim 48, wherein each distinct physical location on the substrate contains multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical location have the same sequence, and each distinct physical location on the substrate contains nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at another distinct physical location on the substrate.

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

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

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

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

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

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

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

63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

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

81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.

82. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:27.

83. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:28.

84. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:29.

85. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:30.

86. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:31.

87. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:32.

88. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:33.

89. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:34.

90. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:35.

91. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:36.

92. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:37.

93. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:38.

94. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:39.

95. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:40.

96. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:41.

97. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:42.

98. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:43.

99. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:44.

100. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:45.

101. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:46.

102. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:47.

103. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:48.

104. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:49.

105. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:50.

106. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:51.

107. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:52.

<110> INCYTE GENOMICS, INC.
LEE, Ernestine A.
YUE, Henry
LAL, Preeti G.
WALIA, Narinder K.
BAUGHN, Mariah R.
WARREN, Bridget A.
LEE, Sally
SANJANWALA, Madhu S.
YAO, Monique G.
RAMKUMAR, Jayalaxmi
THORNTON, Michael
GANDHI, Ameena R.
POLICKY, Jennifer L.
ELLIOTT, Vicki S.
ARVIZU, Chandra
RAUMANN, Brigitte E.
BRUNS, Christopher M.
NAINA, Amir
HAFALIA, April J.A.
NGUYEN, Danniel B.
XU, Yuming
LU, Dyung Aina M.
ISON, Craig H.
GRIFFIN, Jennifer A.
REDDY, Roopa M.
BURFORD, Neil

<120> TRANSPORTERS AND ION CHANNELS

<130> PI-0217 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/232,685; 60/234,842; 60/236,882; 60/239,057; 60/240,540;
60/241,700
<151> 2000-09-15; 2000-09-22; 2000-09-29; 2000-10-05; 2000-13-10;
2000-10-18

<160> 52
<170> PERL Program

<210> 1
<211> 619
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1687189CD1

<400> 1
Met Pro Trp Gln Ala Phe Arg Arg Phe Gly Gln Lys Leu Val Arg
1 5 10 15
Arg Arg Thr Leu Glu Ser Gly Met Ala Glu Thr Arg Leu Ala Arg

```
                  20                  25                  30
Cys Leu Ser Thr Leu Asp Leu Val Ala Leu Gly Val Gly Ser Thr
                  35                  40                  45
Leu Gly Ala Gly Val Tyr Val Leu Ala Gly Glu Val Ala Lys Asp
                  50                  55                  60
Lys Ala Gly Pro Ser Ile Val Ile Cys Phe Leu Val Ala Ala Leu
                  65                  70                  75
Ser Ser Val Leu Ala Gly Leu Cys Tyr Ala Glu Phe Gly Ala Arg
                  80                  85                  90
Val Pro Arg Ser Gly Ser Ala Tyr Leu Tyr Ser Tyr Val Thr Val
                  95                 100                 105
Gly Glu Leu Trp Ala Phe Thr Thr Gly Trp Asn Leu Ile Leu Ser
                 110                 115                 120
Tyr Val Ile Gly Thr Ala Ser Val Ala Arg Ala Trp Ser Ser Ala
                 125                 130                 135
Phe Asp Asn Leu Ile Gly Asn His Ile Ser Lys Thr Leu Gln Gly
                 140                 145                 150
Ser Ile Ala Leu His Val Pro His Val Leu Ala Glu Tyr Pro Asp
                 155                 160                 165
Phe Phe Ala Leu Gly Leu Val Leu Leu Leu Thr Gly Leu Leu Ala
                 170                 175                 180
Leu Gly Ala Ser Glu Ser Ala Leu Val Thr Lys Val Phe Thr Gly
                 185                 190                 195
Val Asn Leu Leu Val Leu Gly Phe Val Met Ile Ser Gly Phe Val
                 200                 205                 210
Lys Gly Asp Val His Asn Trp Lys Leu Thr Glu Glu Asp Tyr Glu
                 215                 220                 225
Leu Ala Met Ala Glu Leu Asn Asp Thr Tyr Ser Leu Gly Pro Leu
                 230                 235                 240
Gly Ser Gly Gly Phe Val Pro Phe Gly Phe Glu Gly Ile Leu Arg
                 245                 250                 255
Gly Ala Ala Thr Cys Phe Tyr Ala Phe Val Gly Phe Asp Cys Ile
                 260                 265                 270
Ala Thr Thr Gly Glu Glu Ala Gln Asn Pro Gln Arg Ser Ile Pro
                 275                 280                 285
Met Gly Ile Val Ile Ser Leu Ser Val Cys Phe Leu Ala Tyr Phe
                 290                 295                 300
Ala Val Ser Ser Ala Leu Thr Leu Met Met Pro Tyr Tyr Gln Leu
                 305                 310                 315
Gln Pro Glu Ser Pro Leu Pro Glu Ala Phe Leu Tyr Ile Gly Trp
                 320                 325                 330
Ala Pro Ala Arg Tyr Val Val Ala Val Gly Ser Leu Cys Ala Leu
                 335                 340                 345
Ser Thr Ser Leu Leu Gly Ser Met Phe Pro Met Pro Arg Val Ile
                 350                 355                 360
Tyr Ala Met Ala Glu Asp Gly Leu Leu Phe Arg Val Leu Ala Arg
                 365                 370                 375
Ile His Thr Gly Thr Arg Thr Pro Ile Ile Ala Thr Val Val Ser
                 380                 385                 390
Gly Ile Ile Ala Ala Phe Met Ala Phe Leu Phe Lys Leu Thr Asp
                 395                 400                 405
Leu Val Asp Leu Met Ser Ile Gly Thr Leu Leu Ala Tyr Ser Leu
                 410                 415                 420
Val Ser Ile Cys Val Leu Ile Leu Arg Tyr Gln Pro Asp Gln Glu
                 425                 430                 435
Thr Lys Thr Gly Glu Glu Val Glu Leu Gln Glu Glu Ala Ile Thr
```

```
                440                 445                 450
Thr Glu Ser Glu Lys Leu Thr Leu Trp Gly Leu Phe Phe Pro Leu
                455                 460                 465
Asn Ser Ile Pro Thr Pro Leu Ser Gly Gln Ile Val Tyr Val Cys
                470                 475                 480
Ser Ser Leu Leu Ala Val Leu Leu Thr Ala Leu Cys Leu Val Leu
                485                 490                 495
Ala Gln Trp Ser Val Pro Leu Leu Ser Gly Asp Leu Leu Trp Thr
                500                 505                 510
Ala Val Val Val Leu Leu Leu Leu Leu Ile Ile Gly Ile Ile Val
                515                 520                 525
Val Ile Trp Arg Gln Pro Gln Ser Ser Thr Pro Leu His Phe Lys
                530                 535                 540
Val Pro Ala Leu Pro Leu Leu Pro Leu Met Ser Ile Phe Val Asn
                545                 550                 555
Ile Tyr Leu Met Met Gln Met Thr Ala Gly Thr Trp Ala Arg Phe
                560                 565                 570
Gly Val Trp Met Leu Ile Gly Phe Ala Ile Tyr Phe Gly Tyr Gly
                575                 580                 585
Ile Gln His Ser Leu Glu Glu Ile Lys Ser Asn Gln Pro Ser Arg
                590                 595                 600
Lys Ser Arg Ala Lys Thr Val Asp Leu Asp Pro Gly Thr Leu Tyr
                605                 610                 615
Val His Ser Val
```

<210> 2
<211> 2436
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7078207CD1

<400> 2

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

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

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

```
Val Cys Val Asp Lys Leu Thr Lys Val Tyr Lys Asp Asp Lys Lys
                995                 1000                1005
Leu Ala Leu Asn Lys Leu Ser Leu Asn Leu Tyr Glu Asn Gln Val
           1010                1015                1020
Val Ser Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Met
           1025                1030                1035
Ser Ile Leu Thr Gly Leu Phe Pro Pro Thr Ser Gly Ser Ala Thr
           1040                1045                1050
Ile Tyr Gly His Asp Ile Arg Thr Glu Met Asp Glu Ile Arg Lys
           1055                1060                1065
Asn Leu Gly Met Cys Pro Gln His Asn Val Leu Phe Asp Arg Leu
           1070                1075                1080
Thr Val Glu Glu His Leu Trp Phe Tyr Ser Arg Leu Lys Ser Met
           1085                1090                1095
Ala Gln Glu Glu Ile Arg Arg Glu Met Asp Lys Met Ile Glu Asp
           1100                1105                1110
Leu Glu Leu Ser Asn Lys Arg His Ser Leu Val Gln Thr Leu Ser
           1115                1120                1125
Gly Gly Met Lys Arg Lys Leu Ser Val Ala Ile Ala Phe Val Gly
           1130                1135                1140
Gly Ser Arg Ala Ile Ile Leu Asp Glu Pro Thr Ala Gly Val Asp
           1145                1150                1155
Pro Tyr Ala Arg Arg Ala Ile Trp Asp Leu Ile Leu Lys Tyr Lys
           1160                1165                1170
Pro Gly Arg Thr Ile Leu Leu Ser Thr His His Met Asp Glu Ala
           1175                1180                1185
Asp Leu Leu Gly Asp Arg Ile Ala Ile Ile Ser His Gly Lys Leu
           1190                1195                1200
Lys Cys Cys Gly Ser Pro Leu Phe Leu Lys Gly Thr Tyr Gly Asp
           1205                1210                1215
Gly Tyr Arg Leu Thr Leu Val Lys Arg Pro Ala Glu Pro Gly Gly
           1220                1225                1230
Pro Gln Glu Pro Gly Leu Ala Ser Ser Pro Pro Gly Arg Ala Pro
           1235                1240                1245
Leu Ser Ser Cys Ser Glu Leu Gln Val Ser Gln Phe Ile Arg Lys
           1250                1255                1260
His Val Ala Ser Cys Leu Leu Val Ser Asp Thr Ser Thr Glu Leu
           1265                1270                1275
Ser Tyr Ile Leu Pro Ser Glu Ala Ala Lys Lys Gly Ala Phe Glu
           1280                1285                1290
Arg Leu Phe Gln His Leu Glu Arg Ser Leu Asp Ala Leu His Leu
           1295                1300                1305
Ser Ser Phe Gly Leu Met Asp Thr Thr Leu Glu Glu Val Phe Leu
           1310                1315                1320
Lys Val Ser Glu Glu Asp Gln Ser Leu Glu Asn Ser Glu Ala Asp
           1325                1330                1335
Val Lys Glu Ser Arg Lys Asp Val Leu Pro Gly Ala Glu Gly Pro
           1340                1345                1350
Ala Ser Gly Glu Gly His Ala Gly Asn Leu Ala Arg Cys Ser Glu
           1355                1360                1365
Leu Thr Gln Ser Gln Ala Ser Leu Gln Ser Ala Ser Ser Val Gly
           1370                1375                1380
Ser Ala Arg Gly Asp Glu Gly Ala Gly Tyr Thr Asp Val Tyr Gly
           1385                1390                1395
Asp Tyr Arg Pro Leu Phe Asp Asn Pro Gln Asp Pro Asp Asn Val
           1400                1405                1410
```

Ser Leu Gln Glu Val Glu Ala Glu Ala Leu Ser Arg Val Gly Gln
1415 1420 1425
Gly Ser Arg Lys Leu Asp Gly Gly Trp Leu Lys Val Arg Gln Phe
1430 1435 1440
His Gly Leu Leu Val Lys Arg Phe His Cys Ala Arg Arg Asn Ser
1445 1450 1455
Lys Ala Leu Phe Ser Gln Ile Leu Leu Pro Ala Phe Phe Val Cys
1460 1465 1470
Val Ala Met Thr Val Ala Leu Ser Val Pro Glu Ile Gly Asp Leu
1475 1480 1485
Pro Pro Leu Val Leu Ser Pro Ser Gln Tyr His Asn Tyr Thr Gln
1490 1495 1500
Pro Arg Gly Asn Phe Ile Pro Tyr Ala Asn Glu Glu Arg Arg Glu
1505 1510 1515
Tyr Arg Leu Arg Leu Ser Pro Asp Ala Ser Pro Gln Gln Leu Val
1520 1525 1530
Ser Thr Phe Arg Leu Pro Ser Gly Val Gly Ala Thr Cys Val Leu
1535 1540 1545
Lys Ser Pro Ala Asn Gly Ser Leu Gly Pro Thr Leu Asn Leu Ser
1550 1555 1560
Ser Gly Glu Ser Arg Leu Leu Ala Ala Arg Phe Phe Asp Ser Met
1565 1570 1575
Cys Leu Glu Ser Phe Thr Gln Gly Leu Pro Leu Ser Asn Phe Val
1580 1585 1590
Pro Pro Pro Pro Ser Pro Ala Pro Ser Asp Ser Pro Ala Ser Pro
1595 1600 1605
Asp Glu Asp Leu Gln Ala Trp Asn Val Ser Leu Pro Pro Thr Ala
1610 1615 1620
Gly Pro Glu Met Trp Thr Ser Ala Pro Ser Leu Pro Arg Leu Val
1625 1630 1635
Arg Glu Pro Val Arg Cys Thr Cys Ser Ala Gln Gly Thr Gly Phe
1640 1645 1650
Ser Cys Pro Ser Ser Val Gly Gly His Pro Pro Gln Met Arg Val
1655 1660 1665
Val Thr Gly Asp Ile Leu Thr Asp Ile Thr Gly His Asn Val Ser
1670 1675 1680
Glu Tyr Leu Leu Phe Thr Ser Asp Arg Phe Arg Leu His Arg Tyr
1685 1690 1695
Gly Ala Ile Thr Phe Gly Asn Val Leu Lys Ser Ile Pro Ala Ser
1700 1705 1710
Phe Gly Thr Arg Ala Pro Pro Met Val Arg Lys Ile Ala Val Arg
1715 1720 1725
Arg Ala Ala Gln Val Phe Tyr Asn Asn Lys Gly Tyr His Ser Met
1730 1735 1740
Pro Thr Tyr Leu Asn Ser Leu Asn Asn Ala Ile Leu Arg Ala Asn
1745 1750 1755
Leu Pro Lys Ser Lys Gly Asn Pro Ala Ala Tyr Gly Ile Thr Val
1760 1765 1770
Thr Asn His Pro Met Asn Lys Thr Ser Ala Ser Leu Ser Leu Asp
1775 1780 1785
Tyr Leu Leu Gln Gly Thr Asp Val Val Ile Ala Ile Phe Ile Ile
1790 1795 1800
Val Ala Met Ser Phe Val Pro Ala Ser Phe Val Val Phe Leu Val
1805 1810 1815
Ala Glu Lys Ser Thr Lys Ala Lys His Leu Gln Phe Val Ser Gly
1820 1825 1830

```
Cys Asn Pro Ile Ile Tyr Trp Leu Ala Asn Tyr Val Trp Asp Met
                1835                1840                1845
Leu Asn Tyr Leu Val Pro Ala Thr Cys Cys Val Ile Ile Leu Phe
                1850                1855                1860
Val Phe Asp Leu Pro Ala Tyr Thr Ser Pro Thr Asn Phe Pro Ala
                1865                1870                1875
Val Leu Ser Leu Phe Leu Leu Tyr Gly Trp Ser Ile Thr Pro Ile
                1880                1885                1890
Met Tyr Pro Ala Ser Phe Trp Phe Glu Val Pro Ser Ser Ala Tyr
                1895                1900                1905
Val Phe Leu Ile Val Ile Asn Leu Phe Ile Gly Ile Thr Ala Thr
                1910                1915                1920
Val Ala Thr Phe Leu Leu Gln Leu Phe Glu His Asp Lys Asp Leu
                1925                1930                1935
Lys Val Val Asn Ser Tyr Leu Lys Ser Cys Phe Leu Ile Phe Pro
                1940                1945                1950
Asn Tyr Asn Leu Gly His Gly Leu Met Glu Met Ala Tyr Asn Glu
                1955                1960                1965
Tyr Ile Asn Glu Tyr Tyr Ala Lys Ile Gly Gln Phe Asp Lys Met
                1970                1975                1980
Lys Ser Pro Phe Glu Trp Asp Ile Val Thr Arg Gly Leu Val Ala
                1985                1990                1995
Met Ala Val Glu Gly Val Val Gly Phe Leu Leu Thr Ile Met Cys
                2000                2005                2010
Gln Tyr Asn Phe Leu Arg Arg Pro Gln Arg Met Pro Val Ser Thr
                2015                2020                2025
Lys Pro Val Glu Asp Asp Val Asp Val Ala Ser Glu Arg Gln Arg
                2030                2035                2040
Val Leu Arg Gly Asp Ala Asp Asn Asp Met Val Lys Ile Glu Asn
                2045                2050                2055
Leu Thr Lys Val Tyr Lys Ser Arg Lys Ile Gly Arg Ile Leu Ala
                2060                2065                2070
Val Asp Arg Leu Cys Leu Gly Val Arg Pro Gly Glu Cys Phe Gly
                2075                2080                2085
Leu Leu Gly Val Asn Gly Ala Gly Lys Thr Ser Thr Phe Lys Met
                2090                2095                2100
Leu Thr Gly Asp Glu Ser Thr Thr Gly Gly Glu Ala Phe Val Asn
                2105                2110                2115
Gly His Ser Val Leu Lys Glu Leu Leu Gln Val Gln Gln Ser Leu
                2120                2125                2130
Gly Tyr Cys Pro Gln Cys Asp Ala Leu Phe Asp Glu Leu Thr Ala
                2135                2140                2145
Arg Glu His Leu Gln Leu Tyr Thr Arg Leu Arg Gly Ile Ser Trp
                2150                2155                2160
Lys Asp Glu Ala Arg Val Val Lys Trp Ala Leu Glu Lys Leu Glu
                2165                2170                2175
Leu Thr Lys Tyr Ala Asp Lys Pro Ala Gly Thr Tyr Ser Gly Gly
                2180                2185                2190
Asn Lys Arg Lys Leu Ser Thr Ala Ile Ala Leu Ile Gly Tyr Pro
                2195                2200                2205
Ala Phe Ile Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Lys
                2210                2215                2220
Ala Arg Arg Phe Leu Trp Asn Leu Ile Leu Asp Leu Ile Lys Thr
                2225                2230                2235
Gly Arg Ser Val Val Leu Thr Ser His Ser Met Glu Glu Cys Glu
                2240                2245                2250
```

```
Ala Leu Cys Thr Arg Leu Ala Ile Met Val Asn Gly Arg Leu Arg
                2255                2260                2265
Cys Leu Gly Ser Ile Gln His Leu Lys Asn Arg Phe Gly Asp Gly
                2270                2275                2280
Tyr Met Ile Thr Val Arg Thr Lys Ser Ser Gln Ser Val Lys Asp
                2285                2290                2295
Val Val Arg Phe Phe Asn Arg Asn Phe Pro Glu Ala Met Leu Lys
                2300                2305                2310
Glu Arg His His Thr Lys Val Gln Tyr Gln Leu Lys Ser Glu His
                2315                2320                2325
Ile Ser Leu Ala Gln Val Phe Ser Lys Met Glu Gln Val Ser Gly
                2330                2335                2340
Val Leu Gly Ile Glu Asp Tyr Ser Val Ser Gln Thr Thr Leu Asp
                2345                2350                2355
Asn Val Phe Val Asn Phe Ala Lys Lys Gln Ser Asp Asn Leu Glu
                2360                2365                2370
Gln Gln Glu Thr Glu Pro Pro Ser Ala Leu Gln Ser Pro Leu Gly
                2375                2380                2385
Cys Leu Leu Ser Leu Leu Arg Pro Arg Ser Ala Pro Thr Glu Leu
                2390                2395                2400
Arg Ala Leu Val Ala Asp Glu Pro Glu Asp Leu Asp Thr Glu Asp
                2405                2410                2415
Glu Gly Leu Ile Ser Phe Glu Glu Glu Arg Ala Gln Leu Ser Phe
                2420                2425                2430
Asn Thr Asp Thr Leu Cys
                2435

<210> 3
<211> 610
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1560619CD1

<400> 3
Met Ser Arg Ser Pro Leu Asn Pro Ser Gln Leu Arg Ser Val Gly
  1               5                  10                  15
Ser Gln Asp Ala Leu Ala Pro Leu Pro Pro Pro Ala Pro Gln Asn
                20                  25                  30
Pro Ser Thr His Ser Trp Asp Pro Leu Cys Gly Ser Leu Pro Trp
                35                  40                  45
Gly Leu Ser Cys Leu Leu Ala Leu Gln His Val Leu Val Met Ala
                50                  55                  60
Ser Leu Leu Cys Val Ser His Leu Leu Leu Leu Cys Ser Leu Ser
                65                  70                  75
Pro Gly Gly Leu Ser Tyr Ser Pro Ser Gln Leu Leu Ala Ser Ser
                80                  85                  90
Phe Phe Ser Cys Gly Met Ser Thr Ile Leu Gln Thr Trp Met Gly
                95                 100                 105
Ser Arg Leu Pro Leu Val Gln Ala Pro Ser Leu Glu Phe Leu Ile
               110                 115                 120
Pro Ala Leu Val Leu Thr Ser Gln Lys Leu Pro Arg Ala Ile Gln
               125                 130                 135
Thr Pro Gly Asn Ser Ser Leu Met Leu His Leu Cys Arg Gly Pro
```

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

```
                560                 565                 570
Pro Glu Asp Pro Gly Asp Glu Glu Gly Gly Ser Ser Glu Pro Glu
                575                 580                 585
Glu Met Ala Asp Leu Leu Pro Gly Ser Gly Glu Pro Cys Pro Glu
                590                 595                 600
Ser Ser Arg Glu Gly Phe Arg Ser Gln Lys
                605                 610

<210> 4
<211> 372
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2614283CD1

<400> 4
Met Glu Ala Lys Glu Lys Gln His Leu Leu Asp Ala Arg Pro Ala
  1                 5                  10                  15
Ile Arg Ser Tyr Thr Gly Ser Leu Trp Gln Glu Gly Ala Gly Trp
                 20                  25                  30
Ile Pro Leu Pro Arg Pro Gly Leu Asp Leu Gln Ala Ile Glu Leu
                 35                  40                  45
Ala Ala Gln Ser Asn His His Cys His Ala Gln Lys Gly Pro Asp
                 50                  55                  60
Ser His Cys Asp Pro Lys Lys Gly Lys Ala Gln Arg Gln Leu Tyr
                 65                  70                  75
Val Ala Ser Ala Ile Cys Leu Leu Phe Met Ile Gly Glu Val Val
                 80                  85                  90
Gly Gly Tyr Leu Ala His Ser Leu Ala Val Met Thr Asp Ala Ala
                 95                 100                 105
His Leu Leu Thr Asp Phe Ala Ser Met Leu Ile Ser Leu Phe Ser
                110                 115                 120
Leu Trp Met Ser Ser Arg Pro Ala Thr Lys Thr Met Asn Phe Gly
                125                 130                 135
Trp Gln Arg Ala Glu Ile Leu Gly Ala Leu Val Ser Val Leu Ser
                140                 145                 150
Ile Trp Val Val Thr Gly Val Leu Val Tyr Leu Ala Val Glu Arg
                155                 160                 165
Leu Ile Ser Gly Asp Tyr Glu Ile Asp Gly Gly Thr Met Leu Ile
                170                 175                 180
Thr Ser Gly Cys Ala Val Ala Val Asn Ile Ile Met Gly Leu Thr
                185                 190                 195
Leu His Gln Ser Gly His Gly His Ser His Gly Thr Thr Asn Gln
                200                 205                 210
Gln Glu Glu Asn Pro Ser Val Arg Ala Ala Phe Ile His Val Ile
                215                 220                 225
Gly Asp Phe Met Gln Ser Met Gly Val Leu Val Ala Ala Tyr Ile
                230                 235                 240
Leu Tyr Phe Lys Pro Glu Tyr Lys Tyr Val Asp Pro Ile Cys Thr
                245                 250                 255
Phe Val Phe Ser Ile Leu Val Leu Gly Thr Thr Leu Thr Ile Leu
                260                 265                 270
Arg Asp Val Ile Leu Val Leu Met Glu Gly Thr Pro Lys Gly Val
                275                 280                 285
```

```
Asp Phe Thr Ala Val Arg Asp Leu Leu Leu Ser Val Glu Gly Val
                290                 295                 300
Glu Ala Leu His Ser Leu His Ile Trp Ala Leu Thr Val Ala Gln
                305                 310                 315
Pro Val Leu Ser Val His Ile Ala Ile Ala Gln Asn Thr Asp Ala
                320                 325                 330
Gln Ala Val Leu Lys Thr Ala Ser Ser Arg Leu Gln Gly Lys Phe
                335                 340                 345
His Phe His Thr Val Thr Ile Gln Ile Glu Asp Tyr Ser Glu Asp
                350                 355                 360
Met Lys Asp Cys Gln Ala Cys Gln Gly Pro Ser Asp
                365                 370
```

<210> 5
<211> 490
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2667691CD1

<400> 5

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

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

<210> 6
<211> 377
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3211415CD1

<400> 6
Met Leu Pro Leu Ser Ile Lys Asp Asp Glu Tyr Lys Pro Pro Lys
  1               5                  10                  15
Phe Asn Leu Phe Gly Lys Ile Ser Gly Trp Phe Arg Ser Ile Leu
                20                  25                  30
Ser Asp Lys Thr Ser Arg Asn Leu Phe Phe Phe Leu Cys Leu Asn
                35                  40                  45
Leu Ser Phe Ala Phe Val Glu Leu Leu Tyr Gly Ile Trp Ser Asn
                50                  55                  60
Cys Leu Gly Leu Ile Ser Asp Ser Phe His Met Phe Phe Asp Ser
                65                  70                  75
Thr Ala Ile Leu Ala Gly Leu Ala Ala Ser Val Ile Ser Lys Trp
                80                  85                  90
```

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

<210> 7
<211> 340
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4739923CD1

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

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

<210> 8
<211> 1274
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 55030459CD1

<400> 8

```
Met Ala Arg Gln Pro Glu Glu Glu Glu Thr Ala Val Ala Arg Ala
  1               5                  10                  15
Arg Arg Pro Pro Leu Trp Leu Leu Cys Leu Val Ala Cys Trp Leu
                20                  25                  30
```

```
Leu Gly Ala Gly Ala Glu Ala Asp Phe Ser Ile Leu Asp Glu Ala
                35                  40                  45
Gln Val Leu Ala Ser Gln Met Arg Arg Leu Ala Ala Glu Glu Leu
                50                  55                  60
Gly Val Val Thr Met Gln Arg Ile Phe Asn Ser Phe Val Tyr Thr
                65                  70                  75
Glu Lys Ile Ser Asn Gly Glu Ser Glu Val Gln Gln Leu Ala Lys
                80                  85                  90
Lys Ile Arg Glu Lys Phe Asn Arg Tyr Leu Asp Val Val Asn Arg
                95                  100                 105
Asn Lys Gln Val Val Glu Ala Ser Tyr Thr Ala His Leu Thr Ser
                110                 115                 120
Pro Leu Thr Ala Ile Gln Asp Cys Cys Thr Ile Pro Pro Ser Met
                125                 130                 135
Met Glu Phe Asp Gly Asn Phe Asn Thr Asn Val Ser Arg Thr Ile
                140                 145                 150
Ser Cys Asp Arg Leu Ser Thr Thr Val Asn Ser Arg Ala Phe Asn
                155                 160                 165
Pro Gly Arg Asp Leu Asn Ser Val Leu Ala Asp Asn Leu Lys Ser
                170                 175                 180
Asn Pro Gly Ile Lys Trp Gln Tyr Phe Ser Ser Glu Glu Gly Ile
                185                 190                 195
Phe Thr Val Phe Pro Ala His Lys Phe Arg Cys Lys Gly Ser Tyr
                200                 205                 210
Glu His Arg Ser Arg Pro Ile Tyr Val Ser Thr Val Arg Pro Gln
                215                 220                 225
Ser Lys His Ile Val Val Ile Leu Asp His Gly Ala Ser Val Thr
                230                 235                 240
Asp Thr Gln Leu Gln Ile Ala Lys Asp Ala Ala Gln Val Ile Leu
                245                 250                 255
Ser Ala Ile Asp Glu His Asp Lys Ile Ser Val Leu Thr Val Ala
                260                 265                 270
Asp Thr Val Arg Thr Cys Ser Leu Asp Gln Cys Tyr Lys Thr Phe
                275                 280                 285
Leu Ser Pro Ala Thr Ser Glu Thr Lys Arg Lys Met Ser Thr Phe
                290                 295                 300
Val Ser Ser Val Lys Ser Ser Asp Ser Pro Thr Gln His Ala Val
                305                 310                 315
Gly Phe Gln Lys Ala Phe Gln Leu Ile Arg Ser Thr Asn Asn Asn
                320                 325                 330
Thr Lys Phe Gln Ala Asn Thr Asp Met Val Ile Ile Tyr Leu Ser
                335                 340                 345
Ala Gly Ile Thr Ser Lys Asp Ser Ser Glu Glu Asp Lys Lys Ala
                350                 355                 360
Thr Leu Gln Val Ile Asn Glu Glu Asn Ser Phe Leu Asn Asn Ser
                365                 370                 375
Val Met Ile Leu Thr Tyr Ala Leu Met Asn Asp Gly Val Thr Gly
                380                 385                 390
Leu Lys Glu Leu Ala Phe Leu Arg Asp Leu Ala Glu Gln Asn Ser
                395                 400                 405
Gly Lys Tyr Gly Val Pro Asp Arg Thr Ala Leu Pro Val Ile Lys
                410                 415                 420
Gly Ser Met Met Val Leu Asn Gln Leu Ser Asn Leu Glu Thr Thr
                425                 430                 435
Val Gly Arg Phe Tyr Thr Asn Leu Pro Asn Arg Met Ile Asp Glu
                440                 445                 450
```

```
Ala Val Phe Ser Leu Pro Phe Ser Asp Glu Met Gly Asp Gly Leu
                455                 460                 465
Ile Met Thr Val Ser Lys Pro Cys Tyr Phe Gly Asn Leu Leu Leu
                470                 475                 480
Gly Ile Val Gly Val Asp Val Asn Leu Ala Tyr Ile Leu Glu Asp
                485                 490                 495
Val Thr Tyr Tyr Gln Asp Ser Leu Ala Ser Tyr Thr Phe Leu Ile
                500                 505                 510
Asp Asp Lys Gly Tyr Thr Leu Met His Pro Ser Leu Thr Arg Pro
                515                 520                 525
Tyr Leu Leu Ser Glu Pro Pro Leu His Thr Asp Ile Ile His Tyr
                530                 535                 540
Glu Asn Ile Pro Lys Phe Glu Leu Val Arg Gln Asn Ile Leu Ser
                545                 550                 555
Leu Pro Leu Gly Ser Gln Ile Ile Ala Val Pro Val Asn Ser Ser
                560                 565                 570
Leu Ser Trp His Ile Asn Lys Leu Arg Glu Thr Gly Lys Glu Ala
                575                 580                 585
Tyr Asn Val Ser Tyr Ala Trp Lys Met Val Gln Asp Thr Ser Phe
                590                 595                 600
Ile Leu Cys Ile Val Val Ile Gln Pro Glu Ile Pro Val Lys Gln
                605                 610                 615
Leu Lys Asn Leu Asn Thr Val Pro Ser Ser Lys Leu Leu Tyr His
                620                 625                 630
Arg Leu Asp Leu Leu Gly Gln Pro Ser Ala Cys Leu His Phe Lys
                635                 640                 645
Gln Leu Ala Thr Leu Glu Ser Pro Thr Ile Met Leu Ser Ala Gly
                650                 655                 660
Ser Phe Ser Ser Pro Tyr Glu His Leu Ser Gln Pro Glu Thr Lys
                665                 670                 675
Arg Met Val Glu His Tyr Thr Ala Tyr Leu Ser Asp Asn Thr Arg
                680                 685                 690
Leu Ile Ala Asn Pro Gly Leu Lys Phe Ser Val Arg Asn Glu Val
                695                 700                 705
Met Ala Thr Ser His Val Thr Asp Glu Trp Met Thr Gln Met Glu
                710                 715                 720
Met Ser Ser Leu Asn Thr Tyr Ile Val Arg Arg Tyr Ile Ala Thr
                725                 730                 735
Pro Asn Gly Val Leu Arg Ile Tyr Pro Gly Ser Leu Met Asp Lys
                740                 745                 750
Ala Phe Asp Pro Thr Arg Arg Gln Trp Tyr Leu His Ala Val Ala
                755                 760                 765
Asn Pro Gly Leu Ile Ser Leu Thr Gly Pro Tyr Leu Asp Val Gly
                770                 775                 780
Gly Ala Gly Tyr Val Val Thr Ile Ser His Thr Ile His Ser Ser
                785                 790                 795
Ser Thr Gln Leu Ser Ser Gly His Thr Val Ala Val Met Gly Ile
                800                 805                 810
Asp Phe Thr Leu Arg Tyr Phe Tyr Lys Val Leu Met Asp Leu Leu
                815                 820                 825
Pro Val Cys Asn Gln Asp Gly Gly Asn Lys Ile Arg Cys Phe Ile
                830                 835                 840
Met Glu Asp Arg Gly Tyr Leu Val Ala His Pro Thr Leu Ile Asp
                845                 850                 855
Pro Lys Gly His Ala Pro Val Glu Gln Gln His Ile Thr His Lys
                860                 865                 870
```

Glu Pro Leu Val Ala Asn Asp Ile Leu Asn His Pro Asn Phe Val
875 880 885
Lys Lys Asn Leu Cys Asn Ser Phe Ser Asp Arg Thr Val Gln Arg
890 895 900
Phe Tyr Lys Phe Asn Thr Ser Leu Ala Gly Asp Leu Thr Asn Leu
905 910 915
Val His Gly Ser His Cys Ser Lys Tyr Arg Leu Ala Arg Ile Pro
920 925 930
Gly Thr Asn Ala Phe Val Gly Ile Val Asn Glu Thr Cys Asp Ser
935 940 945
Leu Ala Phe Cys Ala Cys Ser Met Val Asp Arg Leu Cys Leu Asn
950 955 960
Cys His Arg Met Glu Gln Asn Glu Cys Glu Cys Pro Cys Glu Cys
965 970 975
Pro Leu Glu Val Asn Glu Cys Thr Gly Asn Leu Thr Asn Ala Glu
980 985 990
Asn Arg Asn Pro Ser Cys Glu Val His Gln Glu Pro Val Thr Tyr
995 1000 1005
Thr Ala Ile Asp Pro Gly Leu Gln Asp Ala Leu His Gln Cys Val
1010 1015 1020
Asn Ser Arg Cys Ser Gln Arg Leu Glu Ser Gly Asp Cys Phe Gly
1025 1030 1035
Val Leu Asp Cys Glu Trp Cys Met Val Asp Ser Asp Gly Lys Thr
1040 1045 1050
His Leu Asp Lys Pro Tyr Cys Ala Pro Gln Lys Glu Cys Phe Gly
1055 1060 1065
Gly Ile Val Gly Ala Lys Ser Pro Tyr Val Asp Asp Met Gly Ala
1070 1075 1080
Ile Gly Asp Glu Val Ile Thr Leu Asn Met Ile Lys Ser Ala Pro
1085 1090 1095
Val Gly Pro Val Ala Gly Gly Ile Met Gly Cys Ile Met Val Leu
1100 1105 1110
Val Leu Ala Val Tyr Ala Tyr Arg His Gln Ile His Arg Arg Ser
1115 1120 1125
His Gln His Met Ser Pro Leu Ala Ala Gln Glu Met Ser Val Arg
1130 1135 1140
Met Ser Asn Leu Glu Asn Asp Arg Asp Glu Arg Asp Asp Asp Ser
1145 1150 1155
His Glu Asp Arg Gly Ile Ile Ser Asn Thr Arg Phe Ile Ala Ala
1160 1165 1170
Val Ile Glu Arg His Ala His Ser Pro Glu Arg Arg Arg Arg Tyr
1175 1180 1185
Trp Gly Arg Ser Gly Thr Glu Ser Asp His Gly Tyr Ser Thr Met
1190 1195 1200
Ser Pro Gln Glu Asp Ser Glu Asn Pro Pro Cys Asn Asn Asp Pro
1205 1210 1215
Leu Ser Ala Gly Val Asp Val Gly Asn His Asp Glu Asp Leu Asp
1220 1225 1230
Leu Asp Thr Pro Pro Gln Thr Ala Ala Leu Leu Ser His Lys Phe
1235 1240 1245
His His Tyr Arg Ser His His Pro Thr Leu His His Ser His His
1250 1255 1260
Leu Gln Ala Ala Val Thr Val His Thr Val Asp Ala Glu Cys
1265 1270

<210> 9

<211> 595
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6113039CD1

<400> 9
Met Lys Phe Phe Ser Tyr Ile Leu Val Tyr Arg Arg Phe Leu Phe
1 5 10 15
Val Val Phe Thr Val Leu Val Leu Leu Pro Leu Pro Ile Val Leu
20 25 30
His Thr Lys Glu Ala Glu Cys Ala Tyr Thr Leu Phe Val Val Ala
35 40 45
Thr Phe Trp Leu Thr Glu Ala Leu Pro Leu Ser Val Thr Ala Leu
50 55 60
Leu Pro Ser Leu Met Leu Pro Met Phe Gly Ile Met Pro Ser Lys
65 70 75
Lys Val Ala Ser Ala Tyr Phe Lys Asp Phe His Leu Leu Leu Ile
80 85 90
Gly Val Ile Cys Leu Ala Thr Ser Ile Glu Lys Trp Asn Leu His
95 100 105
Lys Arg Ile Ala Leu Lys Met Val Met Met Val Gly Val Asn Pro
110 115 120
Ala Trp Leu Thr Leu Gly Phe Met Ser Ser Thr Ala Phe Leu Ser
125 130 135
Met Trp Leu Ser Asn Thr Ser Thr Ala Ala Met Val Met Pro Ile
140 145 150
Ala Glu Ala Val Val Gln Gln Ile Ile Asn Ala Glu Ala Glu Val
155 160 165
Glu Ala Thr Gln Met Thr Tyr Phe Asn Gly Ser Thr Asn His Gly
170 175 180
Leu Glu Ile Asp Glu Ser Val Asn Gly His Glu Ile Asn Glu Arg
185 190 195
Lys Glu Lys Thr Lys Pro Val Pro Gly Tyr Asn Asn Asp Thr Gly
200 205 210
Lys Ile Ser Ser Lys Val Glu Leu Glu Lys Asn Ser Gly Met Arg
215 220 225
Thr Lys Tyr Arg Thr Lys Lys Gly His Val Thr Arg Lys Leu Thr
230 235 240
Cys Leu Cys Ile Ala Tyr Ser Ser Thr Ile Gly Gly Leu Thr Thr
245 250 255
Ile Thr Gly Thr Ser Thr Asn Leu Ile Phe Ala Glu Tyr Phe Asn
260 265 270
Thr Arg Tyr Pro Asp Cys Arg Cys Leu Asn Phe Gly Ser Trp Phe
275 280 285
Thr Phe Ser Phe Pro Ala Ala Leu Ile Ile Leu Leu Leu Ser Trp
290 295 300
Ile Trp Leu Gln Trp Leu Phe Leu Gly Phe Asn Phe Lys Glu Met
305 310 315
Phe Lys Cys Gly Lys Thr Lys Thr Val Gln Gln Lys Ala Cys Ala
320 325 330
Glu Val Ile Lys Gln Glu Tyr Gln Lys Leu Gly Pro Ile Arg Tyr
335 340 345
Gln Glu Ile Val Thr Leu Val Leu Phe Ile Ile Met Ala Leu Leu

```
                350                 355                 360
Trp Phe Ser Arg Asp Pro Gly Phe Val Pro Gly Trp Ser Ala Leu
                365                 370                 375
Phe Ser Glu Tyr Pro Gly Phe Ala Thr Asp Ser Thr Val Ala Leu
                380                 385                 390
Leu Ile Gly Leu Leu Phe Phe Leu Ile Pro Ala Lys Thr Leu Thr
                395                 400                 405
Lys Thr Thr Pro Thr Gly Glu Ile Val Ala Phe Asp Tyr Ser Pro
                410                 415                 420
Leu Ile Thr Trp Lys Glu Phe Gln Ser Phe Met Pro Trp Asp Ile
                425                 430                 435
Ala Ile Leu Val Gly Gly Gly Phe Ala Leu Ala Asp Gly Cys Glu
                440                 445                 450
Glu Ser Gly Leu Ser Lys Trp Ile Gly Asn Lys Leu Ser Pro Leu
                455                 460                 465
Gly Ser Leu Pro Ala Trp Leu Ile Ile Leu Ile Ser Ser Leu Met
                470                 475                 480
Val Thr Ser Leu Thr Glu Val Ala Ser Asn Pro Ala Thr Ile Thr
                485                 490                 495
Leu Phe Leu Pro Ile Leu Ser Pro Leu Ala Glu Ala Ile His Val
                500                 505                 510
Asn Pro Leu Tyr Ile Leu Ile Pro Ser Thr Leu Cys Thr Ser Phe
                515                 520                 525
Ala Phe Leu Leu Pro Val Ala Asn Pro Pro Asn Ala Ile Val Phe
                530                 535                 540
Ser Tyr Gly His Leu Lys Val Ile Asp Met Val Lys Ala Gly Leu
                545                 550                 555
Gly Val Asn Ile Val Gly Val Ala Val Val Met Leu Gly Ile Cys
                560                 565                 570
Thr Trp Ile Val Pro Met Phe Asp Leu Tyr Thr Tyr Pro Ser Trp
                575                 580                 585
Ala Pro Ala Met Ser Asn Glu Thr Met Pro
                590                 595

<210> 10
<211> 475
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7101781CD1

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

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

<210> 11
<211> 927
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7473036CD1

<400> 11

```
Met Gln Pro Ala Arg Gly Pro Leu Ala Ser Glu Pro Arg Thr Val
  1               5                  10                  15
Leu Val Leu Arg Phe Cys Ala Ser Leu Met Glu Met Lys Leu Pro
             20                  25                  30
Gly Gln Glu Gly Phe Glu Ala Ser Ser Ala Pro Arg Asn Ile Pro
         35                  40                  45
Ser Gly Glu Leu Asp Ser Asn Pro Asp Pro Gly Thr Gly Pro Ser
     50                  55                  60
Pro Asp Gly Pro Ser Asp Thr Glu Ser Lys Glu Leu Gly Val Pro
 65                  70                  75
Lys Asp Pro Leu Leu Phe Ile Gln Leu Asn Glu Leu Leu Gly Trp
             80                  85                  90
Pro Gln Ala Leu Glu Trp Arg Glu Thr Gly Arg Trp Val Leu Phe
         95                 100                 105
Glu Glu Lys Leu Glu Val Ala Ala Gly Arg Trp Ser Ala Pro His
        110                 115                 120
Val Pro Thr Leu Ala Leu Pro Ser Leu Gln Lys Leu Arg Ser Leu
    125                 130                 135
Leu Ala Glu Gly Leu Val Leu Leu Asp Cys Pro Ala Gln Ser Leu
    140                 145                 150
Leu Glu Leu Val Gly Ser Thr His Pro Arg Lys Ala Ser Asp Asn
    155                 160                 165
Glu Glu Ala Pro Leu Arg Glu Gln Cys Gln Asn Pro Leu Arg Gln
    170                 175                 180
Lys Leu Pro Pro Gly Ala Glu Ala Gly Thr Val Leu Ala Gly Glu
    185                 190                 195
Leu Gly Phe Leu Ala Gln Pro Leu Gly Ala Phe Val Arg Leu Arg
    200                 205                 210
Asn Pro Val Val Leu Gly Ser Leu Thr Glu Val Ser Leu Pro Ser
    215                 220                 225
Arg Phe Phe Cys Leu Leu Leu Gly Pro Cys Met Leu Gly Lys Gly
    230                 235                 240
Tyr His Glu Met Gly Arg Ala Ala Ala Val Leu Leu Ser Asp Pro
    245                 250                 255
Gln Phe Gln Trp Ser Val Arg Arg Ala Ser Asn Leu His Asp Leu
    260                 265                 270
Leu Ala Ala Leu Asp Ala Phe Leu Glu Glu Val Thr Val Leu Pro
    275                 280                 285
Pro Gly Arg Trp Asp Pro Thr Ala Arg Ile Pro Pro Pro Lys Cys
    290                 295                 300
Leu Pro Ser Gln His Lys Arg Leu Pro Ser Gln Gln Arg Glu Ile
    305                 310                 315
Arg Gly Pro Ala Val Pro Arg Leu Thr Ser Ala Glu Asp Arg His
    320                 325                 330
Arg His Gly Pro His Ala His Ser Pro Glu Leu Gln Arg Thr Gly
    335                 340                 345
Ser Asp Phe Leu Asp Ala Leu His Leu Gln Cys Phe Ser Ala Val
    350                 355                 360
Leu Tyr Ile Tyr Leu Ala Thr Val Thr Asn Ala Ile Thr Phe Gly
```

```
                365                 370                 375
Gly Leu Leu Gly Asp Ala Thr Asp Gly Ala Gln Gly Val Leu Glu
                380                 385                 390
Ser Phe Leu Gly Thr Ala Val Ala Gly Ala Ala Phe Cys Leu Met
                395                 400                 405
Ala Gly Gln Pro Leu Thr Ile Leu Ser Ser Thr Gly Pro Val Leu
                410                 415                 420
Val Phe Glu Arg Leu Leu Phe Ser Phe Ser Arg Asp Tyr Ser Leu
                425                 430                 435
Asp Tyr Leu Pro Phe Arg Leu Trp Val Gly Ile Trp Val Ala Thr
                440                 445                 450
Phe Cys Leu Val Leu Val Ala Thr Glu Ala Ser Val Leu Val Arg
                455                 460                 465
Tyr Phe Thr Arg Phe Thr Glu Glu Gly Phe Cys Ala Leu Ile Ser
                470                 475                 480
Leu Ile Phe Ile Tyr Asp Ala Val Gly Lys Met Leu Asn Leu Thr
                485                 490                 495
His Thr Tyr Pro Ile Gln Lys Pro Gly Ser Ser Ala Tyr Gly Cys
                500                 505                 510
Leu Cys Gln Tyr Pro Gly Pro Gly Gly Asn Glu Ser Gln Trp Ile
                515                 520                 525
Arg Thr Arg Pro Lys Asp Arg Asp Asp Ile Val Ser Met Asp Leu
                530                 535                 540
Gly Leu Ile Asn Ala Ser Leu Leu Pro Pro Pro Glu Cys Thr Arg
                545                 550                 555
Gln Gly Gly His Pro Arg Gly Pro Gly Cys His Thr Val Pro Asp
                560                 565                 570
Ile Ala Phe Phe Ser Leu Leu Leu Phe Leu Thr Ser Phe Phe Phe
                575                 580                 585
Ala Met Ala Leu Lys Cys Val Lys Thr Ser Arg Phe Phe Pro Ser
                590                 595                 600
Val Val Arg Lys Gly Leu Ser Asp Phe Ser Ser Val Leu Ala Ile
                605                 610                 615
Leu Leu Gly Cys Gly Leu Asp Ala Phe Leu Gly Leu Ala Thr Pro
                620                 625                 630
Lys Leu Met Val Pro Arg Glu Phe Lys Pro Thr Leu Pro Gly Arg
                635                 640                 645
Gly Trp Leu Val Ser Pro Phe Gly Ala Asn Pro Trp Trp Trp Ser
                650                 655                 660
Val Ala Ala Ala Leu Pro Ala Leu Leu Leu Ser Ile Leu Ile Phe
                665                 670                 675
Met Asp Gln Gln Ile Thr Ala Val Ile Leu Asn Arg Met Glu Tyr
                680                 685                 690
Arg Leu Gln Lys Gly Ala Gly Phe His Leu Asp Leu Phe Cys Val
                695                 700                 705
Ala Val Leu Met Leu Leu Thr Ser Ala Leu Gly Leu Pro Trp Tyr
                710                 715                 720
Val Ser Ala Thr Val Ile Ser Leu Ala His Met Asp Ser Leu Arg
                725                 730                 735
Arg Glu Ser Arg Ala Cys Ala Pro Gly Glu Arg Pro Asn Phe Leu
                740                 745                 750
Gly Ile Arg Glu Gln Arg Leu Thr Gly Leu Val Val Phe Ile Leu
                755                 760                 765
Thr Gly Ala Ser Ile Phe Leu Ala Pro Val Leu Lys Phe Ile Pro
                770                 775                 780
Met Pro Val Leu Tyr Gly Ile Phe Leu Tyr Met Gly Val Ala Ala
```

```
                785                 790                 795
Leu Ser Ser Ile Gln Phe Thr Asn Arg Val Lys Leu Leu Leu Met
                800                 805                 810
Pro Ala Lys His Gln Pro Asp Leu Leu Leu Leu Arg His Val Pro
                815                 820                 825
Leu Thr Arg Val His Leu Phe Thr Ala Ile Gln Leu Ala Cys Leu
                830                 835                 840
Gly Leu Leu Trp Ile Ile Lys Ser Thr Pro Ala Ala Ile Ile Phe
                845                 850                 855
Pro Leu Met Leu Leu Gly Leu Val Gly Val Arg Lys Ala Leu Glu
                860                 865                 870
Arg Val Phe Ser Pro Gln Glu Leu Leu Trp Leu Asp Glu Leu Met
                875                 880                 885
Pro Glu Glu Glu Arg Ser Ile Pro Glu Lys Gly Leu Glu Pro Glu
                890                 895                 900
His Ser Phe Ser Gly Ser Asp Ser Glu Asp Ser Glu Leu Met Tyr
                905                 910                 915
Gln Pro Lys Ala Pro Glu Ile Asn Ile Ser Val Asn
                920                 925

<210> 12
<211> 516
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7476943CD1

<400> 12
Met Pro Ser Gly Ser His Trp Thr Ala Asn Ser Ser Lys Ile Ile
  1                 5                  10                  15
Thr Trp Leu Leu Glu Gln Pro Gly Lys Glu Glu Lys Arg Lys Thr
                 20                  25                  30
Met Ala Lys Val Asn Arg Ala Arg Ser Thr Ser Pro Pro Asp Gly
                 35                  40                  45
Gly Trp Gly Trp Met Ile Val Ala Gly Cys Phe Leu Val Thr Ile
                 50                  55                  60
Cys Thr Arg Ala Val Thr Arg Cys Ile Ser Ile Phe Phe Val Glu
                 65                  70                  75
Phe Gln Thr Tyr Phe Thr Gln Asp Tyr Ala Gln Thr Ala Trp Ile
                 80                  85                  90
His Ser Ile Val Asp Cys Val Thr Met Leu Cys Ala Pro Leu Gly
                 95                 100                 105
Ser Val Val Ser Asn His Leu Ser Cys Gln Val Gly Ile Met Leu
                110                 115                 120
Gly Gly Leu Leu Ala Ser Thr Gly Leu Ile Leu Ser Ser Phe Ala
                125                 130                 135
Thr Ser Leu Lys His Leu Tyr Leu Thr Leu Gly Val Leu Thr Gly
                140                 145                 150
Leu Gly Phe Ala Leu Cys Tyr Ser Pro Ala Ile Ala Met Val Gly
                155                 160                 165
Lys Tyr Phe Ser Arg Arg Lys Ala Leu Ala Tyr Gly Ile Ala Met
                170                 175                 180
Ser Gly Ser Gly Ile Gly Thr Phe Ile Leu Ala Pro Val Val Gln
                185                 190                 195
```

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

<210> 13
<211> 514
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8003355CD1

<400> 13
Met His Gly Gly Gln Gly Pro Leu Leu Leu Leu Leu Leu Leu Ala
```

```
  1               5                  10                  15
Val Cys Leu Gly Ala Gln Gly Arg Asn Gln Glu Glu Arg Leu Leu
                20                  25                  30
Ala Asp Leu Met Gln Asn Tyr Asp Pro Asn Leu Arg Pro Ala Glu
                35                  40                  45
Arg Asp Ser Asp Val Val Asn Val Ser Leu Lys Leu Thr Leu Thr
                50                  55                  60
Asn Leu Ile Ser Leu Asn Glu Arg Glu Glu Ala Leu Thr Thr Asn
                65                  70                  75
Val Trp Ile Glu Val Gln Trp Cys Asp Tyr Arg Leu Arg Arg Asp
                80                  85                  90
Pro Arg Asp Tyr Glu Gly Leu Trp Val Leu Arg Val Pro Ser Thr
                95                  100                 105
Met Val Trp Arg Pro Asp Ile Val Leu Glu Asn Asn Ala Asp Gly
                110                 115                 120
Val Phe Glu Val Ala Leu Tyr Cys Asn Val Leu Val Ser Pro Asp
                125                 130                 135
Gly Cys Ile Tyr Trp Leu Pro Pro Ala Ile Phe Arg Ser Ala Cys
                140                 145                 150
Ser Ile Ser Val Thr Tyr Phe Pro Phe Asp Trp Gln Asn Cys Ser
                155                 160                 165
Leu Ile Phe Gln Ser Gln Thr Tyr Ser Thr Asn Glu Ile Asp Leu
                170                 175                 180
Gln Leu Ser Gln Glu Asp Gly Gln Thr Ile Glu Trp Ile Phe Ile
                185                 190                 195
Asp Pro Glu Ala Phe Thr Glu Asn Gly Glu Trp Ala Ile Gln His
                200                 205                 210
Arg Pro Ala Lys Met Leu Leu Asp Pro Ala Ala Pro Ala Gln Glu
                215                 220                 225
Ala Gly His Gln Lys Val Val Phe Tyr Leu Leu Ile Gln Arg Lys
                230                 235                 240
Pro Leu Phe Tyr Val Ile Asn Ile Ile Ala Pro Cys Val Leu Ile
                245                 250                 255
Ser Ser Val Ala Ile Leu Ile His Phe Leu Pro Ala Lys Ala Gly
                260                 265                 270
Gly Gln Lys Cys Thr Val Ala Ile Asn Val Leu Leu Ala Gln Thr
                275                 280                 285
Val Phe Leu Phe Leu Val Ala Lys Lys Val Pro Glu Thr Ser Gln
                290                 295                 300
Ala Val Pro Leu Ile Ser Lys Tyr Leu Thr Phe Leu Leu Val Val
                305                 310                 315
Thr Ile Leu Ile Val Val Asn Ala Val Val Val Leu Asn Val Ser
                320                 325                 330
Leu Arg Ser Pro His Thr His Ser Met Ala Arg Gly Val Phe Leu
                335                 340                 345
Arg Leu Leu Pro Gln Leu Leu Arg Met His Val Arg Pro Leu Ala
                350                 355                 360
Pro Ala Ala Val Gln Asp Thr Gln Ser Arg Leu Gln Asn Gly Ser
                365                 370                 375
Ser Gly Trp Ser Ile Thr Thr Gly Glu Glu Val Ala Leu Cys Leu
                380                 385                 390
Pro Arg Ser Glu Leu Leu Phe Gln Gln Trp Gln Arg Gln Gly Leu
                395                 400                 405
Val Ala Ala Ala Leu Glu Lys Leu Glu Lys Gly Pro Glu Leu Gly
                410                 415                 420
Leu Ser Gln Phe Cys Gly Ser Leu Lys Gln Ala Ala Pro Ala Ile
```

```
                425                 430                 435
Gln Ala Cys Val Glu Ala Cys Asn Leu Ile Ala Cys Ala Arg His
                440                 445                 450
Gln Gln Ser His Phe Asp Asn Gly Asn Glu Glu Trp Phe Leu Val
                455                 460                 465
Gly Arg Val Leu Asp Arg Val Cys Phe Leu Ala Met Leu Ser Leu
                470                 475                 480
Phe Ile Cys Gly Thr Ala Gly Ile Phe Leu Met Ala His Tyr Asn
                485                 490                 495
Arg Val Pro Ala Leu Pro Phe Pro Gly Asp Pro Arg Pro Tyr Leu
                500                 505                 510
Pro Ser Pro Asp


<210> 14
<211> 691
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3116448CD1

<400> 14
Met Glu Leu Arg Ser Thr Ala Ala Pro Arg Ala Glu Gly Tyr Ser
  1                 5                  10                  15
Asn Val Gly Phe Gln Asn Glu Glu Asn Phe Leu Glu Asn Glu Asn
                 20                  25                  30
Thr Ser Gly Asn Asn Ser Ile Arg Ser Arg Ala Val Gln Ser Arg
                 35                  40                  45
Glu His Thr Asn Thr Lys Gln Asp Glu Glu Gln Val Thr Val Glu
                 50                  55                  60
Gln Asp Ser Pro Arg Asn Arg Glu His Met Glu Asp Asp Asp Glu
                 65                  70                  75
Glu Met Gln Gln Lys Gly Cys Leu Glu Arg Arg Tyr Asp Thr Val
                 80                  85                  90
Cys Gly Phe Cys Arg Lys His Lys Thr Thr Leu Arg His Ile Ile
                 95                 100                 105
Trp Gly Ile Leu Leu Ala Gly Tyr Leu Val Met Val Ile Ser Ala
                110                 115                 120
Cys Val Leu Asn Phe His Arg Ala Leu Pro Leu Phe Val Ile Thr
                125                 130                 135
Val Ala Ala Ile Phe Phe Val Val Trp Asp His Leu Met Ala Lys
                140                 145                 150
Tyr Glu His Arg Ile Asp Glu Met Leu Ser Pro Gly Arg Arg Leu
                155                 160                 165
Leu Asn Ser His Trp Phe Trp Leu Lys Trp Val Ile Trp Ser Ser
                170                 175                 180
Leu Val Leu Ala Val Ile Phe Trp Leu Ala Phe Asp Thr Ala Lys
                185                 190                 195
Leu Gly Gln Gln Gln Leu Val Ser Phe Gly Gly Leu Ile Met Tyr
                200                 205                 210
Ile Val Leu Leu Phe Leu Phe Ser Lys Tyr Pro Thr Arg Val Tyr
                215                 220                 225
Trp Arg Pro Val Leu Trp Gly Ile Gly Leu Gln Phe Leu Leu Gly
                230                 235                 240
```

```
Leu Leu Ile Leu Arg Thr Asp Pro Gly Phe Ile Ala Phe Asp Trp
                245                 250                 255
Leu Gly Arg Gln Val Gln Thr Phe Leu Glu Tyr Thr Asp Ala Gly
                260                 265                 270
Ala Ser Phe Gly Phe Gly Glu Lys Tyr Lys Asp His Phe Phe Gly
                275                 280                 285
Phe Lys Val Leu Ala Ile Val Val Phe Phe Ser Thr Val Met Ser
                290                 295                 300
Met Leu Tyr Tyr Leu Gly Leu Met Gln Trp Ile Ile Arg Lys Val
                305                 310                 315
Gly Trp Ile Met Leu Val Thr Thr Gly Ser Ser Pro Ile Glu Ser
                320                 325                 330
Val Val Ala Ser Gly Asn Ile Phe Val Gly Gln Thr Glu Ser Pro
                335                 340                 345
Leu Leu Val Arg Pro Tyr Leu Pro Tyr Ile Thr Lys Ser Glu Leu
                350                 355                 360
His Ala Ile Met Thr Ala Gly Phe Ser Thr Ile Ala Gly Ser Val
                365                 370                 375
Leu Gly Ala Tyr Ile Ser Phe Gly Val Pro Ser Ser His Leu Leu
                380                 385                 390
Thr Ala Ser Val Met Ser Ala Pro Ala Ser Leu Ala Ala Ala Lys
                395                 400                 405
Leu Phe Trp Pro Glu Thr Glu Lys Pro Lys Ile Thr Leu Lys Asn
                410                 415                 420
Ala Met Lys Met Glu Ser Gly Asp Ser Gly Asn Leu Leu Glu Ala
                425                 430                 435
Ala Thr Gln Gly Ala Ser Ser Ser Ile Ser Leu Val Ala Asn Ile
                440                 445                 450
Ala Val Asn Leu Ile Ala Phe Leu Ala Leu Leu Ser Phe Met Asn
                455                 460                 465
Ser Ala Leu Ser Trp Phe Gly Asn Met Phe Asp Tyr Pro Gln Leu
                470                 475                 480
Ser Phe Glu Leu Ile Cys Ser Tyr Ile Phe Met Pro Phe Ser Phe
                485                 490                 495
Met Met Gly Val Glu Trp Gln Asp Ser Phe Met Val Ala Arg Leu
                500                 505                 510
Ile Gly Tyr Lys Thr Phe Phe Asn Glu Phe Val Ala Tyr Glu His
                515                 520                 525
Leu Ser Lys Trp Ile His Leu Arg Lys Glu Gly Gly Pro Lys Phe
                530                 535                 540
Val Asn Gly Val Gln Gln Tyr Ile Ser Ile Arg Ser Glu Ile Ile
                545                 550                 555
Ala Thr Tyr Ala Leu Cys Gly Phe Ala Asn Ile Gly Ser Leu Gly
                560                 565                 570
Ile Val Ile Gly Gly Leu Thr Ser Met Ala Pro Ser Arg Lys Arg
                575                 580                 585
Asp Ile Ala Ser Gly Ala Val Arg Ala Leu Ile Ala Gly Thr Val
                590                 595                 600
Ala Cys Phe Met Thr Ala Cys Ile Ala Gly Ile Leu Ser Ser Thr
                605                 610                 615
Pro Val Asp Ile Asn Cys His His Val Leu Glu Asn Ala Phe Asn
                620                 625                 630
Ser Thr Phe Pro Gly Asn Thr Thr Lys Val Ile Ala Cys Cys Gln
                635                 640                 645
Ser Leu Leu Ser Ser Thr Val Ala Lys Gly Pro Gly Glu Val Ile
                650                 655                 660
```

Pro Gly Gly Asn His Ser Leu Tyr Ser Leu Lys Gly Cys Cys Thr
665 670 675
Leu Leu Asn Pro Ser Thr Phe Asn Cys Asn Gly Ile Ser Asn Thr
680 685 690
Phe

<210> 15
<211> 342
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 622868CD1

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

290 295 300
Trp Glu Glu Ala Ala Lys Leu Glu Ser Glu Asn Glu Ser Lys Ser
305 310 315
Ser Lys Leu Leu Leu Thr Thr Asn Asn Ser Gly Leu Glu Lys Thr
320 325 330
Glu Ala Ile Thr Pro Arg Asp Ser Gly Leu Gly Glu
335 340

<210> 16
<211> 791
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7476494CD1

<400> 16
Met Gly His Phe Glu Lys Gly Gln His Ala Leu Leu Asn Glu Gly
1 5 10 15
Glu Glu Asn Glu Met Glu Ile Phe Gly Tyr Arg Thr Gln Gly Cys
20 25 30
Arg Lys Ser Leu Cys Leu Ala Gly Ser Ile Phe Ser Phe Gly Ile
35 40 45
Leu Pro Leu Val Phe Tyr Trp Arg Pro Ala Trp His Val Trp Ala
50 55 60
His Cys Val Pro Cys Ser Leu Gln Glu Ala Asp Thr Val Leu Leu
65 70 75
Arg Thr Thr Val Arg Cys Ile Lys Val Gln Lys Ile Arg Tyr Val
80 85 90
Trp Asn Tyr Leu Glu Gly Gln Phe Gln Lys Ile Gly Ser Leu Glu
95 100 105
Asp Trp Leu Ser Ser Ala Lys Ile His Gln Lys Phe Gly Ser Gly
110 115 120
Leu Thr Arg Glu Glu Gln Glu Ile Arg Arg Leu Met Cys Gly Pro
125 130 135
Asn Thr Ile Asp Val Glu Val Thr Pro Ile Trp Lys Leu Leu Ile
140 145 150
Lys Glu Val Leu Asn Pro Phe Tyr Ile Phe Gln Leu Phe Ser Val
155 160 165
Cys Leu Trp Phe Ser Glu Asp Tyr Lys Glu Tyr Ala Phe Ala Ile
170 175 180
Ile Ile Met Ser Ile Ile Ser Ile Ser Leu Thr Val Tyr Asp Leu
185 190 195
Arg Glu Gln Ser Val Lys Leu His His Leu Val Glu Ser His Asn
200 205 210
Ser Ile Thr Val Ser Val Cys Gly Arg Lys Ala Gly Val Gln Glu
215 220 225
Leu Glu Ser Arg Val Leu Val Pro Gly Asp Leu Leu Ile Leu Thr
230 235 240
Gly Asn Lys Val Leu Met Pro Cys Asp Ala Val Leu Ile Glu Gly
245 250 255
Ser Cys Val Val Asp Glu Gly Met Leu Thr Gly Glu Ser Ile Pro
260 265 270
Val Thr Lys Thr Pro Leu Pro Lys Met Asp Ser Ser Val Pro Trp
275 280 285

```
Lys Thr Gln Ser Glu Ala Asp Tyr Lys Arg His Val Leu Phe Cys
                290                 295                 300
Gly Thr Glu Val Ile Gln Ala Lys Ala Ala Cys Ser Gly Thr Val
                305                 310                 315
Arg Ala Val Val Leu Gln Thr Gly Phe Asn Thr Ala Lys Gly Asp
                320                 325                 330
Leu Val Arg Ser Ile Leu Tyr Pro Lys Pro Val Asn Phe Gln Leu
                335                 340                 345
Tyr Arg Asp Ala Ile Arg Phe Leu Leu Cys Leu Val Gly Thr Ala
                350                 355                 360
Thr Ile Gly Met Ile Tyr Thr Leu Cys Val Tyr Val Leu Ser Gly
                365                 370                 375
Glu Pro Pro Glu Glu Val Val Arg Lys Ala Leu Asp Val Ile Thr
                380                 385                 390
Ile Ala Val Pro Pro Ala Leu Pro Ala Ala Leu Thr Thr Gly Ile
                395                 400                 405
Ile Tyr Ala Gln Arg Arg Leu Lys Lys Arg Gly Ile Phe Cys Ile
                410                 415                 420
Ser Pro Gln Arg Ile Asn Val Cys Gly Gln Leu Asn Leu Val Cys
                425                 430                 435
Phe Asp Lys Thr Gly Thr Leu Thr Arg Asp Gly Leu Asp Leu Trp
                440                 445                 450
Gly Val Val Ser Cys Asp Arg Asn Gly Phe Gln Glu Val His Ser
                455                 460                 465
Phe Ala Ser Gly Gln Ala Leu Pro Trp Gly Pro Leu Cys Ala Ala
                470                 475                 480
Met Ala Ser Cys His Ser Leu Ile Leu Leu Asp Gly Thr Ile Gln
                485                 490                 495
Gly Asp Pro Leu Asp Leu Lys Met Phe Glu Ala Thr Thr Trp Glu
                500                 505                 510
Met Ala Phe Ser Gly Asp Asp Phe His Ile Lys Gly Val Pro Ala
                515                 520                 525
His Ala Met Val Val Lys Pro Cys Arg Thr Ala Ser Gln Val Pro
                530                 535                 540
Val Glu Gly Ile Ala Ile Leu His Gln Phe Pro Phe Ser Ser Ala
                545                 550                 555
Leu Gln Arg Met Thr Val Ile Val Gln Glu Met Gly Gly Asp Arg
                560                 565                 570
Leu Ala Phe Met Lys Gly Ala Pro Glu Arg Val Ala Ser Phe Cys
                575                 580                 585
Gln Pro Glu Thr Val Pro Thr Ser Phe Val Ser Glu Leu Gln Ile
                590                 595                 600
Tyr Thr Thr Gln Gly Phe Arg Val Ile Ala Leu Ala Tyr Lys Lys
                605                 610                 615
Leu Glu Asn Asp His His Ala Thr Thr Leu Thr Arg Glu Thr Val
                620                 625                 630
Glu Ser Asp Leu Ile Phe Leu Gly Leu Leu Ile Leu Glu Asn Arg
                635                 640                 645
Leu Lys Glu Glu Thr Lys Pro Val Leu Glu Glu Leu Ile Ser Ala
                650                 655                 660
Arg Ile Arg Thr Val Met Ile Thr Gly Asp Asn Leu Gln Thr Ala
                665                 670                 675
Ile Thr Val Ala Arg Lys Ser Gly Met Val Ser Glu Ser Gln Lys
                680                 685                 690
Val Ile Leu Ile Glu Ala Asn Glu Thr Thr Gly Ser Ser Ser Ala
                695                 700                 705
```

```
Ser Ile Ser Trp Thr Leu Val Glu Glu Lys Lys His Ile Met Tyr
                710                 715                 720
Gly Asn Gln Asp Asn Tyr Ile Asn Ile Arg Asp Glu Val Ser Asp
                725                 730                 735
Lys Gly Arg Glu Gly Ser Tyr His Phe Ala Leu Thr Gly Lys Ser
                740                 745                 750
Phe His Val Ile Ser Gln His Phe Ser Ser Leu Leu Pro Lys Ile
                755                 760                 765
Leu Ile Asn Gly Thr Ile Phe Ala Arg Met Ser Pro Gly Gln Lys
                770                 775                 780
Ser Ser Leu Val Glu Glu Phe Gln Lys Leu Glu
                785                 790

<210> 17
<211> 1108
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477260CD1

<400> 17
Met Val Thr Gly Gly Gln His His Pro Gly Ala Gly Leu Ser Phe
  1                 5                  10                  15
Thr Glu Leu Glu Asn Thr Phe Pro Leu Cys Leu Pro Pro Thr Pro
                 20                  25                  30
Phe Leu Leu Ala Leu Trp Ser Ser Cys Leu Pro Trp Asp Thr Gln
                 35                  40                  45
Gln Thr Cys Cys Pro Ser Phe Ala Gly Ser Pro Ala Ala Glu Gln
                 50                  55                  60
Leu Gln Asp Ile Leu Gly Glu Glu Asp Glu Ala Pro Asn Pro Thr
                 65                  70                  75
Leu Phe Thr Glu Met Asp Thr Leu Gln His Asp Gly Asp Gln Met
                 80                  85                  90
Glu Trp Lys Glu Ser Ala Arg Trp Ile Lys Phe Glu Glu Lys Val
                 95                 100                 105
Glu Glu Gly Gly Glu Arg Trp Ser Lys Pro His Val Ser Thr Leu
                110                 115                 120
Ser Leu His Ser Leu Phe Glu Leu Arg Thr Cys Leu Gln Thr Gly
                125                 130                 135
Thr Val Leu Leu Asp Leu Asp Ser Gly Ser Leu Pro Gln Ile Ile
                140                 145                 150
Asp Asp Val Ile Glu Lys Gln Ile Glu Asp Gly Leu Leu Arg Pro
                155                 160                 165
Glu Leu Arg Glu Arg Val Ser Tyr Val Leu Leu Arg Arg His Arg
                170                 175                 180
His Gln Thr Lys Lys Pro Ile His Arg Ser Leu Ala Asp Ile Gly
                185                 190                 195
Lys Ser Val Ser Thr Thr Asn Arg Ser Pro Ala Arg Ser Pro Gly
                200                 205                 210
Ala Gly Pro Ser Leu His His Ser Thr Glu Asp Leu Arg Met Arg
                215                 220                 225
Gln Ser Ala Asn Tyr Gly Arg Leu Cys His Ala Gln Ser Arg Ser
                230                 235                 240
Met Asn Asp Ile Ser Leu Thr Pro Asn Thr Asp Gln Arg Lys Asn
```

```
                245                 250                 255
Lys Phe Met Lys Lys Ile Pro Lys Asp Ser Glu Ala Ser Asn Val
                260                 265                 270
Leu Val Gly Glu Val Asp Phe Leu Asp Gln Pro Phe Ile Ala Phe
                275                 280                 285
Val Arg Leu Ile Gln Ser Ala Met Leu Gly Gly Val Thr Glu Val
                290                 295                 300
Pro Val Pro Thr Arg Phe Leu Phe Ile Leu Leu Gly Pro Ser Gly
                305                 310                 315
Arg Ala Lys Ser Tyr Asn Glu Ile Gly Arg Ala Ile Ala Thr Leu
                320                 325                 330
Met Val Asp Asp Leu Phe Ser Asp Val Ala Tyr Lys Ala Arg Asn
                335                 340                 345
Arg Glu Asp Leu Ile Ala Gly Ile Asp Glu Phe Leu Asp Glu Val
                350                 355                 360
Ile Val Leu Pro Pro Gly Glu Trp Asp Pro Asn Ile Arg Ile Glu
                365                 370                 375
Pro Pro Lys Lys Val Pro Ser Ala Asp Lys Arg Lys Ser Leu Phe
                380                 385                 390
Ser Leu Ala Glu Leu Gly Gln Met Asn Gly Ser Val Gly Gly Gly
                395                 400                 405
Gly Gly Ala Pro Gly Gly Gly Asn Gly Gly Gly Gly Gly Gly Gly
                410                 415                 420
Ser Gly Gly Gly Ala Gly Ser Gly Gly Ala Gly Gly Thr Ser Ser
                425                 430                 435
Gly Asp Asp Gly Glu Met Pro Ala Met His Glu Ile Gly Glu Glu
                440                 445                 450
Leu Ile Trp Thr Gly Arg Phe Phe Gly Gly Leu Cys Leu Asp Ile
                455                 460                 465
Lys Arg Lys Leu Pro Trp Phe Pro Ser Asp Phe Tyr Asp Gly Phe
                470                 475                 480
His Ile Gln Ser Ile Ser Ala Ile Leu Phe Ile Tyr Leu Gly Cys
                485                 490                 495
Ile Thr Asn Ala Ile Thr Phe Gly Gly Leu Leu Gly Asp Ala Thr
                500                 505                 510
Asp Asn Tyr Gln Gly Val Met Glu Ser Phe Leu Gly Thr Ala Met
                515                 520                 525
Ala Gly Ser Leu Phe Cys Leu Phe Ser Gly Gln Pro Leu Ile Ile
                530                 535                 540
Leu Ser Ser Thr Gly Pro Ile Leu Ile Phe Glu Lys Leu Leu Phe
                545                 550                 555
Asp Phe Ser Lys Gly Asn Gly Leu Asp Tyr Met Glu Phe Arg Leu
                560                 565                 570
Trp Ile Gly Leu His Ser Ala Val Gln Cys Leu Ile Leu Val Ala
                575                 580                 585
Thr Asp Ala Ser Phe Ile Ile Lys Tyr Ile Thr Arg Phe Thr Glu
                590                 595                 600
Glu Gly Phe Ser Thr Leu Ile Ser Phe Ile Phe Ile Tyr Asp Ala
                605                 610                 615
Ile Lys Lys Met Ile Gly Ala Phe Lys Tyr Tyr Pro Ile Asn Met
                620                 625                 630
Asp Phe Lys Pro Asn Phe Ile Thr Thr Tyr Lys Cys Glu Cys Val
                635                 640                 645
Ala Pro Asp Thr Gly Asp Leu Asn Thr Thr Val Phe Asn Ala Ser
                650                 655                 660
Ala Pro Leu Ala Pro Asp Thr Asn Ala Ser Leu Tyr Asn Leu Leu
```

665 670 675
Asn Leu Thr Ala Leu Asp Trp Ser Leu Leu Ser Lys Lys Glu Cys
680 685 690
Leu Ser Tyr Gly Gly Arg Leu Leu Gly Asn Ser Cys Lys Phe Ile
695 700 705
Pro Asp Leu Ala Leu Met Ser Phe Ile Leu Phe Phe Gly Thr Tyr
710 715 720
Ser Met Thr Leu Thr Leu Lys Lys Phe Lys Phe Ser Arg Tyr Phe
725 730 735
Pro Thr Lys Val Arg Ala Leu Val Ala Asp Phe Ser Ile Val Phe
740 745 750
Ser Ile Leu Met Phe Cys Gly Ile Asp Ala Cys Phe Gly Leu Glu
755 760 765
Thr Pro Lys Leu His Val Pro Ser Val Ile Lys Pro Thr Arg Pro
770 775 780
Asp Arg Gly Trp Phe Val Ala Pro Phe Gly Lys Asn Pro Trp Trp
785 790 795
Val Tyr Pro Ala Ser Ile Leu Pro Ala Leu Leu Val Thr Ile Leu
800 805 810
Ile Phe Met Asp Gln Gln Ile Thr Ala Val Ile Val Asn Arg Lys
815 820 825
Glu Asn Lys Leu Lys Lys Ala Ala Gly Tyr His Leu Asp Leu Phe
830 835 840
Trp Val Gly Ile Leu Met Ala Leu Cys Ser Phe Met Gly Leu Pro
845 850 855
Trp Tyr Val Ala Ala Thr Val Ile Ser Ile Ala His Ile Asp Ser
860 865 870
Leu Lys Met Glu Thr Glu Thr Ser Ala Pro Gly Glu Gln Pro Gln
875 880 885
Phe Leu Gly Val Arg Glu Gln Arg Val Thr Gly Ile Ile Val Phe
890 895 900
Ile Leu Thr Gly Ile Ser Val Phe Leu Ala Pro Ile Leu Lys Cys
905 910 915
Ile Pro Leu Pro Val Leu Tyr Gly Val Phe Leu Tyr Met Gly Val
920 925 930
Ala Ser Leu Asn Gly Ile Gln Phe Trp Glu Arg Cys Lys Leu Phe
935 940 945
Leu Met Pro Ala Lys His Gln Pro Asp His Ala Phe Leu Arg His
950 955 960
Val Pro Leu Arg Arg Ile His Leu Phe Thr Leu Val Gln Ile Leu
965 970 975
Cys Leu Ala Val Leu Trp Ile Leu Lys Ser Thr Val Ala Ala Ile
980 985 990
Ile Phe Pro Val Met Ile Leu Gly Leu Ile Ile Val Arg Arg Leu
995 1000 1005
Leu Asp Phe Ile Phe Ser Gln His Asp Leu Ala Trp Ile Asp Asn
1010 1015 1020
Ile Leu Pro Glu Lys Glu Lys Lys Glu Thr Asp Lys Lys Arg Lys
1025 1030 1035
Arg Lys Lys Gly Ala His Glu Asp Cys Asp Glu Glu Glu Lys Asp
1040 1045 1050
Leu Pro Val Gly Val Thr His Ser Asp Ser Ser Phe Ser Asp Thr
1055 1060 1065
Glu Leu Asp Arg Ser Tyr Ser Arg Asn Pro Val Phe Met Val Pro
1070 1075 1080
Gln Val Lys Ile Glu Met Glu Ser Asp Tyr Asp Phe Thr Asp Met

```
               1085                1090                1095
Asp Lys Tyr Arg Arg Glu Thr Asp Ser Glu Thr Thr Leu
               1100                1105

<210> 18
<211> 480
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1963058CD1

<400> 18
Met Gly Pro Gly Pro Pro Ala Ala Gly Ala Ala Pro Ser Pro Arg
  1               5                  10                  15
Pro Leu Ser Leu Val Ala Arg Leu Ser Tyr Ala Val Gly His Phe
                20                  25                  30
Leu Asn Asp Leu Cys Ala Ser Met Trp Phe Thr Tyr Leu Leu Leu
                35                  40                  45
Tyr Leu His Ser Val Arg Ala Tyr Ser Ser Arg Gly Ala Gly Leu
                50                  55                  60
Leu Leu Leu Leu Gly Gln Val Ala Asp Gly Leu Cys Thr Pro Leu
                65                  70                  75
Val Gly Tyr Glu Ala Asp Arg Ala Ala Ser Cys Cys Ala Arg Tyr
                80                  85                  90
Gly Pro Arg Lys Ala Trp His Leu Val Gly Thr Val Cys Val Leu
                95                 100                 105
Leu Ser Phe Pro Phe Ile Phe Ser Pro Cys Leu Gly Cys Gly Ala
               110                 115                 120
Ala Thr Pro Glu Trp Ala Ala Leu Leu Tyr Tyr Gly Pro Phe Ile
               125                 130                 135
Val Ile Phe Gln Phe Gly Trp Ala Ser Thr Gln Ile Ser His Leu
               140                 145                 150
Ser Leu Ile Pro Glu Leu Val Thr Asn Asp His Glu Lys Val Glu
               155                 160                 165
Leu Thr Ala Leu Arg Tyr Ala Phe Thr Val Val Ala Asn Ile Thr
               170                 175                 180
Val Tyr Gly Ala Ala Trp Leu Leu Leu His Leu Gln Gly Ser Ser
               185                 190                 195
Arg Val Glu Pro Thr Gln Asp Ile Ser Ile Ser Asp Gln Leu Gly
               200                 205                 210
Gly Gln Asp Val Pro Val Phe Arg Asn Leu Ser Leu Leu Val Val
               215                 220                 225
Gly Val Gly Ala Val Phe Ser Leu Leu Phe His Leu Gly Thr Arg
               230                 235                 240
Glu Arg Arg Arg Pro His Ala Glu Glu Pro Gly Glu His Thr Pro
               245                 250                 255
Leu Leu Ala Pro Ala Thr Ala Gln Pro Leu Leu Leu Trp Lys His
               260                 265                 270
Trp Leu Arg Glu Pro Ala Phe Tyr Gln Val Gly Ile Leu Tyr Met
               275                 280                 285
Thr Thr Arg Leu Ile Val Asn Leu Ser Gln Thr Tyr Met Ala Met
               290                 295                 300
Tyr Leu Thr Tyr Ser Leu His Leu Pro Lys Lys Phe Ile Ala Thr
               305                 310                 315
```

```
Ile Pro Leu Val Met Tyr Leu Ser Gly Phe Leu Ser Ser Phe Leu
                320                 325                 330
Met Lys Pro Ile Asn Lys Cys Ile Gly Arg Asn Met Thr Tyr Phe
                335                 340                 345
Ser Gly Leu Leu Val Ile Leu Ala Phe Ala Ala Trp Val Ala Leu
                350                 355                 360
Ala Glu Gly Leu Gly Val Ala Val Tyr Ala Ala Ala Val Leu Leu
                365                 370                 375
Gly Ala Gly Cys Ala Thr Ile Leu Val Thr Ser Leu Ala Met Thr
                380                 385                 390
Ala Asp Leu Ile Gly Pro His Thr Asn Ser Gly Ala Phe Val Tyr
                395                 400                 405
Gly Ser Met Ser Phe Leu Asp Lys Val Ala Asn Gly Leu Ala Val
                410                 415                 420
Met Ala Ile Gln Ser Leu His Pro Cys Pro Ser Glu Leu Cys Cys
                425                 430                 435
Arg Ala Cys Val Ser Phe Tyr His Trp Ala Met Val Ala Val Thr
                440                 445                 450
Gly Gly Val Gly Val Ala Ala Ala Leu Cys Leu Cys Ser Leu Leu
                455                 460                 465
Leu Trp Pro Thr Arg Leu Arg Arg Trp Asp Arg Asp Ala Arg Pro
                470                 475                 480


<210> 19
<211> 381
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2395967CD1

<400> 19
Met Ser Glu Phe Trp Leu Ile Ser Ala Pro Gly Asp Lys Glu Asn
  1               5                  10                  15
Leu Gln Ala Leu Glu Arg Met Asn Thr Val Thr Ser Lys Ser Asn
                20                  25                  30
Leu Ser Tyr Asn Thr Lys Phe Ala Ile Pro Asp Phe Lys Val Gly
                35                  40                  45
Thr Leu Asp Ser Leu Val Gly Leu Ser Asp Glu Leu Gly Lys Leu
                50                  55                  60
Asp Thr Phe Ala Glu Ser Leu Ile Arg Arg Met Ala Gln Ser Val
                65                  70                  75
Val Glu Val Met Glu Asp Ser Lys Gly Lys Val Gln Glu His Leu
                80                  85                  90
Leu Ala Asn Gly Val Asp Leu Thr Ser Phe Val Thr His Phe Glu
                95                 100                 105
Trp Asp Met Ala Lys Tyr Pro Val Lys Gln Pro Leu Val Ser Val
               110                 115                 120
Val Asp Thr Ile Ala Lys Gln Leu Ala Gln Ile Glu Met Asp Leu
               125                 130                 135
Lys Ser Arg Thr Ala Ala Tyr Asn Thr Leu Lys Thr Asn Leu Glu
               140                 145                 150
Asn Leu Glu Lys Lys Ser Met Gly Asn Leu Phe Thr Arg Thr Leu
               155                 160                 165
```

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

<210> 20
<211> 484
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3586648CD1

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

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

<210> 21
<211> 736
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7473396CD1

<400> 21
Met Gln Asn Ile Thr Lys Glu Phe Gly Thr Phe Lys Ala Asn Asp
1 5 10 15
Asn Ile Asn Leu Gln Val Lys Ala Gly Glu Ile His Ala Leu Leu
20 25 30
Gly Glu Asn Gly Ala Gly Lys Ser Thr Leu Met Asn Val Leu Ser
35 40 45
Gly Leu Leu Glu Pro Thr Ser Gly Lys Ile Leu Met Arg Gly Lys
50 55 60
Glu Val Gln Ile Thr Ser Pro Thr Lys Ala Asn Gln Leu Gly Ile
65 70 75
Gly Met Val His Gln His Phe Met Leu Val Asp Ala Phe Thr Val
80 85 90
Thr Glu Asn Ile Val Leu Gly Ser Glu Pro Ser Arg Ala Gly Met
95 100 105
Leu Asp His Lys Lys Ala Arg Lys Glu Ile Gln Lys Val Ser Glu
110 115 120
Gln Tyr Gly Leu Ser Val Asn Pro Asp Ala Tyr Val Arg Asp Ile
125 130 135
Ser Val Gly Met Glu Gln Arg Val Glu Ile Leu Lys Thr Leu Tyr
140 145 150
Arg Gly Ala Asp Val Leu Ile Phe Asp Glu Pro Thr Ala Val Leu
155 160 165
Thr Pro Gln Glu Ile Asp Glu Leu Ile Val Ile Met Lys Glu Leu
170 175 180
Val Lys Glu Gly Lys Ser Ile Ile Leu Ile Thr His Lys Leu Asp
185 190 195
Glu Ile Lys Ala Val Ala Asp Arg Cys Thr Val Ile Arg Arg Gly
200 205 210
Lys Gly Ile Gly Thr Val Asn Val Lys Asp Val Thr Ser Gln Gln
215 220 225
Leu Ala Asp Met Met Val Gly Arg Ala Val Ser Phe Lys Thr Met
230 235 240
Lys Lys Glu Ala Lys Pro Gln Glu Val Val Leu Ser Ile Glu Asn
245 250 255
Leu Val Val Lys Glu Asn Arg Gly Leu Glu Ala Val Lys Asn Leu
260 265 270
Asn Leu Glu Val Arg Ala Gly Glu Val Leu Gly Ile Ala Gly Ile
275 280 285
Asp Gly Asn Gly Gln Ser Glu Leu Ile Gln Ala Leu Thr Gly Leu
290 295 300
Arg Lys Ala Glu Ser Gly His Ile Lys Leu Lys Gly Glu Asp Ile
305 310 315
Thr Asn Lys Lys Pro Arg Lys Ile Thr Glu His Gly Val Gly His
320 325 330
Val Pro Glu Asp Arg His Lys Tyr Gly Leu Val Leu Asp Met Thr
335 340 345
Leu Ser Glu Asn Ile Ala Leu Gln Thr Tyr His Gln Lys Pro Tyr
350 355 360
Ser Lys Asn Gly Met Leu Asn Tyr Ser Val Ile Asn Glu His Ala
365 370 375
Arg Glu Leu Ile Glu Glu Tyr Asp Val Arg Thr Thr Asn Glu Leu

```
                380                 385                 390
Val Pro Ala Lys Ala Leu Ser Gly Gly Asn Gln Gln Lys Ala Ile
                395                 400                 405
Ile Ala Arg Ile Val Asp Arg Asp Pro Asp Leu Leu Ile Val Ala
                410                 415                 420
Asn Pro Thr Arg Gly Leu Asp Val Gly Glu Phe Val Ala Val Thr
                425                 430                 435
Gly Val Ser Gly Ser Gly Lys Ser Thr Leu Val Asn Ser Ile Leu
                440                 445                 450
Lys Lys Ser Leu Ala Gln Lys Leu Asn Lys Asn Ser Ala Lys Pro
                455                 460                 465
Gly Lys Phe Lys Thr Ile Ser Gly Tyr Glu Ser Ile Glu Lys Ile
                470                 475                 480
Ile Asp Ile Asp Gln Ser Pro Ile Gly Arg Thr Pro Arg Ser Asn
                485                 490                 495
Pro Ala Thr Tyr Thr Ser Val Phe Asp Asp Ile Arg Gly Leu Phe
                500                 505                 510
Ala Gln Thr Asn Glu Ala Lys Met Arg Gly Tyr Lys Lys Gly Arg
                515                 520                 525
Phe Ser Phe Asn Val Lys Gly Gly Arg Cys Glu Ala Cys Arg Gly
                530                 535                 540
Asp Gly Ile Ile Lys Ile Glu Met His Phe Leu Pro Asp Val Tyr
                545                 550                 555
Val Pro Cys Glu Val Cys His Gly Lys Arg Tyr Asn Ser Glu Thr
                560                 565                 570
Leu Glu Val His Tyr Lys Gly Lys Ser Ile Ala Asp Ile Leu Glu
                575                 580                 585
Met Thr Val Glu Asp Ala Val Glu Phe Phe Lys His Ile Pro Lys
                590                 595                 600
Ile His Arg Lys Leu Gln Thr Ile Val Asp Val Gly Leu Gly Tyr
                605                 610                 615
Val Thr Met Gly Gln Pro Ala Thr Thr Leu Ser Gly Gly Glu Ala
                620                 625                 630
Gln Arg Met Lys Leu Ala Ser Glu Leu His Lys Ile Ser Asn Gly
                635                 640                 645
Lys Asn Phe Tyr Ile Leu Asp Glu Pro Thr Thr Gly Leu His Ser
                650                 655                 660
Asp Asp Ile Ala Arg Leu Leu His Val Leu Gln Arg Leu Val Asp
                665                 670                 675
Ala Gly Asn Thr Val Leu Val Ile Glu His Asn Leu Asp Val Ile
                680                 685                 690
Lys Thr Ala Asp Tyr Ile Ile Asp Leu Gly Pro Glu Gly Gly Glu
                695                 700                 705
Gly Gly Gly Thr Ile Leu Thr Thr Gly Thr Pro Glu Glu Ile Ile
                710                 715                 720
Asn Val Lys Glu Ser Tyr Thr Gly His Tyr Leu Lys Lys Ile Met
                725                 730                 735
Val
```

<210> 22
<211> 465
<212> PRT
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 7476283CD1

<400> 22
Met Gly Pro Leu Lys Ala Phe Leu Phe Ser Pro Phe Leu Leu Arg
1 5 10 15
Ser Gln Ser Arg Gly Val Arg Leu Val Phe Leu Leu Leu Thr Leu
20 25 30
His Leu Gly Asn Cys Val Asp Lys Ala Asp Asp Glu Asp Asp Glu
35 40 45
Asp Leu Lys Val Asn Lys Thr Trp Val Leu Ala Pro Lys Ile His
50 55 60
Glu Gly Asp Ile Thr Gln Ile Leu Asn Ser Leu Leu Gln Gly Tyr
65 70 75
Asp Asn Lys Leu Arg Pro Asp Ile Gly Val Arg Pro Thr Val Ile
80 85 90
Glu Thr Asp Val Tyr Val Asn Ser Ile Gly Pro Val Asp Pro Ile
95 100 105
Asn Met Glu Tyr Thr Ile Asp Ile Ile Phe Ala Gln Thr Trp Phe
110 115 120
Asp Ser Arg Leu Lys Phe Asn Ser Thr Met Lys Val Leu Met Leu
125 130 135
Asn Ser Asn Met Val Gly Lys Ile Trp Ile Pro Asp Thr Phe Phe
140 145 150
Arg Asn Ser Arg Lys Ser Asp Ala His Trp Ile Thr Thr Pro Asn
155 160 165
Arg Leu Leu Arg Ile Trp Asn Asp Gly Arg Val Leu Tyr Thr Leu
170 175 180
Arg Leu Thr Ile Asn Ala Glu Cys Tyr Leu Gln Leu His Asn Phe
185 190 195
Pro Met Asp Glu His Ser Cys Pro Leu Glu Phe Ser Ser Asp Gly
200 205 210
Tyr Pro Lys Asn Glu Ile Glu Tyr Lys Trp Lys Lys Pro Ser Val
215 220 225
Glu Val Ala Asp Pro Lys Tyr Trp Arg Leu Tyr Gln Phe Ala Phe
230 235 240
Val Gly Leu Arg Asn Ser Thr Glu Ile Thr His Thr Ile Ser Gly
245 250 255
Asp Tyr Val Ile Met Thr Ile Phe Phe Asp Leu Ser Arg Arg Met
260 265 270
Gly Tyr Phe Thr Ile Gln Thr Tyr Ile Pro Cys Ile Leu Thr Val
275 280 285
Val Leu Ser Trp Val Ser Phe Trp Ile Asn Lys Asp Ala Val Pro
290 295 300
Ala Arg Thr Ser Leu Gly Ile Thr Thr Val Leu Thr Met Thr Thr
305 310 315
Leu Ser Thr Ile Ala Arg Lys Ser Leu Pro Lys Val Ser Tyr Val
320 325 330
Thr Ala Met Asp Leu Phe Val Ser Val Cys Phe Ile Phe Val Phe
335 340 345
Ala Ala Leu Met Glu Tyr Gly Thr Leu His Tyr Phe Thr Ser Asn
350 355 360
Gln Lys Gly Lys Thr Ala Thr Lys Asp Arg Lys Leu Lys Asn Lys
365 370 375
Ala Ser Met Thr Pro Gly Leu His Pro Gly Ser Thr Leu Ile Pro
380 385 390

```
Met Asn Asn Ile Ser Val Pro Gln Glu Asp Asp Tyr Gly Tyr Gln
                395                 400                 405
Cys Leu Glu Gly Lys Asp Cys Ala Ser Phe Phe Cys Cys Phe Glu
                410                 415                 420
Asp Cys Arg Thr Gly Ser Trp Arg Glu Gly Arg Ile His Ile Arg
                425                 430                 435
Ile Ala Lys Ile Asp Ser Tyr Ser Arg Ile Phe Phe Pro Thr Ala
                440                 445                 450
Phe Ala Leu Phe Asn Leu Val Tyr Trp Val Gly Tyr Leu Tyr Leu
                455                 460                 465


<210> 23
<211> 235
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477105CD1

<400> 23
Met Gly Ser Val Gly Ser Gln Arg Leu Glu Glu Pro Ser Val Ala
  1               5                  10                  15
Gly Thr Pro Asp Pro Gly Val Val Met Ser Phe Thr Phe Asp Ser
                20                  25                  30
His Gln Leu Glu Glu Ala Ala Glu Ala Ala Gln Gly Gln Gly Leu
                35                  40                  45
Arg Ala Arg Gly Val Pro Ala Phe Thr Asp Thr Thr Leu Asp Glu
                50                  55                  60
Pro Val Pro Asp Asp Arg Tyr His Ala Ile Tyr Phe Ala Met Leu
                65                  70                  75
Leu Ala Gly Val Gly Phe Leu Leu Pro Tyr Asn Ser Phe Ile Thr
                80                  85                  90
Asp Val Asp Tyr Leu His His Lys Tyr Pro Gly Thr Ser Ile Val
                95                 100                 105
Phe Asp Met Ser Leu Thr Tyr Ile Leu Val Ala Leu Ala Ala Val
                110                 115                 120
Leu Leu Asn Asn Val Leu Val Glu Arg Leu Thr Leu His Thr Arg
                125                 130                 135
Ile Thr Ala Gly Tyr Leu Leu Ala Leu Gly Pro Leu Leu Phe Ile
                140                 145                 150
Ser Ile Cys Asp Val Trp Leu Gln Leu Phe Ser Arg Asp Gln Ala
                155                 160                 165
Tyr Ala Ile Asn Leu Ala Ala Val Gly Thr Val Ala Phe Gly Cys
                170                 175                 180
Thr Val Gln Gln Ser Ser Phe Tyr Gly His Arg Leu Ala Gln Pro
                185                 190                 195
Pro Pro Gly Thr Pro Pro His Glu Leu Trp Ser Pro Glu Arg Arg
                200                 205                 210
Gly Ala Ala Pro His Leu Val Thr Leu Arg Ala Ser Pro Ser Val
                215                 220                 225
Leu Ile Leu Arg Asp Cys Phe Ser Gln Thr
                230                 235

<210> 24
```

<211> 662
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7482079CD1

<400> 24
Met Leu Lys Gln Ser Glu Arg Arg Arg Ser Trp Ser Tyr Arg Pro
1 5 10 15
Trp Asn Thr Thr Glu Asn Glu Gly Ser Gln His Arg Arg Ser Ile
20 25 30
Cys Ser Leu Gly Ala Arg Ser Gly Ser Gln Ala Ser Ile His Gly
35 40 45
Trp Thr Glu Gly Asn Tyr Asn Tyr Tyr Ile Glu Glu Asp Glu Asp
50 55 60
Gly Glu Glu Glu Asp Gln Trp Lys Asp Asp Leu Ala Glu Glu Asp
65 70 75
Gln Gln Ala Gly Glu Val Thr Thr Ala Lys Pro Glu Gly Pro Ser
80 85 90
Asp Pro Pro Ala Leu Leu Ser Thr Leu Asn Val Asn Val Gly Gly
95 100 105
His Ser Tyr Gln Leu Asp Tyr Cys Glu Leu Ala Gly Phe Pro Lys
110 115 120
Thr Arg Leu Gly Arg Leu Ala Thr Ser Thr Ser Arg Ser Arg Gln
125 130 135
Leu Ser Leu Cys Asp Asp Tyr Glu Glu Gln Thr Asp Glu Tyr Phe
140 145 150
Phe Asp Arg Asp Pro Ala Val Phe Gln Leu Val Tyr Asn Phe Tyr
155 160 165
Leu Ser Gly Val Leu Leu Val Leu Asp Gly Leu Cys Pro Arg Arg
170 175 180
Phe Leu Glu Glu Leu Gly Tyr Trp Gly Val Arg Leu Lys Tyr Thr
185 190 195
Pro Arg Cys Cys Arg Ile Cys Phe Glu Glu Arg Arg Asp Glu Leu
200 205 210
Ser Glu Arg Leu Lys Ile Gln His Glu Leu Arg Ala Gln Ala Gln
215 220 225
Val Glu Glu Ala Glu Glu Leu Phe Arg Asp Met Arg Phe Tyr Gly
230 235 240
Pro Gln Arg Arg Arg Leu Trp Asn Leu Met Glu Lys Pro Phe Ser
245 250 255
Ser Val Ala Ala Lys Ala Ile Gly Val Ala Ser Ser Thr Phe Val
260 265 270
Leu Val Ser Val Val Ala Leu Ala Leu Asn Thr Val Glu Glu Met
275 280 285
Gln Gln His Ser Gly Gln Gly Glu Gly Gly Pro Asp Leu Arg Pro
290 295 300
Ile Leu Glu His Val Glu Met Leu Cys Met Gly Phe Phe Thr Leu
305 310 315
Glu Tyr Leu Leu Arg Leu Ala Ser Thr Pro Asp Leu Arg Arg Phe
320 325 330
Ala Arg Ser Ala Leu Asn Leu Val Asp Leu Val Ala Ile Leu Pro
335 340 345
Leu Tyr Leu Gln Leu Leu Leu Glu Cys Phe Thr Gly Glu Gly His

```
                350                 355                 360
Gln Arg Gly Gln Thr Val Gly Ser Val Gly Lys Val Gly Gln Val
                365                 370                 375
Leu Arg Val Met Arg Leu Met Arg Ile Phe Arg Ile Leu Lys Leu
                380                 385                 390
Ala Arg His Ser Thr Gly Leu Arg Ala Phe Gly Phe Thr Leu Arg
                395                 400                 405
Gln Cys Tyr Gln Gln Val Gly Cys Leu Leu Leu Phe Ile Ala Met
                410                 415                 420
Gly Ile Phe Thr Phe Ser Ala Ala Val Tyr Ser Val Glu His Asp
                425                 430                 435
Val Pro Ser Thr Asn Phe Thr Thr Ile Pro His Ser Trp Trp Trp
                440                 445                 450
Ala Ala Val Ser Thr Phe Ala Leu Gly Phe Pro Ile Leu Phe Pro
                455                 460                 465
Ser Pro Val Ser Cys Ser Ser Leu Pro Trp Leu Ser Ala Thr Arg
                470                 475                 480
Leu Trp Leu Leu Ile Leu Val Phe Pro Pro Thr Pro Asn Arg Arg
                485                 490                 495
Ile Gln Leu Thr Lys Arg Arg Trp Met Ser Lys Val Val Glu Arg
                500                 505                 510
Glu Leu Ser Arg Ser Val Asn Ser Ser Ser His Met Ser Met Ala
                515                 520                 525
Val Ala Lys Asn Lys Arg Glu Asn Ala Ser Pro Ile Met Gln Thr
                530                 535                 540
Leu His Lys Phe Leu Phe Met Ala Phe Ala Gln Pro Ile Gly Gln
                545                 550                 555
Ser Lys Ser His Gly Gln Ala Ala Ser Gln Arg Ala Gly Gln Val
                560                 565                 570
Ser Ile Ser Thr Val Gly Tyr Gly Asp Met Tyr Pro Glu Thr His
                575                 580                 585
Leu Gly Arg Phe Phe Ala Phe Leu Cys Ile Ala Phe Gly Ile Ile
                590                 595                 600
Leu Asn Gly Met Pro Ile Ser Ile Leu Tyr Asn Lys Phe Ser Asp
                605                 610                 615
Tyr Tyr Ser Lys Leu Lys Ala Tyr Glu Tyr Thr Thr Ile Arg Arg
                620                 625                 630
Glu Arg Gly Glu Val Asn Phe Met Gln Arg Ala Arg Lys Lys Ile
                635                 640                 645
Ala Glu Cys Leu Leu Gly Ser Asn Pro Gln Leu Thr Pro Arg Gln
                650                 655                 660
Glu Asn


<210> 25
<211> 371
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 55145506CD1

<400> 25
Met Asn Asp Glu Asp Tyr Ser Thr Ile Tyr Asp Thr Ile Gln Asn
  1               5                  10                  15
```

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

<210> 26
<211> 468
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
```

<223> Incyte ID No: 5950519CD1

<400> 26

```
Met Arg Gly Ser Pro Gly Asp Ala Glu Arg Arg Gln Arg Trp Gly
  1               5                  10                  15
Arg Leu Phe Glu Glu Leu Asp Ser Asn Lys Asp Gly Arg Val Asp
             20                  25                  30
Val His Glu Leu Arg Gln Gly Leu Ala Arg Leu Gly Gly Gly Asn
             35                  40                  45
Pro Asp Pro Gly Ala Gln Gln Gly Ile Ser Ser Glu Gly Asp Ala
             50                  55                  60
Asp Pro Asp Gly Gly Leu Asp Leu Glu Glu Phe Ser Arg Tyr Leu
             65                  70                  75
Gln Glu Arg Glu Gln Arg Leu Leu Leu Met Phe His Ser Leu Asp
             80                  85                  90
Arg Asn Gln Asp Gly His Ile Asp Val Ser Glu Ile Gln Gln Ser
             95                 100                 105
Phe Arg Ala Leu Gly Ile Ser Ile Ser Leu Glu Gln Ala Glu Lys
            110                 115                 120
Ile Leu His Ser Met Asp Arg Asp Gly Thr Met Thr Ile Asp Trp
            125                 130                 135
Gln Glu Trp Arg Asp His Phe Leu Leu His Ser Leu Glu Asn Val
            140                 145                 150
Glu Asp Val Leu Tyr Phe Trp Lys His Ser Thr Val Leu Asp Ile
            155                 160                 165
Gly Glu Cys Leu Thr Val Pro Asp Glu Phe Ser Lys Gln Glu Lys
            170                 175                 180
Leu Thr Gly Met Trp Trp Lys Gln Leu Val Ala Gly Ala Val Ala
            185                 190                 195
Gly Ala Val Ser Arg Thr Gly Thr Ala Pro Leu Asp Arg Leu Lys
            200                 205                 210
Val Phe Met Gln Val His Ala Ser Lys Thr Asn Arg Leu Asn Ile
            215                 220                 225
Leu Gly Gly Leu Arg Ser Met Val Leu Glu Gly Gly Ile Arg Ser
            230                 235                 240
Leu Trp Arg Gly Asn Gly Ile Asn Val Leu Lys Ile Ala Pro Glu
            245                 250                 255
Ser Ala Ile Lys Phe Met Ala Tyr Glu Gln Ile Lys Arg Ala Ile
            260                 265                 270
Leu Gly Gln Gln Glu Thr Leu His Val Gln Glu Arg Phe Val Ala
            275                 280                 285
Gly Ser Leu Ala Gly Ala Thr Ala Gln Thr Ile Ile Tyr Pro Met
            290                 295                 300
Glu Val Leu Lys Thr Arg Leu Thr Leu Arg Arg Thr Gly Gln Tyr
            305                 310                 315
Lys Gly Leu Leu Asp Cys Ala Arg Arg Ile Leu Glu Arg Glu Gly
            320                 325                 330
Pro Arg Ala Phe Tyr Arg Gly Tyr Leu Pro Asn Val Leu Gly Ile
            335                 340                 345
Ile Pro Tyr Ala Gly Ile Asp Leu Ala Val Tyr Glu Thr Leu Lys
            350                 355                 360
Asn Trp Trp Leu Gln Gln Tyr Ser His Asp Ser Ala Asp Pro Gly
            365                 370                 375
Ile Leu Val Leu Leu Ala Cys Gly Thr Ile Ser Ser Thr Cys Gly
            380                 385                 390
Gln Ile Ala Ser Tyr Pro Leu Ala Leu Val Arg Thr Arg Met Gln
```

```
                395                 400                 405
Ala Gln Ala Ser Ile Glu Gly Gly Pro Gln Leu Ser Met Leu Gly
                410                 415                 420
Leu Leu Arg His Ile Leu Ser Gln Glu Gly Met Arg Gly Leu Tyr
                425                 430                 435
Arg Gly Ile Ala Pro Asn Phe Met Lys Val Ile Pro Ala Val Ser
                440                 445                 450
Ile Ser Tyr Val Val Tyr Glu Asn Met Lys Gln Ala Leu Gly Val
                455                 460                 465
Thr Ser Arg
```

<210> 27
<211> 2229
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1687189CB1

<400> 27

```
gcctgagcgg ccgaactcgg cagctccaac ccaactcggc ttaactccgc ctcaccgagc 60
ccagtccaag actctgtgct ccctaggttt gcaacagctc tctgatcatc ttcttcaatt 120
cctgctagga tgccgtggca agcatttcgc agatttggtc aaaagctggt acgcagacgt 180
acactggagt caggcatggc tgagactcgc cttgccagat gcctaagcac cctggattta 240
gtggccctgg gtgtgggcag cacattgggt gcaggcgtgt atgtcctagc tggcgaggtg 300
gccaaagata aagcagggcc atccattgtg atctgctttt tggtggctgc cctgtcttct 360
gtgttggctg ggctgtgcta tgcggagttt ggtgcccggg ttccccgttc tggttcggca 420
tatctctaca gctatgtcac tgtgggtgaa ctctgggcct tcaccactgg ctggaacctc 480
atcctctcct atgtcattgg tacagccagt gtggcccggg cctggagctc tgcttttgac 540
aacctgattg ggaaccacat ctctaagact ctgcaggggt ccattgcact gcacgtgccc 600
catgtccttg cagaatatcc agatttcttt gctttgggcc tcgtgttgct gctcactgga 660
ttgttggctc tcggggctag tgagtcggcc ctggttacca aagtgttcac aggcgtgaac 720
cttttggttc ttgggttcgt catgatctct ggcttcgtta aggggacgt gcacaactgg 780
aagctcacag aagaggacta cgaattggcc atggctgaac tcaatgacac ctatagcttg 840
ggtcctctgg gctctggagg atttgtgcct ttcggcttcg agggaattct ccgtggagca 900
gcgacctgtt tctatgcatt tgttggtttc gactgtattg ctaccactgg agaagaagcc 960
cagaatcccc agcgttccat cccgatgggc attgtgatct cactgtctgt ctgcttttg 1020
gcgtattttg ctgtctcttc tgcactcacc ctgatgatgc cttactacca gcttcagcct 1080
gagagccctt tgcctgaggc atttcctcac attggatggg ctcctgcccg ctatgttgtg 1140
gctgttggct ccctctgtgc tctttctacc agcctcctgg gctccatgtt ccccatgcct 1200
cgggtgatct acgcgatggc agaggatggc ctcctgttcc gtgtacttgc tcggatccac 1260
accggcacac gcaccccaat catagccacc gtggtctctg gcattattgc agcattcatg 1320
gcattcctct tcaaactcac tgatcttgtg gacctcatgt caattgggac cctgcttgct 1380
tactccctgg tgtcgatttg tgttctcatc ctcaggtatc aacctgatca ggagacaaag 1440
actggggaag aagtggagtt gcaggaggag gcaataacta ctgaatcaga gaagttgacc 1500
ctatggggac tatttttccc actcaactcc atccccactc cactctctgg ccaaattgtc 1560
tatgtttgtt cctcattgct tgctgtcctg ctgactgctc tttgcctggt gctggcccag 1620
tggtcagttc cattgctttc tggagacctg ctgtggactg cagtggttgt gctgctcctg 1680
ctgctcatta ttgggatcat tgtggtcatc tggagacagc cacagagctc cactcccctt 1740
cactttaagg tgcctgcttt gcctctcctc ccactaatga gcatctttgt gaatatttac 1800
cttatgatgc agatgacagc tggtacctgg gcccgatttg gggtctggat gctgattggc 1860
tttgctatct acttcggcta tgggatccag cacagcctgg aagagattaa gagtaaccaa 1920
ccctcacgca agtctagagc caaaactgta gaccttgatc ccggcactct ctatgtccac 1980
tcagtttgac atcgtcacac ctaaatgctg tctggtcccc tgcacaataa tggagagtac 2040
```

tcctgacccc agtgacagct agccctcccc tgtgatggtg gtggtggata ctaatacagt 2100
tctgtacgat gtgaaggatg tgtctttgct atttcttgtc tattttaacc cgtctgcttc 2160
taaatgatgt ctagctgctt accaacttta aaaaatgata ttaaaagaaa gtagaaaaat 2220
aaaaaaaaa 2229

<210> 28
<211> 7610
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7078207CB1

<400> 28
gcgccccgcc cccgcgcggg cgatgcccag cggcgcggcg ggctgcgggg cccggcgggg 60
cgcgcagagg agcgggccgc ggcgctgagg cggcggagcg tggccccgcc atgggcttcc 120
tgcaccagct gcagctgctg ctctggaaga acgtgacgct caaacgccgg agcccgtggg 180
tcctggcctt cgagatcttc atccccctgg tgctgttctt tatcctgctg gggctgcgac 240
agaagaagcc caccatctcc gtgaaggaag tctccttcta cacagcggcg cccctgacgt 300
ctgccggcat cctgcctgtc atgcaatcgc tgtgcccgga cggccagcga gacgagttcg 360
gcttcctgca gtacgccaac tccacggtca cgcagctgct tgagcgcctg gaccgcgtgg 420
tggaggaagg caacctgttt gacccagcgc ggcccagcct gggctcagag ctcgaggccc 480
tacgccagca tctggaggcc ctcagtgcgg gcccgggcac ctcggggagc cacctggaca 540
gatccacagt gtcttccttc tctctggact cggtggccag aaacccgcag gagctctggc 600
gtttcctgac gcaaaacttg tcgctgccca atagcacggc ccaagcactc ttggccgccc 660
gtgtggaccc gcccgaggtc taccacctgc tctttggtcc ctcatctgcc ctggatttcac 720
agtctggcct ccacaagggt caggagccct ggagccgcct agggggcaat cccctgttcc 780
ggatggagga gctgctgctg gctcctgccc tcctggagca gctcacctgc acgccgggct 840
cgggggagct gggccggatc ctcactgtgc ctgagagtca gaagggagcc ctgcagggct 900
accgggatgc tgtctgcagt gggcaggctg ctgcgcgtgc caggcgcttc tctgggctgt 960
ctgctgagct ccggaaccag ctggacgtgg ccaaggtctc ccagcagctg ggcctggatg 1020
ccccaacgg ctcggactcc tcgccacagg cgccaccccc acggaggctg caggcgcttc 1080
tgggggacct gctggatgcc cagaaggttc tgcaggatgt ggatgtcctg tcggccctgg 1140
ccctgctact gccccagggt gcctgcactg gccggacccc cggaccccca gccagtggtg 1200
cgggtggggc ggccaatggc actggggcag gggcagtcat gggccccaac gccaccgctg 1260
aggagggcgc accctctgct gcagcactgg ccaccccgga cacgctgcag ggccagtgct 1320
cagccttcgt acagctctgg gccggcctgc agcccatctt gtgtggcaac aaccgcacca 1380
ttgaacccga ggcgctgcgg cggggcaaca tgagctccct gggcttcacg agcaaggagc 1440
agcggaacct gggcctcctc gtgcacctca tgaccagcaa ccccaaaatc ctgtacgcgc 1500
ctgcgggctc tgaggtcgac cgcgtcatcc tcaaggccaa cgagactttt gcttttgtgg 1560
gcaacgtgac tcactatgcc caggtctggc tcaacatctc ggcggagatc cgcagcttcc 1620
tggagcaggg caggctgcag caacacctgc gctggctgca gcagtatgta gcagagctgc 1680
ggctgcaccc cgaggcactg aacctgtcac tggatgagct gccgccggcc ctgagacagg 1740
acaacttctc gctgcccagt ggcatggccc tcctgcagca gctggatacc attgacaacg 1800
cggcctgcgg ctggatccag ttcatgtcca aggtgagcgt ggacatcttc aagggcttcc 1860
ccgacgagga gagcattgtc aactacaccc tcaaccaggc ctaccaggac aacgtcactg 1920
tttttgccag tgtgatcttc cagacccgga aggacgggctc gctcccgcct cacgtgcact 1980
acaagatccg ccagaactcc agcttcaccg agaaaaccaa cgagatccgc cgcgcctact 2040
ggcggcctgg gcccaatact ggcggccgct tctacttcct ctacggcttc gtctggatcc 2100
aggacatgat ggagcgcgcc atcatcgaca cttttgtggg gcacgacgtg gtggagccag 2160
gcagctacgt gcagatgttc ccctaccct gctacacacg cgatgacttc ctgtttgtca 2220
ttgagcacat gatgccgctg tgcatggtga tctcctgggt ctactccgtg gccatgacca 2280
tccagcacat cgtggcggag aaggagcacc ggctcaagga ggtgatgaag accatgggcc 2340
tgaacaacgc ggtgcactgg gtggcctggt tcatcaccgg ctttgtgcag ctgtccatct 2400
ccgtgacagc actcaccgcc atcctgaagt acggccaggt gcttatgcac agccacgtgg 2460

tcatcatctg gctcttcctg gcagtctacg cggtggccac catcatgttc tgcttcctgg 2520
tgtctgtgct gtactccaag gccaagctgg cctcggcctg cggtggcatc atctacttcc 2580
tgagctacgt gccctacatg tacgtggcga tccgagagga ggtggcgcat gataagatca 2640
cggccttcga gaagtgcatc gcgtccctca tgtccacgac ggcctttggt ctgggctcta 2700
agtacttcgc gctgtatgag gtggccggcg tgggcatcca gtggcacacc ttcagccagt 2760
ccccggtgga gggggacgac ttcaacttgc tcctggctgt caccatgctg atggtggacg 2820
ccgtggtcta tggcatcctc acgtggtaca ttgaggctgt gcacccaggc atgtacgggc 2880
tgccccggcc ctggtacttc ccactgcaga agtcctactg gctgggcagt gggcggacag 2940
aagcctggga gtggagctgg ccgtgggcac gcacccccg cctcagtgtc atggaggagg 3000
accaggcctg tgccatggag agccggcgct ttgaggagac ccgtggcatg gaggaggagc 3060
ccacccacct gcctctggtt gtctgcgtgg acaaactcac caaggtctac aaggacgaca 3120
agaagctggc cctgaacaag ctgagcctga acctctacga gaaccaggtg gtctccttct 3180
tgggccacaa cggggcgggc aagaccacca ccatgtccat cctgaccggc ctgttccctc 3240
caacgtcggg ttccgccacc atctacgggc acgacatccg cacggagatg gatgagatcc 3300
gcaagaacct gggcatgtgc ccgcagcaca atgtgctctt tgaccggctc acggtggagg 3360
aacacctctg gttctactca cggctcaaga gcatggctca ggaggagatc cgcagagaga 3420
tggacaagat gatcgaggac ctggagctct ccaacaaacg gcactcactg gtgcagacat 3480
tgtcgggtgg catgaagcgc aagctgtccg tggccatcgc cttcgtgggc ggctctcgcg 3540
ccatcatcct ggacgagccc acggcgggcg tggacccctа cgcgcgccgc gccatctggg 3600
acctcatcct gaagtacaag ccaggccgca ccatccttct gtccacccac cacatggatg 3660
aggctgacct gcttggggac cgcattgcca tcatctccca tgggaagctc aagtgctgcg 3720
gctccccgct cttcctcaag ggcacctatg gcgacgggta ccgcctcacg ctggtcaagc 3780
ggcccgccga gccggggggc ccccaagagc cagggctggc atccagcccc ccaggtcggg 3840
ccccgctgag cagctgctcc gagctccagg tgtcccagtt catccgcaag catgtggcct 3900
cctgcctgct ggtctcagac acaagcacgg agctctccta catcctgccc agcgaggccg 3960
ccaagaaggg ggctttcgag cgcctcttcc agcacctgga gcgcagcctg gatgcactgc 4020
acctcagcag cttcgggctg atggacacga ccctggagga agtgttcctc aaggtgtcgg 4080
aggaggatca gtcgctggag aacagtgagg ccgatgtgaa ggagtccagg aaggatgtgc 4140
tccctggggc ggagggcccg gcgtctgggg agggtcacgc tggcaatctg gcccggtgct 4200
cggagctgac ccagtcgcag gcatcgctgc agtcggcgtc atctgtgggc tctgcccgtg 4260
gcgacgaggg agctggctac accgacgtct atggcgacta ccgccccctc tttgataacc 4320
cacaggaccc agacaatgtc agcctgcaag aggtggaggc agaggccctg tcgagggtcg 4380
gccagggcag ccgcaagctg gacggcgggt ggctgaaggt gcgccagttc cacgggctgc 4440
tggtcaaacg cttccactgc gcccgccgca actccaaggc actcttctcc cagatcttgc 4500
tgccagcctt cttcgtctgc gtggccatga ccgtggccct gtccgtcccg gagattggtg 4560
atctgccccc gctggtcctg tcaccttccc agtaccacaa ctacacccag ccccgtggca 4620
atttcatccc ctacgccaac gaggagcgcc gcgagtaccg gctgcggcta tcgcccgacg 4680
ccagccccca gcagctcgtg agcacgttcc ggctgccgtc gggggtgggt gccacctgcg 4740
tgctcaagtc tcccgccaac ggctcgctgg ggcccacgtt gaacctgagc agcggggagt 4800
cgcgcctgct ggcggctcgg ttcttcgaca gcatgtgtct ggagtccttc acacaggggc 4860
tgccactgtc caatttcgtg ccacccccac cctcgcccgc cccatctgac tcgccagcgt 4920
ccccggatga ggacctgcag gcctggaacg tctccctgcc gcccaccgct gggccagaaa 4980
tgtggacgtc ggcaccctcc ctgccgcgcc tggtacggga gcccgtccgc tgcacctgct 5040
ctgcgcaggg caccggcttc tcctgcccca gcagtgtggg cgggcacccg ccccagatgc 5100
gggtggtcac aggcgacatc ctgaccgaca tcaccggcca caatgtctct gagtacctgc 5160
tcttcacctc cgaccgcttc cgactgcacc ggtatggggc catcaccttt ggaaacgtcc 5220
tgaagtccat cccagcctca tttggcacca gggccccacc catggtgcgg aagatcgcgg 5280
tgcgcagggc tgcccaggtt ttctacaaca acaagggcta tcacagcatg cccacctacc 5340
tcaacagcct caacaacgcc atcctgcgtg ccaacctgcc caagagcaag ggcaacccgg 5400
cggcttacgg catcaccgtc accaaccacc ccatgaataa gaccagcgcc agcctctccc 5460
tggattacct gctgcagggc acggatgtcg tcatcgccat cttcatcatc gtggccatgt 5520
ccttcgtgcc ggccagcttc gttgtcttcc tcgtggccga gaagtccacc aagggcaagc 5580
atctgcagtt tgtcagcggc tgcaacccca tcatctactg gctggcgaac tacgtgtggg 5640
acatgctcaa ctacctggtc cccgctacct gctgtgtcat catcctgttt gtgttcgacc 5700
tgccggccta cacgtcgccc accaacttcc ctgccgtcct ctccctcttc ctgctctatg 5760
ggtggtccat cacgcccatc atgtacccgg cctccttctg gttcgaggtc cccagctccg 5820

cctacgtgtt cctcattgtc atcaatctct tcatcggcat caccgccacc gtggccacct 5880
tcctgctaca gctcttcgag cacgacaagg acctgaaggt tgtcaacagt tacctgaaaa 5940
gctgcttcct cattttcccc aactacaacc tgggccacgg gctcatggag atggcctaca 6000
acgagtacat caacgagtac tacgccaaga ttggccagtt tgacaagatg aagtccccgt 6060
tcgagtggga cattgtcacc cgcggactgg tggccatggc ggttgagggc gtcgtgggct 6120
tcctcctgac catcatgtgc cagtacaact tcctgcggcg gccacagcgc atgcctgtgt 6180
ctaccaagcc tgtggaggat gatgtggacg tggccagtga gcggcagcga gtgctccggg 6240
gagacgccga caatgacatg gtcaagattg agaacctgac caaggtctac aagtcccgga 6300
agattggccg tatcctggcc gttgaccgcc tgtgcctggg tgtgcgtcct ggcgagtgct 6360
tcgggctcct gggcgtcaac ggtgcgggca agaccagcac cttcaagatg ctgaccggcg 6420
acgagagcac gacggggggc gaggccttcg tcaatggaca cagcgtgctg aaggagctgc 6480
tccaggtgca gcagagcctc ggctactgcc cgcagtgtga cgcgctgttc gacgagctca 6540
cggcccggga gcacctgcag ctgtacacgc ggctgcgtgg gatctcctgg aaggacgagg 6600
cccgggtggt gaagtgggct ctggagaagc tggagctgac caagtacgca gacaagccgg 6660
ctggcaccta cagcggcggc aacaagcgga agctctccac ggccatcgcc ctcattgggt 6720
acccagcctt catcttcctg gacgagccca ccacaggcat ggaccccaag gcccggcgct 6780
tcctctggaa cctcatcctc gacctcatca agacagggcg ttcagtggtg ctgacatcac 6840
acagcatgga ggagtgcgag gcgctgtgca cgcggctggc catcatggtg aacggtcgcc 6900
tgcggtgcct gggcagcatc cagcacctga agaaccggtt tggagatggc tacatgatca 6960
cggtgcggac caagagcagc cagagtgtga aggacgtggt gcggttcttc aaccgcaact 7020
tcccggaagc catgctcaag gagcggcacc acacaaaggt gcagtaccag ctcaagtcgg 7080
agcacatctc gctggcccag gtgttcagca agatggagca ggtgtctggc gtgctgggca 7140
tcgaggacta ctcggtcagc cagaccacac tggacaatgt gttcgtgaac tttgccaaga 7200
agcagagtga caacctggag cagcaggaga cggagccgcc atccgcactg cagtcccctc 7260
tcggctgctt gctcagcctg ctccggcccc ggtctgcccc cacggagctc cgggcacttg 7320
tggcagacga gcccgaggac ctggacacgg aggacgaggg cctcatcagc ttcgaggagg 7380
agcgggccca gctgtccttc aacacggaca cgctctgctg accacccaga gctgggccag 7440
ggaggacacg ctccactgac cacccagagc tgggccaggg actcaacaat ggggacagaa 7500
gtcccccagt gcctgccagg gcctggagtg gaggttcagg accaaggggc ttctggtcct 7560
ccagcccctg tactcggcca tgtcctgcgg tcactgcggt tgccggccct 7610

<210> 29
<211> 2219
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1560619CB1

<400> 29
ggcagcatga gccgatcacc cctcaatccc agccaactcc gatcagtggg ctcccaggat 60
gccctggccc ccttgcctcc acctgctccc cagaatccct ccacccactc ttgggaccct 120
ttgtgtggat ctctgccttg gggcctcagc tgtcttctgg ctctgcagca tgtcttggtc 180
atggcttctc tgctctgtgt ctcccacctg ctcctgcttt gcagtctctc cccaggagga 240
ctctcttact ccccttctca gctcctggcc tccagcttct tttcatgtgg tatgtctacc 300
atcctgcaaa cttggatggg cagcaggctg cctcttgtcc aggctccatc cttagagttc 360
cttatccctg ctctggtgct gaccagccag aagctacccc gggccatcca gacacctgga 420
aactcctccc tcatgctgca cctttgtagg ggacctagct gccatggcct ggggcactgg 480
aacacttctc tccaggaggt gtccggggca gtggtagtat ctgggctgct gcagggcatg 540
atggggctgc tggggagtcc cggccacgtg ttcccccact gtgggcccct ggtgctggct 600
cccagcctgg ttgtggcagg gctctctgcc cacagggagg tagcccagtt ctgcttcaca 660
cactgggggt tggccttgct ggttatcctg ctcatggtgg tctgttctca gcacctgggc 720
tcctgccagt ttcatgtgtg cccctggagg cgagcttcaa cgtcatcaac tcacactcct 780
ctccctgtct tccggctcct ttcggtgctg atcccagtgg cctgtgtgtg gattgtttct 840
gcctttgtgg gattcagtgt tatcccccag gaactgtctg cccccaccaa ggcaccatgg 900

atttggctgc ctcacccagg tgagtggaat tggcctttgc tgacgcccag agctctggct 960
gcaggcatct ccatggcctt ggcagcctcc accagttccc tgggctgcta tgccctgtgt 1020
ggccggctgc tgcatttgcc tccccacct ccacatgcct gcagtcgagg gctgagcctg 1080
gaggggctgg gcagtgtgct ggccgggctg ctgggaagcc ccatgggcac tgcatccagc 1140
ttccccaacg tgggcaaagt gggtcttatc caggctggat ctcagcaagt ggctcactta 1200
gtggggctac tctgcgtggg gctttggactc tcccccaggt tggctcagct cctcaccacc 1260
atcccactgc ctgttgttgg tggggtgctg ggggtgaccc aggctgtggt tttgtctgct 1320
ggattctcca gcttctacct ggctgacata gactctgggc gaaatatctt cattgtgggc 1380
ttctccatct tcatggcctt gctgctgcca agatggtttc gggaagcccc agtcctgttc 1440
agcacaggct ggagcccctt ggatgtatta ctgcactcac tgctgacaca gcccatcttc 1500
ctggctggac tctcaggctt cctactagag aacacgattc ctggcacaca gcttgagcga 1560
ggcctaggtc aagggctacc atctcctttc actgcccaag aggctcgaat gcctcagaag 1620
cccagggaga aggctgctca agtgtacaga cttcctttcc ccatccaaaa cctctgtccc 1680
tgcatccccc agcctctcca ctgcctctgc ccactgcctg aagaccctgg ggatgaggaa 1740
ggaggctcct ctgagccaga agagatggca gacttgctgc ctggctcagg ggagccatgc 1800
cctgaatcta gcagagaagg gtttaggtcc cagaaatgac cagaacgcct acttctgccc 1860
tggttaattt agccctaact ctcatctgct ggagagtcag ctcccaaact gttctttctt 1920
gtaggcagag gatatgtgtg tgtgtattac atgggactgt ctagaggttc catttcccaa 1980
tagggtgggt tgcctttcct tgtcttaatt aggcctaact gttccagagc agaggccatg 2040
atttagtgga ccatgaatga ttgagatttt gcctgtgtac tatcaatgcc acttgaacca 2100
cagcattcac tttaatactt actgagcatc tcccatgtgc aaggtcctgg aactacaggg 2160
ataagacagg gtccatgccg tctcaaggca tttacggttt aaaaagacct ttgtaatta 2219

<210> 30
<211> 1280
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2614283CB1

<400> 30
ccgcaggagc cgggccggag tgagcgcacc tcgcggggcc cctcggggca ggtgggtgag 60
cgccacccgg agtcccgcgc gcaactttca gggcgcactc ggcggggcgg ctgcgcggct 120
gccgggactc ggcgcgggac tgcatggagg ccaaggagaa gcagcatctg ttggacgcca 180
ggccggcaat ccggtcatac acgggatctc tgtggcagga aggggctggc tggattcctc 240
tgccccgacc tggcctggac ttgcaggcca ttgagctggc tgcccagagc aaccatcact 300
gccatgctca gaagggtcct gacagtcact gtgaccccaa gaaggggaag gcccagcgcc 360
agctgtatgt agcctctgcc atctgcctgt tgttcatgat cggagaagtc gttggtgggt 420
acctggcaca cagcttggct gtcatgactg acgcagcaca cctgctcact gactttgcca 480
gcatgctcat cagcctcttc tccctctgga tgtcctcccg gccagccacc aagaccatga 540
actttggctg gcagagagct gagatcttgg gagccctggt ctctgtactg tccatctggg 600
tcgtgacggg ggtactggtg tacctggctg tggagcggct gatctctggg gactatgaaa 660
ttgacggggg gaccatgctg atcacgtcgg gctgcgctgt ggctgtgaac atcataatgg 720
ggttgaccct tcaccagtct ggccatgggc acagccacgg caccaccaac cagcaggagg 780
agaaccccag cgtccgagct gccttcatcc atgtgatcgg cgactttatg cagagcatgg 840
gtgtcctagt ggcagcctat attttatact tcaagccaga atacaagtat gtagacccca 900
tctgcacctt cgtcttctcc atcctggtcc tggggacaac cttgaccatc ctgagagatg 960
tgatcctggt gttgatggaa gggaccccca agggcgttga cttcacagct gttcgtgatc 1020
tgctgctgtc ggtggagggg gtagaagccc tgcacagcct gcatatctgg gcactgacgg 1080
tggcccagcc tgttctgtct gtccacatcg ccattgctca gaatacagac gcccaggctg 1140
tgctgaagac agccagcagc cgcctccaag ggaagttcca cttccacacc gtgaccatcc 1200
agatcgagga ctactcggag gacatgaagg actgtcaggc atgccagggc ccctcagact 1260
gactgctcag ccaggcacca 1280

<210> 31
<211> 2727
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2667691CB1

<400> 31
cagtagtggt gggacggcac tagctgctgg ggcctgccgc cccgggagtg gctgcagcag 60
cgccaggaat cgaggatggt aaaatgaccc aggggaagaa gaagaaacgg gccgcgaacc 120
gcagtatcat gctggccaag aagatcatca ttaaggacgg aggcacgcct caaggaatag 180
gttctcctag tgtctatcat gcagttatcg tcatcttttt ggagtttttt gcttggggac 240
tattgacagc acccaccttg gtggtattac atgaaacctt tcctaaacat acatttctga 300
tgaacggctt aattcaagga gtaaagggtt tgttgtcatt ccttagtgcc ccgcttattg 360
gtgctctttc tgatgtttgg ggccgaaaat ccttcttgct gctaacggtg tttttcacat 420
gtgccccaat tcctttaatg aagatcagcc catggtggta ctttgctgtt atctctgttt 480
ctgggggttt tgcagtgact ttttctgtgg tatttgcata cgtagcagat ataacccaag 540
agcatgaaag aagtatggct tatggactgg tttcagcaac atttgctgca agtttagtca 600
ccagtcctgc aattggagct tatcttggac gagtatatgg ggacagcttg gtggtggtct 660
tagctacagc aatagctttg ctagatattt gttttatcct tgttgctgtg ccagagtcgt 720
tgcctgagaa aatgcggcca gcatcctggg gagcacccat ttcctgggaa caagctgacc 780
cttttgcgtc cttaaaaaaa gtcggccaag attccatagt gctgctgatc tgcattacag 840
tgtttctctc ctacctaccg gaggcaggcc aatattccag ctttttttta tacctcagac 900
agataatgaa attttcacca gaaagtgttg cagcgtttat agcagtcctt ggcattcttt 960
ccattattgc acagaccata gtcttgagtt tacttatgag gtcaattgga aataagaaca 1020
ccattttact gggtctagga tttcaaatat tacagttggc atggtatggc tttggttcag 1080
aaccttggat gatgtgggct gctggggcag tagcagccat gtctagcatc acctttcctg 1140
ctgtcagtgc acttgtttca cgaactgctg atgctgatca acagggtgtc gttcaaggaa 1200
tgataacagg aattcgagga ttatgcaatg gtctgggacc ggccctctat ggattcattt 1260
tctacatatt ccatgtggaa cttaaagaac tgccaataac aggaacagac ttgggaacaa 1320
acacaagccc tcagcaccac tttgaacaga attccatcat ccctggccct cccttcctat 1380
ttggagcctg ttcagtactg ctggctctgc ttgttgcctt gtttattccg gaacatacca 1440
atttaagctt aaggtccagc agttggagaa agcactgtgg cagtcacagc catcctcata 1500
atacacaagc gccaggagag gccaaagaac ctttactcca ggacacaaat gtgtgacgac 1560
tgaaatcagg aagatttttc tatcagcacc caggtcttag ttttcacctc tagttctgga 1620
tgtacattcc atttccatcc acagtgtact ttaagattgt cttaagaaat gtatctgcat 1680
gaactccgtg ggaactaaag gaagtgggaa cttagaacca gacagttttc caaagatgtt 1740
acaatttctt ttgaaaaacc ttttgtttat tagcaccaat ttcttgccac taagctattt 1800
gttttattat acatccttta attaaaaact atatatgtaa cttcttagat attagcaaat 1860
gtctctgcta ccatttcctt aaggtgttga gctttaactc tatgctgact cagtgagaca 1920
cagtaggtag tatggttgtg gacctatttg ttttaacatt gtaaaatttt gagtcagatt 1980
ttaatattgt aaaatcttgg gtcaaataat tcaaagcctt aatgcagatg cactaaaaca 2040
aagaaatggt aaatgaattg tttgcattta aaaaaaaaaa ctcttaagaa aactgtacta 2100
aatctgaatc atgttttgag cttgtttgca gtacttttaa acattattca ctactgtttt 2160
tgaagtgaga aagtatcagc catttagcat ttaagttggg gtatttagag cctgtaatct 2220
aaatgctggc tcaaatttat tccccagcta cttcttatac cactattctt ttaatgtttg 2280
cataatcata agcacctcaa cacttgaata cataatctaa aaattatata gtaaagctgg 2340
tagccttgaa aatgtcagtg tgatatctat tatgtagata aatatatata gtggcctttc 2400
aggactgtca cagtaacact ttatttacag agctaatgtt tgtcctaaat tttcaggacc 2460
ctagaggaga gctttataca attaccgatg tgaatttctc taaagtgtat atttttgtgt 2520
ccagttatat tatttaaaaa agtgttactt tgtaaaaatt gtatataaag aactgtatag 2580
tttacactgt tttcatcttg tgtgtggtta ttgcttaatg ctttttaaac ttggaacact 2640
cactatggtt aaataaggtc ttaaaagaaa tgtaaatatt ctgttaataa agttaaatat 2700
tttaatgatt tttttttaa aaaaaaa 2727

<210> 32
<211> 1631
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3211415CB1

<400> 32
ttgcgcttga atgttcgttg actggcccgt cggttttaca aggcccggac aagcgctggg 60
gattcccgtt tgaggcgtca ctactgtcac tgccatcacc ccacggagcc acttctagag 120
gggagtagac ccggcccttc gccgggcaga gaagatgttg cccctgtcca tcaaagacga 180
tgaatacaaa ccacccaagt tcaatttgtt cggcaagatc tcgggctggt ttaggtctat 240
actgtccgac aagacttccc ggaacctgtt tttcttcctg tgcctgaacc tctctttcgc 300
ttttgtggaa ctactctacg gcatctggag caactgctta ggcttgattt ccgactcttt 360
tcacatgttt ttcgatagca ctgccatttt ggctggactg gcagcttctg ttatttcaaa 420
atggagagat aatgatgctt tctcctatgg gtatgttaga gcggaagttc tggctggctt 480
tgtcaatggc ctatttttga tcttcactgc ttttttttatt ttctcagaag gagttgagag 540
agcattagcc cctccagatg tccaccatga gagactgctt cttgtttcca ttcttgggtt 600
tgtggtaaac ctaataggaa tatttgtttt caaacatgga ggtcatggac attctcatgg 660
ctctggtggc cacggacaca gtcattccct ctttaatggt gctctagatc aggcacatgg 720
ccatgtcgat cattgccata gccatgaagt gaaacatggt gctgcacata gccatgatca 780
tgctcatgga catggacact ttcattctca tgatggcccg tccttaaaag aaacaacagg 840
acccagcaga cagattttac aaggtgtatt tttacatatc ctagcagata cacttggaag 900
tattggtgta attgctttctg ccatcatgat gcaaaatttt ggtctgatga tagcagatcc 960
tatctgttca attcttatag ccattcttat agttgtaagt gttattcctc ttttaagaga 1020
atctgttgga atattaatgc agagaactcc tcccctatta gaaaatagtc tgcctcagtg 1080
ctatcagagg gtacagcagt tgcaaggagt ttacagttta caggaacagc acttctggac 1140
tttatgttct gacgtttatg ttgggacctt gaaattaata gtagcacctg atgctgatgc 1200
taggtggatt ttaagccaaa cacataatat ttttactcag gctggagtga gacagctcta 1260
cgtacagatt gactttgcag ccatgtagtg aatggaaaga aattatgcac cttttatgga 1320
ccaaattttt ctggcccaac ccgatgagat gaagcatttc aaacttgagg agaagagaga 1380
ctgcaggacg aggtggacag aaaaaccgtc agtaacaccg aggacatcta aaatctagtg 1440
atgaacacga tgacaaccaa agtagcacaa gagaagaggc aagtcacagg gaggggtttg 1500
ggaaccttat gtcacgctta agaactggga tgggaggttt taaacaaaaa aaaaaaaaaa 1560
aaaaggggcg gccgccgact attgaaccct tcgcccgggg aataaattcc ggcccggtac 1620
ctcgaggggg g 1631

<210> 33
<211> 2673
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4739923CB1

<400> 33
ggacatttta aaagggccgg agattgcggg cgtcagtggc catggcggat acagcgacta 60
cagcatcggc ggcggcggct agtgccgcta gcgcctcgag cgatgcacct ccttccaaac 120
tgggcaaacc ccgcttccag cagacgtcct tctatggccg cttcaggcac ttcttggata 180
tcatcgaccc tcgcacactc tttgtcactg agagacgtct cagagaggct gtgcagctgc 240
tggaggacta taagcatggg accctgcgcc cgggggtcac caatgaacag ctctggagtg 300
cacagaaaat caagcaggct attctacatc cggacaccaa tgagaagatc ttcatgccat 360
ttagaatgcc aggttatatt ccttttggga cgccaattgt agtcggtctt ctcttgccca 420

accagacact ggcatccact gtcttctggc agtggctgaa ccagagccac aatgcctgtg 480
tcaactatgc aaaccgcaat gcgaccaagc cttcacctgc atccaagttc atccagggat 540
acctgggagc tgtcatcagc gccgtctcca ttgctgtggg ccttaatgtc ctggttcaga 600
aagccaacaa gctcacccca gccacccgcc ttctcatcca gaggtttgtg ccgttccctg 660
ctgtagccag tgccaatatc tgcaatgtgg tcctgatgcg gtacggggag ctggaggaag 720
ggattgatgt cctggacagc gatggcaacc tcgtgggctc ctccaagatc gcagcccgac 780
acgccctgct ggagacggcg ctgacgcgag tggtcctgcc catgcccatc ctggtgctac 840
ccccgatcgt catgtccatg ctggagaaga cggctctcct gcaggcacgc ccccggctgc 900
tcctccctgt gcaaagcctc gtgtgcctgg cagccttcgg cctggccctg ccgctggcca 960
tcagcctctt cccgcaaatg tcagagattg aaacatccca attagagccg gagatagccc 1020
aggccacgag cagccggaca gtggtgtaca acaaggggtt gtgagtgtgg tcagcggcct 1080
ggggacggag cactgtgcag ccggggagct gaggggcagg gccgtagact cacggctgca 1140
cctgcaggga gcagcacgcc aaccccagca gtcctgggcc ccctgggaga gtgctcaacc 1200
tacagtggag ggagactgac ccattcacat tttaacatag gcaagaggag ttctaacaca 1260
tttcgtacaa aaaaataatc aagtgcattt ctgggcctta tgtggggttg tcaaaactcc 1320
actcagcaca attatgtgtg aagctgaaaa attgtagagt gcccatgggg tagaagtaga 1380
atccttttat actttggttc ctttttatt tttatttttt atcagaatca aatctgagcc 1440
ttagtttcag ctgaccagaa gtggcaggag gacaggtgga ggcgagccag attaggcctg 1500
gagttgggct ggtttggtgg ccaggctggt aaatttagga tattacaatg gccagcccag 1560
atggctccct gggctggcat ggggagggga gagaaggtgg tctgcacccc acaggatgaa 1620
ctagccatga ctagggtcct caggcagtgg cccagggaat cagggagcac tggaggcctc 1680
tgcaagattc tgtgggcagc tggctctgaa tgtagccagc ccacatccat tccagagctg 1740
cagaaccagt ccctctgagt gaagtccagt gaccctggag ctagggtccc ccttcgtgga 1800
gctctaactg attcagggcc cctgaagtga ccccagctcc aggcaggaaa ccccgagaag 1860
gaatggtgct tggcaggaac catggacctg cacttggcct cttcgggaag atcctccttc 1920
caggccccag gctggatgct gggttctggg gctgagagtg gggctagact gggctggctg 1980
ccttctgcgg agcttctcca gccaccaagg ctgcctgccc catcccatcc ctttctaagc 2040
aggaaggtct catgcctgag aatgcctagg ccagctcctt agacacctat cagagaagca 2100
gcatcaacct aggagcagtg ggccctggct ctgtcactaa atggcagtga gatgtcagac 2160
aaattcattc ccttctctca gcctccactc cctgtctgaa accagaagac tggatcaagg 2220
ggttcccact ggctcctcca gcaagacctc gtctttgctt gtcctgctca gatgctggtc 2280
atcctgggca tgtccccagt gtggactctg gactgggaag ggggcaggcc cctttggacc 2340
tgcagttggc ctcagcagaa ggccttgcct tgtgtatgtg actccatatc ccgggagcag 2400
ttgacctttg ccaaacactt tacagttctg gaggaggagg taacatagat gcctgggcct 2460
gatggtgggg ccatacccat gtgtcgcctc tcactctggc agcctcagag gccccttgct 2520
gctggctccc atctccctcc catttgcaga ccaggaagga agagcaagct gtacaaaggg 2580
aagcagagcc tggggtgggt gtgagcaggg tgacccctca tctgaaaggc ccaaaccagg 2640
gggaagcacc agcctcagtg cagccccctc ctg 2673

<210> 34
<211> 3958
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 55030459CB1

<400> 34
atggcccgcc agccggagga agaggagacg gccgtggccc gggcgcggcg gccgcccctc 60
tggctgctct gcctggtcgc gtgctggctc ctgggcgccg gggccgaagc cgacttctcc 120
atcctggacg aggcgcaagt gctggcgagc cagatgcgga ggctggcggc cgaggagctg 180
ggggtcgtca ccatgcagcg gatattcaac tcctttgttt acactgagaa aatctcaaat 240
ggagaaagtg aagtacagca gctagccaaa aaaatccgag agaagttcaa ccgttacttg 300
gatgtggtca atcggaacaa gcaagttgta gaagcatcct atacggctca cctaacctct 360
cccctaactg caattcaaga ctgctgtact atcccacctt ccatgatgga attcgatggg 420

```
aactttaata ccaatgtgtc tagaacaatt agttgtgatc gactttctac tactgttaat 480
agccgggcct tcaatccagg acgagactta aattcagttc ttgcagacaa cctgaaatcc 540
aaccctggaa ttaagtggca atatttcagt tcagaagaag gaattttcac tgttttccca 600
gcacacaagt tccggtgtaa gggcagctac gaacaccgca gtagacccat ctacgtctct 660
acagtccggc cgcagtcaaa gcacatagta gtgattctgg accacggggc ttcagtcaca 720
gacactcagc ttcagattgc caaggacgct gctcagggtca tcctcagcgc catcgatgaa 780
catgacaaga tttctgtgtt aactgtggca gataccgtcc ggacttgctc actagaccag 840
tgctataaga ccttcttgtc tccagccacc agtgagacaa aaaggaaaat gtccacctttt 900
gttagcagcg tgaagtcttc agacagtcct acccagcacg cagtgggatt ccaaaaggca 960
tttcagctga ttcgaagtac aaacaataac acaaagttcc aagcaaatac agacatggtc 1020
atcatttacc tgtcagctgg cattacatca aaggactctt cggaagaaga taaaaaagcg 1080
actctccaag tcatcaatga agaaaatagc tttctaaaca actctgtaat gattctcacc 1140
tatgccctca tgaacgatgg ggtgactggt ttgaaagagc tggcttttct gagggatcta 1200
gctgaacaga attcagggaa gtacggtgtg ccagaccgga cggccttgcc tgtgatttaag 1260
ggcagcatga tggtgctgaa tcagttgagc aacctggaga ccacagtggg caggttctac 1320
acaaaccttc ccaaccggat gattgatgaa gccgtcttca gcctgccctt ctctgatgag 1380
atgggagatg gtttgataat gactgtgagt aaaccctgtt attttggaaa cctacttctg 1440
ggaattgtag gtgtggacgt gaatctggct tacattcttg aagacgtgac gtattaccaa 1500
gactctttgg cttcctatac ttttctcata gacgacaaag gatatacact tatgcaccca 1560
tctcttacca ggccatattt attgtcagag ccccacttc atactgacat catacattat 1620
gaaaatattc caaaatttga attagttcgg caaaatatcc taagcctccc tctgggcagc 1680
cagattatcg cagtccctgt gaactcatcc ctgtcttggc acataaacaa gctgagagaa 1740
actggaaagg aagcctacaa tgttagctat gcctggaaga tggtacaaga cacttccttt 1800
attctgtgta ttgtggtgat acaaccagaa atacctgtga aacaactgaa gaacctcaac 1860
actgttccca gcagcaagct gctgtaccac cggctggatc tccttggcca gcccagtgct 1920
tgcctccact tcaaacagct ggcaacccta gaaagtccca ccatcatgct gtctgctggc 1980
agcttttcct ccccctatga gcacctcagc cagccagaga caaagcgcat ggtagagcac 2040
tacaccgcct atctcagcga caacacccgc ctcattgcta acccgggcct caaattctct 2100
gtcagaaatg aagtaatggc taccagccac gtcacagatg aatggatgac acaaatggaa 2160
atgagtagcc tgaacactta cattgtccgc cgttacatag caacacccaa tggcgtcctc 2220
agaatttatc ctggttccct catggacaaa gcatttgatc ccactaggag acaatggtat 2280
ctccatgcag tagctaatcc agggttgatt tctttgactg gtccttactt agatgttgga 2340
ggagctggtt atgttgtgac aatcagtcac acaattcatt catccagtac acagctgtct 2400
tctgggcaca ctgtggctgt gatgggcatt gacttcacac tcagatactt ctacaaagtt 2460
ctgatggacc tattacctgt ctgtaaccaa gatggtggca acaaaatаag gtgcttcata 2520
atggaggaca ggggttatct ggtggcgcac ccgactctca tcgaccccaa aggacatgca 2580
cctgtggagc agcagcacat cacccacaag gagcccctgg tagcaaatga tatcctcaac 2640
caccccaact ttgtaaagaa aaacctgtgc aacagcttca gtgacagaac ggtccagagg 2700
ttttataaat tcaacaccag ccttgcgggg gatttgacga accttgtgca tggcagccac 2760
tgttccaaat acagattagc aaggatccca ggaaccaacg cgtttgttgg cattgtcaac 2820
gaaacctgcg actctcttgc cttctgtgcc tgcagcatgg tggaccgact ctgtctcaac 2880
tgtcaccgaa tggaacaaaa tgaatgtgaa tgtccttgtg agtgccctct agaggtcaat 2940
gagtgcactg gcaaccctcac caatgcagag aaccgaaacc ccagctgcga ggtccaccag 3000
gagccggtga catacacagc tattgaccct ggcctgcaag atgctcttca ccagtgtgtc 3060
aacagcaggt gcagtcagag gctggaaagt ggggactgtt ttggggtgct ggattgtgaa 3120
tggtgcatgg tggacagtga tggaaagact cacctggaca aaccctactg tgccccccag 3180
aaagaatgct tcgggggggat tgtgggagcc aaaagtccct acgttgatga catgggagca 3240
ataggtgatg aggtgatcac attaaacatg attaaaagcg cccctgtggg tcctgtggct 3300
ggagggatca tgggatgcat catggtcttg gtcctggcgg tgtatgccta ccgccaccag 3360
attcatcgcc ggagccatca gcatatgtct cctcttgctg cccaagaaat gtcagtgcgt 3420
atgtccaacc tggagaatga cagagatgaa agggacgacg acagccacga agacagaggc 3480
atcatcagca acactcggtt tatagctgcg gtcatcgaac gacatgcaca cagtccagaa 3540
agaaggcgcc gctactgggg tcgatcagga acagaaagtg atcatggtta cagcaccatg 3600
agcccacagg aggacagtga aaatcctcca tgcaacaatg accccttgtc agccggggtc 3660
gatgtgggaa accatgatga ggacttagac ctggataccc cccctcagac tgctgcccta 3720
ctaagtcaca agttccacca ctaccggtca caccaccctа cacttcatca tagccaccac 3780
```

ttacaggcgg ccgtcacggt acacactgtc gatgcagaat gctaacaatc tcctcacctc 3840
cacgccaaga tgagatctgg gagctacaga atgttctgga aagaaaaaga accggcttaa 3900
aacccacaag cagagacctc ccttgtgttt gtgctttgtg cagagttgtt tgagtcat 3958

<210> 35
<211> 2000
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6113039CB1

<400> 35
gctcaggaca atgaaattct tcagttacat tctggtttat cgccgatttc tcttcgtggt 60
tttcactgtg ttggttttac tacctctgcc catcgtcctc cacaccaagg aagcagaatg 120
tgcctacaca ctctttgtgg tcgccacatt ttggctcaca gaagcattgc ctctgtcggt 180
aacagctttg ctacctagtt taatgttacc catgttttggg atcatgcctt ctaagaaggt 240
ggcatctgct tatttcaagg attttcactt actgctaatt ggagttatct gtttagcaac 300
atccatagaa aaatggaatt tgcacaagag aattgctctg aaaatggtga tgatggttgg 360
tgtaaatcct gcatggctga cgctggggtt catgagcagc actgcctttt tgtctatgtg 420
gctcagcaac acctcgacgg ctgccatggt gatgcccatt gcggaggctg tagtgcagca 480
gatcatcaat gcagaagcag aggtcgaggc cactcagatg acttacttca acggatcaac 540
caaccacgga ctagaaattg atgaaagtgt taatggacat gaaataaatg agaggaaaga 600
gaaaacaaaa ccagttccag gatacaataa tgatacaggg aaaatttcaa gcaaggtgga 660
gttggaaaag aactcaggca tgagaaccaa atatcgaaca aagaagggcc acgtgacacg 720
taaacttacg tgtttgtgca ttgcctactc ttctaccatt ggtggactga caacaatcac 780
tggtacctcc accaacttga tctttgcaga gtatttcaat acacgctatc ctgactgtcg 840
ttgcctcaac tttggatcat ggtttacgtt ttccttccca gctgcccttа tcattctact 900
cttatcctgg atctggcttc agtggctttt cctaggattc aattttaagg agatgttcaa 960
atgtggcaaa accaaaacag tccaacaaaa agcttgtgct gaggtgatta agcaagaata 1020
ccaaaagctt gggccaataa ggtatcaaga aattgtgacc ttggtcctct tcattataat 1080
ggctctgcta tggtttagtc gagaccccgg atttgttcct ggttggtctg cacttttttc 1140
agagtaccct ggttttgcta cagattcaac tgttgctttа cttatagggc tgctattctt 1200
tcttatccca gctaagacac tgactaaaac tacacctaca ggagaaattg ttgcttttga 1260
ttactctcca ctgattactt ggaaagaatt ccagtcattc atgccctggg atatagccat 1320
tcttgttggt ggagggtttg ccctggcaga tggttgtgag gagtctggat tatctaagtg 1380
gataggaaat aaattatctc ctctgggttc attaccagca tggctaataa ttctgatatc 1440
ttctttgatg gtgacatctt taactgaggt agccagcaat ccagctacca ttacactctt 1500
tctcccaata ttatctccat tggccgaagc cattcatgtg aaccctcttt atattctgat 1560
accttctact ctgtgtactt catttgcatt cctcctacca gtagcaaatc cacccaatgc 1620
tattgtcttt tcatatggtc atctgaaagt cattgacatg gttaaagctg gacttggtgt 1680
caacattgtt ggtgttgctg tggttatgct tggcatatgt acttggattg tacccatgtt 1740
tgacctctac acttaccctt cgtgggctcc tgctatgagt aatgagacca tgccataata 1800
agcacaaaat ttctgactat cttgcggtaa tttctggaag acattaatga ttgactgtaa 1860
aatgtggctc taaataacta atgacacaca tttaaatcag ttatggtgta gctgctgcaa 1920
ttcccgtgaa tacccgaaac ctgctgttat aactcagagt ccatatttgt tattgcagtg 1980
caactaaaga gcatctatgt 2000

<210> 36
<211> 1997
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7101781CB1

<400> 36
ctgctcgcca ctggccggcg cgctcccggc gcacggagca cactcgcgct cccggcgcac 60
ggagcacact cgcgctccgg gactgaaacc tgagcagccg tagcagccga atttgggagc 120
atatccttgt cactgcagcc agaaagccct tcgatcccca tcagagaggt cacatgagcc 180
ccgaggtcac ctgccccgcgg aggggccacc tgcctcgctt ccacccgagg acctgggttg 240
agcccgtggt ggcatcgtcc caggtggctg cctccctcta cgatgcgggg ctactcctcg 300
tggtgaaggc gtcctacgga accggaggct cctccaacca cagtgccagc ccatcgcccc 360
gggggggctct agaggaccaa cagcagagag ccatctccaa tttctacatt atctacaacc 420
ttgtggtggg cctgtccccc ctgctgtccg cctacgggct gggatggctc agcgaccgct 480
accaccgaaa gatctccatc tgcatgtcgc tgctgggctt cctgctctcc cgcctcgggc 540
tgctgctcaa ggtgctgctg gactggccag tggaggtgct gtacggggcg gcggcgctga 600
acgggctatt cggcggcttc tccgccttct ggtccggggt catggcgctg ggatcgctgg 660
gctcctccga gggccgccgc tctgtgcgcc tcatcctcat tgacctgatg ctgggcttgg 720
cggggttctg cgggagcatg gcttccgggc atctcttcaa gcagatggct gggcactctg 780
ggcagggcct gatactgacg gcctgcagcg tgagctgtgc ctcgtttgcc ctgctctaca 840
gcctttggt gctaaaggtc cctgagtcgg tggccaaacc cagccaggag ctccccgccg 900
tggataccgt gtctggcacg gttggcacat accgcactct ggatcctgat cagttggacc 960
aacagtatgc agtggggcac cctccatctc ctggaaaagc aaaaccccat aaaaccacca 1020
ttgccttgct ctttgtgggt gctatcatat atgacctggc ggtggtgggc acagtggacg 1080
tgatcctct ttttgtgctg agggagcctc tcggttggaa ccaagtgcag gtgggctatg 1140
gtatggctgc agggtacacc atcttcatca ccagcttcct gggtgtcctg gtcttctccc 1200
gctgctttcg ggacaccacc atgatcatga ttgggatggt ctcctttggg tcaggagccc 1260
tcctcttggc ttttgtgaaa gagacataca tgttctatat tgctcgagcc gtcatgctgt 1320
ttgctctcat ccccgtcaca accatccgat cagctatgtc caaactcata aagggctcct 1380
cttatggaaa ggtgttcgtc atactgcagc tgtccttggc tctgaccggc gtggtgacat 1440
ccaccttgta caacaagatc taccagctca ccatggacat gtttgtgggc tcctgctttg 1500
ctctctcctc ctttctctcc trcctggcca tcattccaat tagcatcgtg gcctataaac 1560
aagtcccatt gtcaccatat ggagacatca tagagaaatg aagatgctta cctgcaggaa 1620
ctgaaaacat cagccatggc caggccccca gaagacaaaa gaagggacca gggaactggt 1680
gacctaagca acccactgct taagaaacct gcgttccagc cagagttggc ctcagaatga 1740
cctgctctgg ctcagggatc cctggtggat ggggaaaagc actttcctgg tgatggaaaa 1800
acgttctcag ctttaagaca cccccattag gcagacactg ggttctgtga cagcagagca 1860
tgaccttaag ggttacaggg aggctgcaca cagcagtccc aggccctgtg aggggcttca 1920
gactccagct acagcgagcc tgccccttt cttcaaggga ctgtcttgag ggctccaaag 1980
tatagctaac tagtcac 1997

<210> 37
<211> 3069
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7473036CB1

<400> 37
atgcaaccag ccagagggcc cctggcttca gaacctagga ctgtactggt tctgagattc 60
tgtgcaagcc tcatggaaat gaagctgcca ggccaggaag ggtttgaagc ctccagtgct 120
cctagaaata ttccttcagg ggagctggac agcaaccctg accctggcac cggccccagc 180
cctgatggcc cctcagacac agagagcaag gaactgggag tacccaaaga ccctctgctc 240
ttcattcagc tgaatgagct gctgggctgg ccccaggcgc tggagtggag agagacaggc 300
aggtgggtac tgtttgagga gaagttggag gtggctgcag gccggtggag tgcccccac 360
gtgcccaccc tggcactgcc cagcctccag aagctccgca gcctgctggc cgagggcctt 420
gtactgctgg actgcccagc tcagagcctc ctggagctcg tgggctctac tcatccaaga 480

aaggcttctg acaatgagga agcccccctg agggaacagt gtcagaaccc cctgagacag 540
aagctacctc caggagctga ggcagggact gtgctggcag gggagctggg cttcctggca 600
cagccactgg gagcctttgt tcgactgcgg aaccctgtgg tactggggtc ccttactgag 660
gtgtccctcc caagcaggtt tttctgcctt ctcctgggcc cctgtatgct gggaaagggc 720
taccatgaga tgggacgggc agcagctgtc ctcctcagtg acccgcaatt ccagtggtca 780
gttcgtcggg ccagcaacct tcatgacctt ctggcagccc tggatgcatt cctagaggag 840
gtgacagtgc ttcccccagg tcggtgggac ccaacagccc ggattccccc gcccaaatgt 900
ctgccatctc agcacaaaag gcttccctcg caacagcggg agatcagagg tcccgccgtc 960
ccgcgcctga cctcggctga ggacaggcac cgccatgggc cacacgcaca cagcccggag 1020
ttgcagcgga ccggcagcga tttcttggac gccctgcatc tccagtgctt ctcggccgta 1080
ctctacattt acctggccac tgtcactaat gccatcactt ttgggggtct gctgggagat 1140
gccactgatg gtgcccaggg agtgctggaa agtttcctgg gcacagcagt ggctggagct 1200
gccttctgcc tgatggcagg ccagcccctc accattctga gcagcacggg gccagtgctg 1260
gtctttgagc gcctgctctt ctctttcagc agagattaca gcctggacta cctgcccttc 1320
cgcctatggg tgggcatctg ggtggctacc ttttgcctgg tgctggtggc cacagaggcc 1380
agtgtgctgg tgcgctactt cacccgcttc actgaggaag gtttctgtgc cctcatcagc 1440
ctcatcttca tctacgatgc tgtgggcaaa atgctgaact tgacccatac ctatcctatc 1500
cagaagcctg ggtcctctgc ctacgggtgc ctctgccaat acccaggccc aggaggaaat 1560
gagtctcaat ggataaggac aaggccaaaa gacagagacg acattgtaag catggactta 1620
ggcctgatca atgcatcctt gctgccgcca cctgagtgca cccggcaggg aggccaccct 1680
cgtggccctg gctgtcatac agtcccagac attgccttct tctcccttct cctcttcctt 1740
acttctttct tctttgctat ggccctcaag tgtgtaaaga ccagccgctt cttcccctct 1800
gtggtgcgca aagggctcag cgacttctcc tcagtcctgg ccatcctgct cggctgtggc 1860
cttgatgctt tcctgggcct agccacacca aagctcatgg tacccagaga gttcaagccc 1920
acactccctg ggcgtggctg gctggtgtca ccttttggag ccaacccctg gtggtggagt 1980
gtggcagctg ccctgcctgc cctgctgctg tctatcctca tcttcatgga ccaacagatc 2040
acagcagtca tcctcaaccg catggaatac agactgcaga agggagctgg cttccacctg 2100
gacctcttct gtgtggctgt gctgatgcta ctcacatcag cgcttggact gccttggtat 2160
gtctcagcca ctgtcatctc cctggctcac atggacagtc ttcggagaga gagcagagcc 2220
tgtgcccccg gggagcgccc caacttcctg ggtatcaggg aacagaggct gacaggcctg 2280
gtggtgttca tccttacagg agcctccatc ttcctggcac ctgtgctcaa gttcattcca 2340
atgcctgtgc tctatggcat cttcctgtat atgggggtgg cagcgctcag cagcattcag 2400
ttcactaata gggtgaagct gttgttgatg ccagcaaaac accagccaga cctgctactc 2460
ttgcggcatg tgcctctgac cagggtccac ctcttcacag ccatccagct tgcctgtctg 2520
gggctgcttt ggataatcaa gtctaccctc gcagccatca tcttcccctc catgttgctg 2580
ggccttgtgg gggtccgaaa ggccctggag agggtcttct caccacagga actcctctgg 2640
ctggatgagc tgatgccaga ggaggagaga agcatccctg agaaggggct ggagccagaa 2700
cactcattca gtggaagtga cagtgaagat tcagagctga tgtatcagcc aaaggctcca 2760
gaaatcaaca tttctgtgaa ttagctggag taggagtctg ggagtggaga ccccaggaaa 2820
cagcatgagt tcacaggtgc ttactcagga agtcaggaca tttttggcct ttggcttaac 2880
ttccagatgc tcagtcggct tggggaagga ctgaagggca gctgccaaga cctcagttac 2940
ctcctgacct gagggtggag agtggcagga agcaagcatg tttgctgtgc acttaggaaa 3000
ggctggtgag ccagagggac tgatcaggcc ccattcactc tctactcatt aaaaggtcct 3060
gagccacaa 3069

<210> 38
<211> 2241
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7476943CB1

<400> 38
gccggcggcc cccgctcccg gatccccagc gccctggcca agaagcttcc tcggctcccc 60

ctcttccctc tccctgacac ggttgtgcag agggcgcggt ggctcaggcc ctggcaacca 120
ccattctact ttttgtgtct atgagtttga ctaccctaag gacctcacat ggcgagtaac 180
ccatgggcca ggtagcgttc tatgccaacc ttgaatgcca tcaggaagtc actggacagc 240
aaactcttcc aagatcataa cttggctgtt ggagcaacct ggaaaagaag aaaaaagaaa 300
aaccatggca aaagtaaata gagctcggtc tacctcccct ccagatggag gctggggctg 360
gatgattgtg gctggctgtt tccttgttac catctgcaca cgggcagtca caagatgtat 420
ctcaattttt tttgtggagt tccagacata cttcactcag gattacgcac aaacggcatg 480
gatccattcc attgtagatt gtgtgaccat gctctgtgct ccacttggga gtgttgtcag 540
taaccattta tcctgtcaag tgggaatcat gctgggtggc ttgcttgcat ctactggact 600
catcctgagc tcatttgcca cgagtctgaa gcatctctac ctcactctgg gagttcttac 660
aggtcttgga tttgcacttt gttactctcc agctattgcc atggttggca agtacttcag 720
cagacggaaa gcccttgctt atggtatcgc catgtcagga agtggcattg gcaccttcat 780
cctggctcct gtggttcagc tccttattga acagttttcc tggcggggag ccttactcat 840
tcttgggggc tttgtcttga atctctgtgt atgtggtgcc ttgatgaggc caattactct 900
taaagaggac cacacaactc cagagcagaa ccatgtgtgt agaactcaga aagaagacat 960
taagcgggtg tctccctatt catctttgac caaagaatgg gcacagactt gcctctgttg 1020
ctgtttgcag caagagtaca gtttttttact catgtcagac tttgttgtgt tagccgtctc 1080
cgttctgttt atggcttatg gctgcagccc tctctttgtg tacttggtgc cttatgcttt 1140
gagtgttgga gtgagtcatc agcaagctgc ttttcttatg tccatacttg gagtgattga 1200
cattattggc aatatcacat ttggatggct gaccgacaga aggtgtctga agaattacca 1260
gtatgtttgc tacctctttg ccgtgggaat ggatgggctc tgctatctct gcctcccaat 1320
gcttcaaagt ctccctctgc tcgtgccttt ctcttgtacc tttggctact ttgatggtgc 1380
ctatgtgact ttgatcccag tagtgaccac agagatagtg gggaccacct ctttgtcatc 1440
agcgcttggt gtggtatact tccttcacgc agtgccatac ttggtgagcc cacccatcgc 1500
aggacggctg gtagatacca ccggcagcta cactgcagca ttcctcctct gtggattttc 1560
aatgatattt agttctgtgt tgcttggctt tgctagactt ataaagagaa tgagaaaaac 1620
ccagttgcag ttcattgcca aagaatctga tcctaagctg cagctatgga ccaatggatc 1680
agtggcttat tctgtggcaa gagaattaga tcagaaacat ggggagcctg tggctacagc 1740
agtgcctggc tacagcctca catgaccaaa ggccttgagc cccagaatct tcaggtttga 1800
gagaggtggg gccaccagat tcttcatgtt tctgaaactt tttattttgg cagaaggatt 1860
gccttccaag gaaattatta ttattgtttt gttaacatat taatatttat aagggaaaac 1920
agcacataat aaggaaagct ggactagccc agagccttct catttgggat ttgtgctcat 1980
aactgaactc gtatctttgg tcaatgggca tagctctgta agaaatgtaa ggacacagct 2040
gatataatta gctgtaatta gggataattt caaagcataa ccaaagcaga tgacactggg 2100
cagcagcttt gttccagtct caggcccttc atgttccctc ctcagaaaga aatggaacat 2160
taacgtggta gctttggtta cttggtctgg ttagagaagg aggccagtga gtgggggggtg 2220
aagtgaaaag caaataaagt a 2241

<210> 39
<211> 1593
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8003355CB1

<400> 39
ccttggagct gttgtcccac ccctgtcact gcagagagct gaggcaccat gcatgggggc 60
caggggccgc tgctcctcct gctgctgctg gctgtctgcc tgggggccca gggccggaac 120
caggaggagc gcctgctcgc agacctgatg caaaactacg accccaacct gcggcccgcg 180
gaacgagact cggatgtggt caatgtcagc ctgaagctaa ccctcaccaa cctcatctcc 240
ctgaacgagc gagaggaagc cctcaccacc aatgtctgga tagaggtgca gtggtgcgac 300
tatcgcctgc gccgggatcc gcgagactac gaaggcctgt gggtgctgag ggtgccgtcc 360
accatggtgt ggcggccgga tatcgtgctg gagaacaacg cggacggtgt cttcgaggtg 420
gccctctact gcaatgtgct cgtgtccccc gacggctgta tctactggct gccgcctgcc 480

atcttccgtt ccgcctgctc tatctcagtc acctacttcc ccttcgactg gcagaactgc 540
tcccttatct tccagtccca gacttacagc accaatgaga ttgatctgca gctgagtcag 600
gaagatggcc agaccatcga gtggattttc attgaccctg aggccttcac agagaatggg 660
gagtgggcca tccagcaccg accagccaag atgctcctgg acccagcggc gccagcccag 720
gaagcaggcc accagaaggt ggtgttctac ctgctcatcc agcgcaagcc cctcttctac 780
gtcatcaaca tcatcgcccc ctgtgtgctc atctcctctg tcgccatcct catccacttc 840
cttcctgcca aggctggggg ccagaagtgt accgtcgcca tcaacgtgct cctggcccag 900
actgtcttcc tcttccttgt ggccaagaag gtgcctgaaa cctcccaggc ggtgccactc 960
atcagcaagt acctgacctt cctcctggtg gtgaccatcc tcattgtcgt gaatgctgtg 1020
gttgtgctca atgtctcctt gcggtctcca cacacacact ccatggcccg aggggtgttc 1080
ctgaggctct tgccccagct gctgaggatg cacgttcgcc cgctggcccc ggcagctgtg 1140
caggacaccc agtcccggct acagaatggc tcctcgggat ggtcgatcac aactggggag 1200
gaggtggccc tctgcctgcc tcgcagtgaa ctcctcttcc agcagtggca gcggcaaggg 1260
ctggtggcgg cagcgctgga gaagctagag aaaggcccgg agttaggggct gagccagttc 1320
tgtggcagcc tgaagcaggc tgccccagcc atccaggcct gtgtggaagc ctgcaacctc 1380
attgcctgtg cccggcacca gcagagtcac tttgacaatg ggaatgagga gtggttcctg 1440
gtgggccgag tgctggaccg cgtctgcttc ctggccatgc tctcgctctt catctgtggc 1500
acagctggca tcttcctcat ggcccactac aaccgggtgc cggccctgcc attccctgga 1560
gatccacgcc cctacctgcc ctcaccagac tga 1593

<210> 40
<211> 2121
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3116448CB1

<400> 40
gtacaaagga cctccagacc agagccagcc agcagcaaaa agagcatgga gctgaggagt 60
acagcagccc ccagagctga gggctacagc aacgtgggct tccagaatga agaaaacttt 120
cttgagaacg agaacacatc aggaaacaac tcaataagaa gcagagctgt gcaaagcagg 180
gagcacacaa acaccaaaca ggatgaagaa caggtcacag ttgagcagga ttctccaaga 240
aacagagaac acatggagga tgatgatgag gagatgcaac aaaaagggtg tttggaaagg 300
aggtatgaca cggtatgtgg tttctgtagg aaacacaaaa caactcttcg gcacatcatc 360
tggggcattt tattagcagg ttatctggtt atggtgattt cggcctgtgt gctgaacttt 420
cacagagccc ttcctctttt tgtgatcacc gtggctgcca tcttctttgt tgtctgggat 480
cacctgatgg ccaaatacga acatcgaatt gatgagatgc tgtctcctgg cagaaggctt 540
ctaaacagcc attggttctg gctgaagtgg gtgatctgga gctccctggt cctagcagtt 600
attttctggt tggcctttga cactgccaaa ttgggtcaac agcagctggt gtccttcggt 660
gggctcataa tgtacattgt cctgttattt ctattttcca agtacccaac cagagtttac 720
tggagacctg tcttatgggg aatcgggcta cagtttcttc ttgggctctt gattctaagg 780
actgaccctg gatttatagc ttttgattgg ttgggcagac aagttcagac ttttctggag 840
tacacagatg ctggtgcttc atttggcttt ggtgagaaat acaaagacca cttctttgga 900
tttaaggtcc tggcgatcgt ggttttcttc agcactgtga tgtccatgct gtactacctg 960
ggactgatgc agtggattat tagaaaggtt ggatggatca tgctagttac tacgggatca 1020
tctcctattg aatctgtagt tgcttctggc aatatatttg ttggacaaac ggagtctcca 1080
ctgctggtcc gaccatattt accttacatc accaagtctg aactccacgc catcatgacc 1140
gccgggttct ctaccattgc tggaagcgtg ctaggtgcat acatttcttt tggggttcca 1200
tcctcccact tgttaacagc gtcagttatg tcagcacctg cgtcattggc tgctgctaaa 1260
ctcttttggc ctgagacaga aaaacctaaa ataaccctca agaatgccat gaaaatggaa 1320
agtggtgatt cagggaatct tctagaagct gcaacacagg gagcatcctc ctccatctcc 1380
ctggtggcca acatcgctgt gaatctgatt gccttcctgg ccctgctgtc ttttatgaat 1440
tcagccctgt cctggtttgg aaacatgttt gactacccac agctgagttt tgagctaatc 1500
tgctcctaca tcttcatgcc cttttccttc atgatgggag tggaatggca ggacagcttt 1560

atggttgcca gactcatagg ttataagacc ttcttcaatg aatttgtggc ttatgagcac 1620
ctctcaaaat ggatccactt gaggaaagaa ggtggaccca aatttgtaaa cggtgtgcag 1680
caatatatat caattcgttc tgagataatc gccacttacg ctctctgtgg ttttgccaat 1740
atcgggtccc taggaatcgt gatcggcgga ctcacatcca tggctccttc cagaaagcgt 1800
gatatcgcct cgggggcagt gagagctctg attgcgggga ccgtggcctg cttcatgaca 1860
gcctgcatcg caggcatact ctccagcact cctgtggaca tcaactgcca tcacgtttta 1920
gagaatgcct tcaactccac tttccctgga aacacaacca aggtgatagc ttgttgccaa 1980
agtctgttga gcagcactgt tgccaagggt cctggtgaag tcatcccagg aggaaaccac 2040
agtctgtatt ctttgaaggg ctgctgcaca ttgttgaatc catcgacctt taactgcaat 2100
gggatctcta atacattttg a 2121

<210> 41
<211> 1225
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 622868CB1

<400> 41
aattcattgg catgaagtct cggacatggg cgtctgtcca tttgcattcc ttttttgcag 60
ttggaaccct gctggtggct ttgacaggat acttggtcag gacctggtgg ctttaccaga 120
tgatcctctc cacagtgact gtccccttta tcctgtgctg ttgggtgctc ccagagacac 180
ctttttggct tctctcagag ggacgatatg aagaagcaca aaaaatagtt gacatcatgg 240
ccaagtggaa cagggcaagc tcctgtaaac tgtcagaact tttatcactg gacctacaag 300
gtcctgttag taatagcccc actgaagttc agaagcacaa cctatcatat ctgttttata 360
actggagcat tacgaaaagg acacttaccg tttggctaat ctggttcact ggaagtttgg 420
gattctactc gttttccttg aatttctgtta acttaggagg caatgaatac ttaaacctct 480
tcctcctggg tgtagtggaa attcccgcct acaccttcgt gtgcatcgcc acggacaagg 540
tcgggaggag aacagtcctg gcctactctc ttttctgcag tgcactggcc tgtggtgtcg 600
ttatggtgat cccccagaaa cattatattt tgggtgtggt gacagctatg gttggaaaat 660
ttgccatcgg ggcagcattt ggcctcattt atctttatac agctgagctg tatccaacca 720
ttgtaagatc gctggctgtg ggaagcggca gcatggtgtg tcgcctggcc agcatcctgg 780
cgccgttctc tgtggacctc agcagcattt ggatcttcat accacagttg tttgttggga 840
ctatggccct cctgagtgga gtgttaacac taaagcttcc agaaaccctt gggaaacggc 900
tagcaactac ttgggaggag gctgcaaaac tggagtcaga gaatgaaagc aagtcaagca 960
aattacttct cacaactaat aatagtgggc tggaaaaaac ggaagcgatt acccccaggg 1020
attctggtct tggtgaataa atgtgccatg cctgctgtct agcacctgaa atattattta 1080
ccctaatgcc tttgtattag aggaatctta ttctcatctc ccatatgttg tttgtatgtc 1140
tttttaataa attttgtaag aaaattttaa agcaaatatg ttataaaaga aataaaaact 1200
aagatgaaaa ttctcagttt taaaa 1225

<210> 42
<211> 2693
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7476494CB1

<400> 42
tcctttgaga cagctgctct gagagaatgc aataagcagg gagcagccag caattcctcc 60
tagcagaggg cgactcgtgg gaggagttca gtttgccaag tattgtcatt tgttgagaga 120
aggtgtgtgc tcaaggagga gttttaacct ggaggatcat taactctttt agtcagctga 180

ggagctgcgg tggctcggcg agttggagtt catcctggaa gcgtctgcac gacaaggtca 240
gggatgaggt gtggaataac tttttcatgg gacactttga gaagggccag cacgctctgc 300
tcaatgaagg agaagagaat gagatggaga tatttggcta tcggactcaa ggctgccgga 360
aaagtctctg ccttgccgga tccatcttct catttggaat cctcccctg gtgttttact 420
ggagaccagc atggcacgta tgggcacatt gtgtcccatg ttccttgcaa gaagcagaca 480
ctgtgttgct gaggacaacg gtgagatgca tcaaagtgca gaaaataaga tatgtttgga 540
actacttaga aggacagttc cagaaaattg gttctttgga agactggctc agttctgcca 600
agatacatca aaaatttgga tcaggcttga caagagaaga acaggagatt aggaggttaa 660
tgtgtgggcc taatactatc gatgttgaag ttacaccaat ttggaaactg ctcatcaagg 720
aggttctaaa tccatttta atatttcaac tcttcagtgt ctgtttgtgg tttagtgaag 780
actataagga atatgctttt gccatcataa tcatgtccat aatttccata tctttgacag 840
tatatgatct cagagagcaa tctgtaaaac tccaccatct cgtcgagtca cataatagca 900
ttacggtctc tgtatgtggg agaaaagctg gagttcaaga gctggaatca cgcgtcctgg 960
tgcctggaga tttattaatt ttgacaggga acaaagtgct aatgccatgt gatgccgttc 1020
tgattgaagg cagctgtgtg gtggatgaag gcatgctgac aggagaaagt attccagtca 1080
ccaaaactcc gttacccaag atggatagct ctgtgccctg gaaaacacag agtgaagcgg 1140
attacaagcg gcatgtcctc ttctgtggaa cagaggttat ccaggccaag gcagcttgct 1200
ctgggaccgt gagagccgtg gtactgcaga ctggattcaa cactgcaaag ggagaccttg 1260
tgagatccat tctctaccct aagccagtga attttcagtt gtacagggat gccatcaggt 1320
tcctcctgtg ccttgtagga acagccacca ttgggatgat ctatactctg tgtgtctatg 1380
tgcttagtgg ggaacctcca gaggaggtgg tgaggaaagc ccttgacgtc atcacaattg 1440
cggttcctcc ggctctacct gctgctctga ccacaggcat tatctatgcc cagaggaggc 1500
tgaagaagag aggcatcttc tgcattagcc cccagaggat caacgtatgt ggacagttaa 1560
accttgtctg ctttgacaag acaggcacct taacaaggga cggcttggac ctctggggag 1620
tcgtgtcctg tgataggaat ggctttcagg aagttcacag ctttgcctca ggccaggctt 1680
tgccatgggg cccactgtgt gcagcgatgg ccagctgcca ctctctgatc cttcttgatg 1740
ggaccatcca gggagaccct ctggacctca aaatgtttga agccaccacc tgggaaatgg 1800
ctttttctgg ggacgatttc cacatcaagg gagtgccggc acatgccatg gtagttaagc 1860
cctgcagaac agccagccag gtcccagtgg aaggaattgc aatcctgcat cagttcccat 1920
tctcatcggc actgcaaaga atgacagtca ttgtccaaga gatgggaggt gaccgactgg 1980
cattcatgaa aggtgcacca gagagggtgg ccagcttttg ccaacctgag acagtaccca 2040
ctagttttgt tagcgaactt cagatttaca cgacacaggg cttccgagtc atagcactgg 2100
cctacaagaa gctggaaaat gaccatcacg ctactacctt gacgagggag acggtagaat 2160
cagacctgat atttctgggg ctgctgatct tggagaatcg attgaaggaa gagacaaaac 2220
ctgtcttgga agagctcatc tcagcccgga taaggactgt aatgatcaca ggtgacaatc 2280
ttcagactgc aataacagtg gccagaaaat ctggaatggt ttctgaaagc cagaaagtca 2340
ttctcattga ggcaaatgaa accaccgggt cctcatcagc atctatatct tggacgttag 2400
tagaagagaa gaaacacatt atgtatggga atcaggacaa ttacattaac atcagggatg 2460
aagtctctga taaaggcaga gaaggaagtt accattttgc cctaactgga aaatcctttc 2520
atgttataag tcaacatttc agcagcctac tgccaaagat attgatcaat gggaccatct 2580
ttgcaagaat gtctcctggg cagaagtcca gtctggtgga agaatttcag aaactggagt 2640
aggttctttg ccagtgcagg tggcatgaac tgcatggagg cataacagtc agg 2693

<210> 43
<211> 3569
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477260CB1

<400> 43
agaccagtgt tggaggatgg cttgcttggg gcccggtggg aagaagaacc ccctcgggtg 60
gtaagctgaa gttgggtcag agtgcttctc acttctctct tcagttctgg ttcttgctac 120
tgccctggct tcgaccattc ccccatcatt cttcactgcc agctgcaaga ccctggatct 180

```
gaatgcagac taaatctttt catctctttt atcttaaaga gtatctcgcc cacctctgat 240
tcatggttac aggaggccag catcacccag gagctggact tagttttaca gaattagaaa 300
atacttttcc cttgtgcttg cctcctactc catttctgtt ggccttgtgg tcctcctgcc 360
ttccatggga cactcagcag acctgctgcc cctcttttgc agggtcccca gctgctgagc 420
agctccagga catcctgggg gaggaagatg aggctcccaa ccccaccctc tttacagaga 480
tggatactct gcagcatgac ggagaccaga tggagtggaa ggagtcagcc aggtggataa 540
agtttgaaga aaaggtagag gaaggcggcg aacgctggag caagccccac gtgtccacac 600
tatccctgca cagcctcttc gagctccgta cctgcctgca gacggggacg gtgctgctgg 660
atttggacag tggctcctta ccacagatca tagatgatgt cattgagaag cagattgagg 720
atggtctcct gcggccagag ctccgggaga gggtcagtta cgtcctcctg aggaggcacc 780
gccaccaaac caagaagccc atccaccgct ccttagctga cattgggaag tcagtctcca 840
ccacaaatcg cagtcctgcc cggagccctg gtgctggccc gagtctacac cactccacgg 900
aagacctgcg gatgcggcag agtgcaaatt acggacgtct gtgtcatgcc cagagcagaa 960
gcatgaatga catttctctc accccaaaca cagaccagcg gaaaaacaaa ttcatgaaga 1020
agatccccaa ggactcagaa gcgtccaacg tgctcgtggg cgaggtggac ttcctagacc 1080
agccattcat cgcgttcgtg cgcctcatcc agtcggccat gctgggagga gtgaccgagg 1140
tgcctgtccc caccagattt ctgtttatac tactgggacc ttctgggaga gcaaaatcct 1200
acaatgaaat tggccgtgcc attgcaaccc tcatggtaga tgatctcttc agtgacgtgg 1260
cctacaaagc ccgcaatcgg gaagatctga tcgcaggaat tgatgaattt ctggatgagg 1320
tcatcgtcct tcctcctgga gaatgggacc caaatatccg gattgagccc cccaagaagg 1380
tgccctctgc tgacaagagg aaatctctgt tctccctagc agagctgggc cagatgaatg 1440
gctctgtggg aggaggcggc ggagctcctg gaggaggcaa tggaggtggt ggtggtggtg 1500
gcagtggcgg cggggctggc agtggcgggg ccggcggaac aagcagcggg gatgatggag 1560
agatgccagc catgcatgaa atcggggagg aacttatctg gacaggaagg ttcttcggtg 1620
gcctgtgtct ggatatcaag aggaagttgc cctggttccc aagtgacttc tatgatggct 1680
tccacattca gtccatctct gccatcctat tcatctacct cggctgtatc accaacgcga 1740
tcacctttgg tgggcttctg ggggatgcca ccgacaatta tcagggagtg atggagagct 1800
tcctgggcac tgccatggct ggctccttgt tctgcctctt ctcgggacag cctctcatca 1860
ttctcagcag cacggggccc atcctcatct ttgagaagct cctcttcgac ttcagcaaag 1920
gcaatggcct ggactacatg gagttccgcc tctggattgg cctacactca gctgttccagt 1980
gccttatcct agtggccaca gatgccagct ttatcatcaa atatatcacc cgcttcaccg 2040
aggagggctt ctccaccctt atcagcttca tcttcatcta cgatgccatc aagaagatga 2100
tcggtgcctt caagtactac cctatcaata tggacttcaa gccaaacttc atcactacct 2160
acaagtgcga gtgtgtcgcc cctgacacag gtgacctgaa tacaaccgtg ttcaatgctt 2220
cagccccatt ggcaccagac accaacgctt ctctgtacaa cctccttaac ctcacagcgt 2280
tggactggtc cctgctgagc aagaaggagt gtctgagcta cggtgggcgc ctgcttggga 2340
attcctgcaa gtttatccca gacctggcgc tcatgtcctt catcctttttc tttgggacat 2400
actccatgac cctgaccctg aagaagttca aattcagccg ctattttcct accaaggtcc 2460
gggccctggt ggctgacttt tccattgttt tctccatcct gatgttctgt ggaatcgatg 2520
cctgttttgg cctagaaact cccaagctgc atgtgcccag tgtcatcaag ccaacgcggc 2580
ctgaccgagg ctggttcgtg gccccctttg ggaagaaccc gtggtgggta tacccagcaa 2640
gcatcctgcc cgccctgctg gtgaccatcc tgatcttcat ggaccagcag atcactgccg 2700
tcattgtcaa ccggaaggag aacaaactga agaaggctgc cggctaccat ctggacctgt 2760
tctgggtggg catcctcatg gctttgtgct cctttatggg gctcccctgg tacgtggctg 2820
ccacggtcat ctccatcgcc cacatcgaca gcctcaagat ggagacagag accagtgccc 2880
ctggggagca gccccagttt ctgggagtca gggaacagag agtaaccggc atcatcgtct 2940
tcatcctgac gggaatctct gtcttcctgg ctcccatcct aaagtgtatc cccctgccgg 3000
tgctgtacgg agtcttcctc tacatgggcg tggcctccct gaatggcatc cagttctggg 3060
aacgctgcaa gctcttcctg atgccagcca agcaccagcc ggaccatgcc ttcctgcggc 3120
acgtgccgct gcgccggatc cacctcttca ccctggtgca gatcctctgc ctggcggtgc 3180
tctggatcct caaatccacg gtggctgcca tcatcttccc ggtcatgatc ctgggcctca 3240
tcatcgttcg aaggcttctg gatttcatct tttcccagca cgacctggcc tggattgaca 3300
acatcctccc agagaaggaa aaaaaggaga cagacaagaa gaggaagaga aaaaaagggg 3360
cccacgagga ctgtgatgag gaggaaaaag atcttccagt tggagttact cactctgatt 3420
cttccttcag tgacacagaa cttgaccgaa gctactcacg gaacccagtg ttcatggtgc 3480
cacaggtgaa gatagagatg gagtcagact atgacttcac agacatggat aaataccgaa 3540
```

gagaaactga cagtgagacc accctctag 3569

<210> 44
<211> 3920
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1963058CB1

<400> 44
cggacgcggc ggacgtgggt gagggcgcgg ccgtaagaga gcgggacgcg gggtgcccgg 60
cgcgtggtgg gggtccccgg cgcctgcccc cacggcaccc aagaaggcct ggccagggta 120
ccctccgcgg agcccggggg tgggggggcgc ggggccggcg ccgcgatggg cccgggaccc 180
ccagcggccg gagcggcgcc gtccccgcgg ccgctgtccc tggtggcgcg gctgagctac 240
gccgtgggcc acttcctcaa cgacctgtgc gcgtccatgt ggttcaccta cctgctgctc 300
tacctgcact cggtgcgcgc ctacagctcc cgcggcgcgg ggctgctgct gctgctgggc 360
caggtggccg acgggctgtg cacaccgctc gtgggctacg aggccgaccg cgccgccagc 420
tgctgcgccc gctacggccc gcgcaaggcc tggcacctgg tcggcaccgt ctgcgtcctg 480
ctgtccttcc ccttcatctt cagcccctgc ctgggctgtg gggcggccac gcccgagtgg 540
gctgccctcc tctactacgg cccgttcatc gtgatcttcc agtttggctg ggcctccaca 600
cagatctccc acctcagcct catcccggag ctcgtcacca acgaccatga gaaggtggag 660
ctcacggcac tcaggtatgc gttcaccgtg gtggccaaca tcaccgtcta cggcgccgcc 720
tggctcctgc tgcacctgca gggctcgtcg cgggtggagc ccacccaaga catcagcatc 780
agcgaccagc tgggggggcca ggacgtgccc gtgttccgga acctgtccct gctggtggtg 840
ggtgtcggcg ccgtgttctc actgctattc cacctgggca cccgggagag gcgccggccg 900
catgcggagg agccaggcga gcacaccccc ctgttggccc ctgccacggc ccagcccctg 960
ctgctctgga agcactggct ccgggagccg gctttctacc aggtgggcat actgtacatg 1020
accaccaggc tcatcgtgaa cctgtcccag acctacatgg ccatgtacct cacctactcg 1080
ctccacctgc ccaagaagtt catcgcgacc attcccctgg tgatgtacct cagcggcttc 1140
ttgtcctcct tcctcatgaa gcccatcaac aagtgcattg ggaggaacat gacctacttc 1200
tcaggcctcc tggtgatcct ggcctttgcc gcctgggtgg cgctggcgga gggactgggt 1260
gtggccgtgt acgcagcggc tgtgctgctg ggtgctggct gtgccaccat cctcgtcacc 1320
tcgctggcca tgacggccga cctcatcggt ccccacacga acagcggagc gttcgtgtac 1380
ggctccatga gcttcttgga taaggtggcc aatgggctgg cagtcatggc catccagagc 1440
ctgcaccctt gcccctcaga gctctgctgc agggcctgcg tgagctttta ccactgggcg 1500
atggtggctg tgacgggcgg cgtgggcgtg gccgctgccc tgtgtctctg tagcctcctg 1560
ctgtggccga cccgcctgcg acgctgggac cgtgatgccc ggccctgact cctgacagcc 1620
tcctgcacct gtgcaaggga actgtgggga cgcacgagga tgccccccag ggccttgggg 1680
aaaagccccc actgcccctc actcttctct ggacccccac cctccatcct cacccagctc 1740
ccgggggtgg ggtcgggtga gggcagcagg gatgcccgcc agggacttgc aaggaccccc 1800
tgggttttga gggtgtccca ttctcaactc taatccatcc cagccctctg gaggatttgg 1860
ggtgcccctc tcggcaggga acaggaagta ggaatcccag aagggtctgg gggaacccta 1920
accctgagct cagtccagtt caccctcac ctccagcctg ggggtctcca gacactgcca 1980
gggccccctc aggacggctg gagcctggag gagacagcca cggggtggtg ggctgggcct 2040
ggaccccacc gtggtgggca gcagggctgc ccggcaggct tggtggactc tgctggcagc 2100
aaataaagag atgacggcag cctggctcct gtctgcctgc gggggggctc tgggcagggg 2160
tagcctgggc atctcagccc tgccctggtt gtgggcggcc agcgagccca gtgtctgcct 2220
ctgtcccgag cctctggtcc cctgggacta ggttagtgcc ccctcatctg ggtgcagaga 2280
cagtgggtgc atcctggtag catgccttta tcggggagtg ggtgtgaggg aaggcgggga 2340
ccgctggcag gtggaggggc agtatggttc caggacccac tcccggtagt tctgggtggt 2400
gccgggcggg cgctggggtg ccgacaggga gggcacgtag tctgatgccc tccacagtgg 2460
ctccaccccg taccggttcc tgttgagcac tgtaggtggg actcgggtca ccatgtgccc 2520
ccacctcctc gccgttggcc agcaaggggc tcctggatcg ccccgggcag tttcaccctg 2580
gcctaggtgg ccttgtcccc ctggcctccc aaggacccac cctgcaccta gcctcaccgt 2640

attccttgcc ccggattggc ctgtctttcc acagcgcgct cccccaccgg gtgctggggg 2700
cctggtactg ggcagggacg atggggtcat gccaggcggt ctcccgcagg tgctgggtgt 2760
aggctgcggt ggggcggggg ctggcggtca ttcctgtccc cctctggcag gcccgctgcc 2820
cagggcgggg gggggggcac tcacccgatg gcatgctgca ctcacggtgt tggtagcagc 2880
tgtgccagcg gttgtaggcc tcgcggaagg ggctggccgg ggcccgtggc aggttgtacc 2940
aggcttccca ggcgtccgag ttggtcaggc ctgtgtacca cagctggccg gctgcgtccc 3000
gtcccatggg cgtgtacttc cagcgtgtgg cctgcctgat ggctggtggc cagcgggaac 3060
cctccaggga caggtagtca tcgctgaggg tgggcagggg gcagagactg agcccatgtc 3120
tacagcgagt gctttgaccc ctttgcgatg tctgccaggg tggatgatgt agaggcctgg 3180
cccacggcgt ggggtctccc tccctcgcca cttggagtct gtccttcagc cctgtacccc 3240
tcaccccaga gtgggtgctt gaggagagag gctgactccc cctctcccca cacatcgcac 3300
cccaagcacc caagtcagca ctaaaccttt ctgttctcag cttttcttgc ctggagaaga 3360
gggaggggag aggacaaggg ccctggctac tcctggattc ctacagtcct tgtccagcct 3420
ccaagaccca caagtccctt cctctgggaa gccccctgg cctggaggtg caccaggaag 3480
aagtggtctg gggctggcac taagccatgg cccagggaag actgggggac ccactaggcc 3540
aggtgtgtgg ctcacgcttg taaacccagc actttgggag gctgaggcag gtggatcact 3600
tgaggtcagg agttcgagac cagcctggcc agcatggtaa aaccccatct ctactaaaaa 3660
tacgaaaatt aagccaggca tggtgtgggg gcgggggggca cctgtaatcc cagctactca 3720
ggaggctgag gcaggagaat cgcttgaacc caggaagtgg agtttgcagt aagctgagat 3780
cgtgcccttt gactccagcc ctgggaaaag agtgagactc cgtcctcaaa aaccaaaggg 3840
ccaggagact caaagaatgt cttatgcttt gaaccttgct ccttggaata atgtcccagg 3900
gaagtcatcc cagaaaacaa 3920

<210> 45
<211> 1361
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2395967CB1

<400> 45
ctggaagcat gtcggagttt tggttaattt ctgcccctgg cgataaggaa aatttgcaag 60
ctctggagag gatgaatact gtaacctcca agtccaacct gtcttataat accaaattcg 120
ctattcctga cttcaaggtg gggaccttgg attccctggt tggcctctct gatgagttgg 180
ggaaactcga caccttgct gaaagcctca taaggagaat ggctcagagc gtggtggaag 240
tcatggagga ctcaaagggg aaggtccagg agcacctcct ggcaaacgga gttgacttaa 300
catcctttgt gacccacttt gaatgggaca tggccaaata tcctgtcaag cagccgctcg 360
tgagtgtggt ggacacaata gccaagcaac tggcgcagat cgagatggac ctgaagtccc 420
gaacggccgc ctacaacact ctgaagacaa acctggagaa cctggaaaag aaatccatgg 480
ggaacctctt cacccggaca ctgagtgata ttgtgagcaa agaggacttc gtgctggatt 540
ctgaatatct cgtcacactt ctggtcatcg tccccaaacc aaactactca caatggcaaa 600
aaacctacga atctctctca gacatggtag tccctcgatc aaccaaactc attactgagg 660
acaaggaagg gggccttttc actgtgactc tgtttcgaaa agtgattgaa gatttcaaaa 720
ccaaggccaa agaaaacaag ttcactgttc gtgaatttta ctatgatgag aaggaaattg 780
aaagggaaag ggaggagatg gccagattgc tgtctgataa gaagcaacag tatggccccc 840
tgctgcgctg gctcaaggtg aacttcagtg aagccttcat tgcctggatc cacatcaagg 900
ccctgagagt gtttgtggag tccgtgctca ggtatggact accagtgaac ttccaggcag 960
tgctcctgca gccgcataag aagtcatcca ccaagcgttt aagagaggtt ctaaactctg 1020
tcttccgaca tctggatgaa gtagccgcta caagtatact ggatgcatct gtggagatcc 1080
cgggactgca actcaataac caagactatt ttccttatgt ctacttccat attgacctta 1140
gtcttcttga ctagaaaggc cagctggcac ctctgtctca tgttcgtgca gattattaca 1200
gacacctctt tcctttagcc agagaatggt tcaaatgtct tacagaacta agatcttttt 1260
cagagaaatt gctcacaaaa gttagtgaca gttgtattta ttttttaag ttacaataaa 1320
atgctctcaa gtcctttgaa tgttccaaca aattcaaaaa a 1361

<210> 46
<211> 1867
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3586648CB1

<400> 46
cagaattagc cggtatagga atgaacgagc atgaagattt gaaattgctc cgattggaag 60
gaagcccagg ttaggtttgg gcacctccaa acgcacccgt tttaaagcca cctggactga 120
ggcgtcgagc tttcagctcc accaaacgct cacctggcct ggcagcgagc ggcggaagag 180
cccgggagcc cctcacagag cgcaccgagc cgggcggaga gctgagccgc aggcacccgc 240
gtctccagga tgataggcga cattgcaaca aatctctaca cccagcagct caggggggctc 300
caagcagagc agcaagttcg aggatccggg cgtggagccg agtgaggccg cagcccagcg 360
ggcctcgggc gaaaaatctt ggaaaatgta taccagtcat gaagatattg ggtatgattt 420
tgaagatggc cccaaagaca aaaagacact gaagccccac ccaaacattg atggcggatg 480
ggcttggatg atggtgctct cctctttctt tgtgcacatc ctcatcatgg gctcccagat 540
ggccctgggt gtcctcaacg tggaatggct ggaagaattc caccagagcc gcggcctgac 600
cgcctgggtc agctccctca gcatgggcat caccttgata gtgggccctt tcatcggctt 660
gttcattaac acctgtgggt gccgccagac tgcgatcatt ggagggctcg tcaactccct 720
gggctgggtg ttgagtgcct atgctgcaaa cgtgcattat ctcttcatta cttttggagt 780
cgcagctggc ctgggcagcg ggatggccta cctgccagcg gtggtcatgg tgggcaggta 840
tttccagaag agacgcgccc tcgcccaggg cctcagcacc acggggaccg gattcggtac 900
gttcctaatg actgtgctgc tgaagtacct gtgcgcagag tacggctgga ggaatgccat 960
gttgatccaa ggtgccgttt ccctaaacct gtgtgtttgt ggggcgctca tgaggcccct 1020
ctctcctggt aaaaacccaa acgacccagg agagaaagat gtgcgtggcc tgccagcgca 1080
ctccacagaa tctgtgaagt caactggaca gcagggaaga acagaagaga aggatggtgg 1140
gctcgggaac gaggagaccc tctgcgacct gcaagcccag gagtgccccg atcaggccgg 1200
gcacaggaag aacatgtgtg ccctccggat tctgaagact gtcagctggc tcaccatgag 1260
agtcaggaag ggcttcgagg actggtattc gggctacttt gggacagcct ctctatttac 1320
aaatcgaatg tttgtagcct ttattttctg ggctttgttt gcatacagca gctttgtcat 1380
ccccttcatt cacctcccag aaatcgtcaa tttgtataac ttatcggagc aaaacgacgt 1440
tttccctctg acgtcaatta tagcaatagt tcacatcttt ggaaaagtga tcctgggcgt 1500
catagccgac ttgccttgca ttagtgtttg gaatgtcttc ctgttggcca acttcaccct 1560
tgtcctcagt atttttattc tgccgttgat gcacacgtac gctggcctgg cggtcatctg 1620
tgcgctgata gggttttcca gtggttattt ctccctaatg cccgtagtga ctgaagactt 1680
ggttggcatt gaacacctgg ccaatgccta cggcatcatc atctgtgcta atggcatctc 1740
tgcattgctg ggaccacctt ttgcaggtaa actctctgag gttttaagag ctcagagtgc 1800
atgtacatat ggtgcgttat gttataaagt cccagattaa gaaacaaaaa aaaaaaaaaa 1860
agatcgg 1867

<210> 47
<211> 2211
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7473396CB1

<400> 47
atgcagaata ttaccaaaga atttggaaca ttcaaggcaa atgacaacat caatttacaa 60
gtaaaggcag gagagattca tgcgttgctt ggagaaaacg gtgctggcaa atctacattg 120
atgaacgtgc tttccggatt attagagccg acatcaggga aaattttgat gcgtgggaaa 180

gaagtacaga tcacaagccc gacaaaagcc aatcaattag ggattgggat ggtccatcag 240
cactttatgc ttgttgatgc ctttactgta acagaaaaca tcgtgttggg aagcgaacct 300
agtcgtgcag ggatgcttga ccataaaaaa gcgcgaaaag agatccaaaa agtttctgaa 360
caatatggat tatcagtcaa cccggatgct tatgttcgtg atatttcagt tgggatggaa 420
caacgggtag aaattttaaa aacactttac cgaggagcag atgtactgat ttttgatgag 480
ccgacagctg tattgacccc tcaggaaatt gatgaattaa tcgtgatcat gaaggaatta 540
gtcaaagaag gcaagtcaat cattttgatt acgcataagt tagatgaaat caaagcagta 600
gctgaccgtt gtacagttat ccgccgtgga aaaggaatcg gtacagtcaa cgttaaagac 660
gttacctcac agcaattagc tgatatgatg gtcggaagag cggtttcatt caaaacgatg 720
aaaaaagaag cgaagcctca agaagtcgtt ttgtctattg aaaatctagt ggtaaaagaa 780
aatcgtggat tagaagccgt gaaaaacctg aacttagagg ttcgtgctgg cgaagtactt 840
ggtatcgctg gaatcgatgg aaacgggcag tcggagttga tccaagcttt gactggtttg 900
cgaaaggcag aaagcggaca tatcaagcta aaaggggaag acatcaccaa taaaaaacct 960
cgaaagatca ctgaacatgg tgtaggacat gtgccagaag accgtcataa atacgggttg 1020
gtcctagata tgacattgtc tgaaaacatt gccctgcaaa cgtatcatca aaaaccttac 1080
agtaaaaacg gtatgctgaa ttattcagtg ataaatgaac atgccagaga attgatcgaa 1140
gaatatgatg ttcgaacaac gaatgaactt gttcctgcaa aagctttatc aggcggaaat 1200
cagcaaaaag caatcatcgc tcggatagtc gaccgagatc ctgatctgtt gatcgttgca 1260
aatccaactc gtgggctgga tgtaggagaa tttgtagcag tcacaggtgt gtctggttct 1320
ggaaagagta cattggtcaa tagtatctta aagaaatcgt tagcgcaaaa attaaataag 1380
aattctgcta agccaggtaa attcaagaca atttccggct acgaaagtat cgaaaagatc 1440
atcgatatcg atcaaagccc aatcggccgg acgccgagaa gtaatccagc gacttataca 1500
agtgtatttg atgatatccg tgggttattt gctcaaacga acgaggcaaa aatgcggggt 1560
tataagaaag ggcgttttag tttcaacgta aaaggcggtc gttgtgaagc ttgtcgcggg 1620
gatggaatta ttaagatcga aatgcacttt ttgcctgatg tctatgttcc ttgtgaagta 1680
tgtcatggca aacgatataa ctctgaaaca ttagaagtgc attacaaagg aaaaagcatt 1740
gctgatattt tggaaatgac agtagaagat gctgtagaat tcttcaagca cattccaaag 1800
attcatcgca aactgcaaac gattgttgat gttggcttag gttatgtgac tatggggcaa 1860
ccagcaacga cattgtccgg tggtgaggca caacggatga aacttgccag tgaattgcac 1920
aaaatctcta atggaaagaa tttctatata ctagatgaac caacgacagg acttcatagc 1980
gatgacatcg cccgcttgtt gcatgtatta caaagattag tagatgctgg taacacagtt 2040
ttagtgattg aacacaatct agatgtaatc aaaacagcag attatatcat tgatttagga 2100
ccagaaggtg gagaaggtgg aggaacgatc cttacgactg gaacaccaga agaaatcatt 2160
aacgtaaaag aaagttatac aggtcactat ttgaaaaaaa taatggtata a 2211

<210> 48
<211> 1446
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7476283CB1

<400> 48
tggctgggag aattgagcta gtgcagcaca cgtaaaaaag cgattccgat gggtcctttg 60
aaagcttttc tcttctcccc ttttcttctg cggagtcaaa gtagaggggt gaggttggtc 120
ttcttgttac tgaccctgca tttgggaaac tgtgttgata aggcagatga tgaagatgat 180
gaggatttaa aggtgaacaa aacctgggtc ttggccccaa aaattcatga aggagatatc 240
acacaaattc tgaattcatt gcttcaaggc tatgacaata aacttcgtcc agatatagga 300
gtgaggccca cagtaattga aactgatgtt tatgtaaaca gcattggacc agttgatcca 360
attaatatgg aatatacaat agatataatt tttgcccaaa cctggtttga cagtcgttta 420
aaattcaata gtaccatgaa agtgcttatg cttaacagta atatggttgg aaaaatttgg 480
attcctgaca ctttcttcag aaactcaaga aaatctgatg ctcactggat aacaactcct 540
aatcgtctgc ttcgaatttg gaatgatgga cgagttctgt atactctaag attgacaatt 600
aatgcagaat gttatcttca gcttcataac tttcccatgg atgaacattc ctgtccactg 660

gaattttcaa gcgatggata ccctaaaaat gaaattgagt ataagtggaa aaagccctcc 720
gtagaagtgg ctgatcctaa atactggaga ttatatcagt ttgcatttgt agggttacgg 780
aactcaactg aaatcactca cacgatctct ggggattatg ttatcatgac aatttttttt 840
gacctgagca gaagaatggg atatttcact attcagacct acattccatg cattctgaca 900
gttgttcttt cttgggtgtc tttttggatc aataaagatg cagtgcctgc aagaacatcg 960
ttgggtatca ctacagttct gactatgaca accctgagta caattgccag gaagtcttta 1020
cctaaggttt cttatgtgac tgcgatggat ctctttgttt ctgtttgttt catttttgtt 1080
tttgcagcct tgatggaata tggaaccttg cattatttta ccagcaacca aaaaggaaag 1140
actgctacta aagacagaaa gctaaaaaat aaagcctcga tgactcctgg tctccatcct 1200
ggatccactc tgattccaat gaataatatt tctgtgccgc aagaagatga ttatgggtat 1260
cagtgtttgg agggcaaaga ttgtgccagc ttcttctgtt gctttgaaga ctgcagaaca 1320
ggatcttgga gggaaggaag gatacacata cgcattgcca aaattgactc ttattctaga 1380
atatttttcc caaccgcttt tgccctgttc aacttggttt attgggttgg ctatctttac 1440
ttataa 1446

<210> 49
<211> 1332
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477105CB1

<400> 49
ttcggctcga gggaccccag gccgggccgg gccgagaggc tgccatgggc tccgtgggga 60
gccagcgcct tgaggagccc agcgtggcag gcacaccaga cccgggcgta gtgatgagct 120
tcaccttcga cagtcaccag ctggaggagg cggcggaggc ggctcagggc cagggccttа 180
gggccagggg cgtcccagct ttcacggata ctacattgga cgagccagtg cccgatgacc 240
gttatcacgc catctacttt gcgatgctgc tggctggcgt gggcttcctg ctgccataca 300
acagcttcat cacggacgtg gactacctgc atcacaagta cccagggacc tccatcgtgt 360
ttgacatgag cctcacctac atcttggtgg cactggcagc tgtcctcctg aacaacgtcc 420
tggtggagag actgaccctg cacaccagga tcaccgcagg ctacctctta gccttgggcc 480
ctctcctttt tatcagcatc tgcgacgtgt ggctgcagct cttctctcgg gaccaggcct 540
acgccatcaa cctggccgct gtgggcaccg tggccttcgg ctgcacagtg cagcaatcca 600
gcttctacgg gcaccgcctg gcccagcctc caccagggac ccctcctcat gaactctgga 660
gccctgagag gagaggggca gccccccacc ttgtcaccct cagggcttcc ccttctgtcc 720
tcattcttag agactgcttc tcccaaacat aacgcgttag ccatgaagga gtcggagccc 780
tgggtccgaa tggacccgcc tgcggtctgc atcagcctct gggaaaccac agcagtgatg 840
ccagctgggc acgtcaggac ctccccacac acccacacga tgccacaggt caggggggctg 900
tgcctgacta gggagccctc ccattgcctt cctggcccgg gatagaagag gggaggtaag 960
tctgggggct acgaagccgg gcccccacac cctggctgaa gtcagcttga cctaggtctt 1020
gaccctcatc cagcaaggga ctcgacagac ccaagggtcc ctggaacgta gggaggggct 1080
gggggtcact ccagcccggg cctcccagaa caccaggccc gtgtgggtgg caccctgagg 1140
tcaggggatc ctaagggtgt ccttccagag acggtgtttc caggggggagg accgcccccg 1200
cttccagatc cccggccccg gctgtgactg ccctgtttca cccctgctgt gtcccatccc 1260
ccgtctgtcc actaactgta ccgcaccggc cattaaaaga tgaaggcaga ccgctggaaa 1320
aaaaaaaaaa aa 1332

<210> 50
<211> 2298
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7482079CB1

<400> 50
atgctcaaac agagtgagag gagacggtcc tggagctaca ggccctggaa cacgacggag 60
aatgagggca gccaacaccg caggagcatt tgctccctgg gtgcccgttc cggctcccag 120
gccagcatcc acggctggac agagggcaac tataactact acatcgagga agacgaagac 180
ggcgaggagg aggaccagtg gaaggacgac ctggcagaag aggaccagca ggcaggggag 240
gtcaccaccg ccaagcccga gggccccagc gaccctccgg ccctgctgtc cacgctgaat 300
gtgaacgtgg gtggccacag ctaccagctg gactactgcg agctggccgg cttccccaag 360
acgcgcctag gtcgcctggc cacctccacc agccgcagcc gccagctaag cctgtgcgac 420
gactacgagg agcagacaga cgaatacttc ttcgaccgcg acccggccgt cttccagctg 480
gtctacaatt tctacctgtc cggggtgctg ctggtgctcg acgggctgtg tccgcgccgc 540
ttcctggagg agctgggcta ctggggcgtg cggctcaagt acacgccacg ctgctgccgc 600
atctgcttcg aggagcggcg cgacgagctg agcgaacggc tcaagatcca gcacgagctg 660
cgcgcgcagg cgcaggtcga ggaggcggag gaactcttcc gcgacatgcg cttctacggc 720
ccgcagcggc gccgcctctg gaacctcatg gagaagccat tctcctcggt ggccgccaag 780
gccatcgggg tggcctccag caccttcgtg ctcgtctccg tggtggcgct ggcgctcaac 840
accgtggagg agatgcagca gcactcgggg cagggcgagg gcggcccaga cctgcggccc 900
atcctggagc acgtggagat gctgtgcatg ggcttcttca cgctcgagta cctgctgcgc 960
ctagcctcca cgcccgacct gaggcgcttc gcgcgcagcg ccctcaacct ggtggacctg 1020
gtggccatcc tgccgctcta ccttcagctg ctgctcgagt gcttcacggg cgagggccac 1080
caacgcggcc agacggtggg cagcgtgggt aaggtgggtc aggtgttgcg cgtcatgcgc 1140
ctcatgcgca tcttccgcat cctcaagctg gcgcgccact ccaccggact gcgtgccttc 1200
ggcttcacgc tgcgccagtg ctaccagcag gtgggctgcc tgctgctctt catcgccatg 1260
ggcatcttca ctttctctgc ggctgtctac tctgtggagc acgatgtgcc cagcaccaac 1320
ttcactacca tcccccactc ctggtggtgg gccgcggtga gtacctttgc cctgggcttt 1380
cccatcctct tccccagccc agtgagctgc tcctccctcc cctggttatc agccaccagg 1440
ctttggcttc tgatcctcgt cttccccccc acccccaatc gccgcataca gctaacaaaa 1500
cggcgatgga tgtcaaaagt ggtggaaaga gaactcagca gatcagtaaa ctccagcagc 1560
cacatgtcga tggctgtggc aaagaacaag agagagaatg caagccccat catgcaaaca 1620
cttcataagt ttcttttcat ggcatttgct cagcccattg gccagagtaa gtcacatggc 1680
caagctgcaa gtcaaagggc agggcaggtg agcatctcca ccgtgggcta cggagacatg 1740
tacccagaga cccacctggg caggtttttt gccttcctct gcattgcttt tgggatcatt 1800
ctcaacggga tgcccatttc catcctctac aacaagtttt ctgattacta cagcaagctg 1860
aaggcttatg agtataccac catacgcagg gagaggggag aggtgaactt catgcagaga 1920
gccagaaaga agatagctga gtgtttgctt ggaagcaacc cacagctcac cccaagacaa 1980
gagaattagt attttatagg acatgtggct ggtagattcc atgaacttca aggcttcatt 2040
gctctttttt taatcattat gattggcagc aaaaggaaat gtgaagcaga catacacaaa 2100
ggccatttcg ttcacaaagt actgcctcta gaaatactca ttttggccca aactcagaat 2160
gtctcatagt tgctctgtgt tgtgtgaaac atctgacctt ctcaatgacg ttgatattga 2220
aaacctgagg ggagcaacag cttagatttt tcttgtagct tctcgtggca tctagctcaa 2280
taaatatttt tggacttg 2298

<210> 51
<211> 2250
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 55145506CB1

<400> 51
agaaacagat ctctcggatc aataagcatg aatgacgaag actacagcac catctatgac 60
acaatccaaa atgagaggac gtatgaggtt ccagaccagc cagaagaaaa tgaaagtccc 120
cattatgatg atgtccatga gtacttaagg ccagaaaatg atttatatgc cactcagctg 180

aatacccatg agtatgattt tgtgtcagtc tataccatta agggtgaaga gaccagcttg 240
gcctctgtcc agtcagaaga cagaggctac ctcctgcctg atgagatata ctctgaactc 300
caggaggctc atccaggtga gccccaggag gacaggggca tctcaatgga agggttatat 360
tcatcagccc aggaccagca actctgcgca gcagaactcc aggagaatgg gagtgtgatg 420
aaggaagatc tgccttctcc ttcaagcttc accattcagc acagtaaggc cttctctacc 480
accaagtatt cctgctattc tgatgctgaa ggtttggaag aaaaggaggg agctcacatg 540
aaccctgaga tttacctctt tgtgaaggct ggaatcgatg gagaaagcat cggcaactgt 600
cctttctctc agcgcctctt catgatcctc tggctgaaag gagtcgtgtt caatgtcacc 660
actgtggatc tgaaaagaaa gccagctgac ctgcacaacc tagcccccgg cacgcacccg 720
cccttcctga ccttcaacgg ggacgtgaag acagacgtca ataagatcga ggagttcctg 780
gaggagacct tgacccctga aaagtacccc aaactggctg caaaacaccg ggaatccaac 840
acagcgggca tcgacatctt ttccaagttt tctgcctaca tcaaaaatac caagcagcag 900
aacaatgctg ctcttgaaag aggcctaacc aaggctctaa agaaattgga tgactacctg 960
aacacccctc taccagagga gattgacgcc aacacttgtg gggaagacaa ggggtcccgg 1020
cgcaagttcc tggatgggga tgagctgacc ctggctgact gcaatctgtt gcccaagctc 1080
catgtggtca agacccacct tctcacttcc tccagcaact tcctaaggaa caagtaccac 1140
tgaaagggat gatataattc cagctcagtc acactgtgtc agagtgatac aatgcaaaga 1200
tcaggagacc cgagttccgg tcctgtattt gctgccaact agcagcatga gctgaggcac 1260
atcatttaat ctttttggaa ttcatttttc tcatgcctag aagaacagaa gtggattgta 1320
ttccttcttg ccttcttttc ctttcttctt tccctccttc tttccttttc tcttgctcaa 1380
acatgtattc actaccactc aaaaaccatt tgttgaacaa agcaaacaaa tgaatctccc 1440
aagccttggg cttcatcctg tgatttcctc aattcccacc tgccttaaat tactcagtga 1500
agccctgtcc ttggagaaaa ttcagtgggt ggttaaccca gagaagctgg agatcaaaaa 1560
gaagatggcc aatgaaagaa caaaggccag cccttggccc ctatctcttt ggatttctgc 1620
tgatccagct tatcagatcc cagaaacctg gcaaacctct aaagttcaca aagagcgaag 1680
gggaagccaa gtcaggcctc cagtttggct tcggatgcca aaacttaatc tgggctgtgg 1740
gagctaactg ttttcatatg aaagagcaaa ttcagaacat gagcatggaa gtccctgcga 1800
acgtcagatc tccgtgtgca tccttacccc cttgctgctt tcatgctcac tctcctcttg 1860
cgtggctcgc tttcaggttt atctccatcc ctggaagcag agttgctctg gcccaggctc 1920
tccatgagag tttggcttga acattcattg tctggccccc tcctagtttct catctcccaa 1980
agtcaagcca atgtgtgaag aaatgaccag ctcagcagcc aaggcccagg gtgcacaggt 2040
cttcgttggg agaggcatct gcaggccttt ccttgcccac tgggatcctt gcctagcata 2100
gtgacgatgt tcagccctgg agacaaacaa gaaggggaac accaacatca atagaagtat 2160
atatttacaa attgcatttc tgctgtattg aaactaacat tctgcccttt aaaatcctga 2220
aaataaaatt tcagtatgaa atgaaaaaaa 2250

<210> 52
<211> 3430
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5950519CB1

<400> 52
gagctgaccc tgcggggtcc cgggggggga gggggagccg cgaagccccc actgaggccg 60
ccgctgccgg gcctcccctc cccccgggc gggcgccatg cgggggagcc cgggcgacgc 120
ggagcggcgg cagcgctggg gtcgcctgtt cgaggagctg gacagtaaca aggatggccg 180
cgtggacgtg cacgagttgc gccaggggct ggccaggctg ggcgggggca acccagaccc 240
cggcgcccaa cagggtatct cctctgaggg tgatgctgac ccagatggcg ggctcgacct 300
ggaggaattt tcccgctatc tgcaggagcg ggaacagcgt ctgctgctca tgtttcacag 360
tcttgaccgg aaccaggatg gtcacattga tgtctctgag atccaacaga gtttccgagc 420
tctgggcatt tccatctcgc tggagcaggc tgagaaaatt ttgcacagca tggaccgaga 480
cggcacaatg accattgact ggcaagaatg gcgcgaccac ttcctgttgc attcgctgga 540
aaatgtggag gacgtgctgt atttctggaa gcattccacg gtcctggaca ttggcgagtg 600

cctgacagtg ccggacgagt tctcaaagca agagaagctg acgggcatgt ggtggaaaca 660
gctggtggcc ggcgcagtgg caggtgccgt gtcacggaca ggcacggccc ctctggaccg 720
cctcaaggtc ttcatgcagg tccatgcctc aaagaccaac cggctgaaca tccttggggg 780
gcttcgaagc atggtccttg agggaggcat ccgctccctg tggcgcggca atggtattaa 840
tgtactcaag attgcccccg agtcagctat caagttcatg gcctatgaac agatcaagag 900
ggccatcctg gggcagcagg agacactgca tgtgcaggag cgcttcgtgg ctggctccct 960
ggctggtgcc acagcccaaa ccatcattta ccctatggag gtgctgaaga cgcggctgac 1020
cttgcgccgg acgggccagt ataaggggct gctggactgc gccaggcgta tcctggagag 1080
ggaggggccc cgtgccttct accgcggcta cctccccaac gtgctgggca tcatccccta 1140
tgcgggcatc gacctggccg tctacgagac tctgaagaac tggtggcttc agcagtacag 1200
ccacgactcg gcagacccag gcatcctcgt gctcctggcc tgcggtacca tatccagcac 1260
ctgcggccag atagccagtt acccgctggc cctggtccgg acccgcatgc aggcacaagc 1320
ctccatcgag ggtggccccc agctgtccat gctgggtctg ctacgtcaca tcctgtccca 1380
ggagggcatg cggggcctct accgggggat cgcccccaac ttcatgaagg ttattccagc 1440
tgtgagcatc tcctatgtgg tctacgagaa catgaagcag gccttggggg tcacgtccag 1500
gtgagggacc cggagcccgt ccccccaatc cctcaccccc cacacctcag ccactggaga 1560
ctgatgatcc aaccacagga tccctactct ttggccacga gatcccagta cccagatcct 1620
ggatcctaga ctcctatgcc ccaaccattg ggtcatggga tcccagcacc cagatcctgg 1680
atcctagact cctatgcccc aaccactggg tcatgcgatc cccacccttc agccactaga 1740
tcccagatcc ccctgtaacc ataactgtgg atcccttact tcagcaactc aagtctgcta 1800
ccctaaccac aagattcaag attatccaca ccccagccct taatccccat cccccaaatc 1860
actggatcct gcagccccac atcctaaggt ggatcccacg cttccctgtg ccccctactg 1920
gatcctggac ctctacgtct taaccactgg atcccacaca aatcagtgaa tggatcccaa 1980
caccccaacc acaggagcac ggattccctg tacctcaaca cccagaccct gcctccctca 2040
ggcaccagat ccagtgtcct agtgaaacgc tggatcctag atccccaacc ccagatcccc 2100
atgcctcgag ccctggatct ccaagctcag ctgctggatt ctggatgtca acaaacctca 2160
ccactggatc ctgacaacca caatgcctgg atcctggggc ccccatcact ggatcccaga 2220
tcccctcact ccacccactg gattcctgca ttggtttttg gtttttttgtt ttttttaac 2280
ctcgacactg ggtctcagat ccttctgctg actgccagat ccctgcattt caagcactac 2340
gccttccacc cccaggcact ggatcccaga ttcccaagcc ttcacccacc agattctggc 2400
tcctaaaaca agtgcggggg ccccagtggc acagcaagtg gatcctggca actgcagctg 2460
ctggattcca gattctgggt ccccaatccc tctgcccagt ccctcaatgt tgaaacctca 2520
tctcttgaag gcagatcctg atattccaag gcactgaatc ccaagccctg aatccccggt 2580
ttctgatctg aatcttccag gcgccgggtc ccaaatgttc aggccccaag tctagatcct 2640
ggcagcccag tcacagagta tcccacacac actggtgccc agagccggct tctcatgaca 2700
tgaaattgca tggtcgaggg agtctgtggg gaaggaagcc caggtcctgg ctgcaacctg 2760
cacggatgct ggattccccc tcaccccacc tctgcatggc caccccctcc cagccctgtg 2820
gggaaactgt tccctggaac cactccactc cctgcatccc cacacttcac agcatcttcc 2880
atcccctcc caccttctag gcgaatagtc cccagagctg tgttcctcca aggggtccga 2940
ggaatcactc actcctggag gctggcaagg agacagtctg aggccaggga cacatgaagg 3000
gatgtcccca ccccagcact atcagggcct ccccaggctt ccagagttga aagccaggag 3060
aaaatcggca aagaccaccc ttccctaaac ccaagcaccc aatgatgcaa aaaacaaaaa 3120
caaaaaaaaa ccaccaaatc cccaaattca ttccagatct atttttctac cagagagagg 3180
agcaaagtcc tcctcccctg cgcccttaca ttctgcactt catagttgga ttctgagctt 3240
aggatcatct ggagacccca tggagggact tggaaagggg aactgggatt tggggagggg 3300
ctggaggact tccgcacgct tccacctcct tcgacctcca ctgcgcccca cctccctgcc 3360
tgtgtgtgtt atttcaaagg aaaagaacaa aaggaataaa ttttctaagc tctttaaaaa 3420
aaaaaaaaaa 3430

【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
21 March 2002 (21.03.2002)

PCT

(10) International Publication Number
WO 02/022684 A3

(51) International Patent Classification⁷: C12N 15/12, 5/10, C12Q 1/68, C07K 14/705, 16/18, 1/22, A61K 38/17, A01K 67/027, G01N 33/68, 33/53, B01J 19/00

(21) International Application Number: PCT/US01/28938

(22) International Filing Date:
14 September 2001 (14.09.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/232,685	15 September 2000 (15.09.2000)	US
60/234,842	22 September 2000 (22.09.2000)	US
60/236,882	29 September 2000 (29.09.2000)	US
60/239,057	5 October 2000 (05.10.2000)	US
60/240,540	13 October 2000 (13.10.2000)	US
60/241,700	18 October 2000 (18.10.2000)	US

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

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LEE, Ernestine, A.** [US/US]; 624 Kains Street, Albany, CA 94706 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **LAL, Preeti, G.** [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **WALIA, Narinder, K.** [US/US]; 890 Davis Street #205, San Leandro, CA 94577 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **WARREN, Bridget, A.** [US/US]; 10130 Parkwood Drive #2, Cupertino, CA 95014 (US). **LEE, Sally** [US/US]; 825 East Evelyn #425, Sunnyvale, CA 94086 (US). **SANJANWALA, Madhu, S.** [US/US]; 210 Sylvia Court, Los Altos, CA 94024 (US). **YAO, Monique, G.** [US/US]; 1189 Woodgate Drive, Carmel, IN 46033 (US). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **THORNTON, Michael** [US/US]; 9 Medway Road, Woodside, CA 94062-2612 (US). **GANDHI, Ameena, R.** [US/US]; 837 Roble Avenue #1, Menlo Park, CA 94025 (US). **POLICKY, Jennifer, L.** [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **ARVIZU, Chandra** [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). **RAUMANN, Brigitte, E.** [US/US]; 5801 South Dorchester Avenue #3B, Chicago, IL 60637 (US). **BRUNS, Christopher, M.** [US/US]; 575 S. Rengstorff Avenue #126, Mountain View, CA 94040 (US). **NAINI, Amir** [US/US]; 51 Spyglass Hill, Oakland, CA 94618 (US). **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **NGUYEN, Danniel, B.** [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). **LU, Dyung, Aina, M.** [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). **ISON, Craig, H.** [US/US]; 1242 Weathersfield Way, San Jose, CA 95118 (US). **GRIFFIN, Jennifer, A.** [US/US]; 33691 Mello Way, Fremont, CA 94555 (US). **REDDY, Roopa, M.** [IN/US]; 1233 W. McKinley Avenue #3, Sunnyvale, CA 94086 (US). **BURFORD, Neil** [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US).

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

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

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

Published:
— *with international search report*

(88) Date of publication of the international search report:
27 March 2003

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

(54) Title: TRANSPORTERS AND ION CHANNELS

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

WO 02/022684 A3

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/28938

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C12N5/10 C12Q1/68 C07K14/705 C07K16/18
C07K1/22 A61K38/17 A01K67/027 G01N33/68 G01N33/53
B01J19/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K C12Q A61K A01K G01N B01J

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EMBL, EPO-Internal, WPI Data, BIOSIS, MEDLINE

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] EMBL; 23 August 1999 (1999-08-23) NCI-CGAP: "wt28g03.x1 NCI CGAP Ut1 Homo sapiens cDNA clone IMAGE:2508820 3' similar to TR:P70423 P70423 AMINO ACID TRANSPORTER, CATIONIC 3 ;, mRNA sequence." Database accession no. AI963203 XP002210463 the whole document --- -/--	1-56,82

[X] Further documents are listed in the continuation of box C. [X] Patent family members are listed in annex.

° Special categories of cited documents :
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
22 August 2002	02 12 2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	van de Kamp, M

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/28938

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HOSOKAWA H ET AL.: "Cloning and characterization of a brain-specific cationic amino acid transporter" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 272, no. 13, 1997, pages 8717-8722, XP002193512 ISSN: 0021-9258 cited in the application the whole document	1-56,82
A	HOSHIDE R ET AL.: "Molecular cloning, tissue distribution, and chromosomal localization of human cationic amino acid transporter 2 (HCAT2)." GENOMICS, vol. 38, no. 2, 1996, pages 174-178, XP002210461 ISSN: 0888-7543 the whole document	1-56,82
A	DEVÉS R ET AL.: "Transporters for cationic amino acids in animal cells: discovery, structure, and function." PHYSIOLOGICAL REVIEWS, vol. 78, no. 2, April 1998 (1998-04), pages 487-545, XP002210462 ISSN: 0031-9333 the whole document page 493, left-hand column, line 28 -page 499, left-hand column, line 47	
A	WO 00 12711 A (INCYTE PHARMA INC ;AZIMZAI YALDA (US); CORLEY NEIL C (US); REDDY R) 9 March 2000 (2000-03-09) the whole document	
A	US 5 989 825 A (KAVANAUGH MICHAEL P ET AL) 23 November 1999 (1999-11-23) the whole document	
P,X	DATABASE EMBL [Online] EMBL; 26 June 2001 (2001-06-26) OTA T ET AL.: "Human protein sequence SEQ ID NO:11182." Database accession no. AAB92739 XP002210464 the whole document	1-5,12, 13,56,82
P,X	-& EP 1 074 617 A (HELIX RES INST) 7 February 2001 (2001-02-07) abstract claims 8-22	1-5,12, 13,56,82

-/--

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/28938

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	DATABASE EMBL [Online] EMBL; 26 June 2001 (2001-06-26) OTA T ET AL.: "Human cDNA sequence SEQ ID NO:11181." Database accession no. AAH14053 XP002210465 the whole document	1-5,12, 13,56,82
P,X	-& EP 1 074 617 A (HELIX RES INST) 7 February 2001 (2001-02-07) abstract claims 8-22	1-5,12, 13,56,82

T	DATABASE EMBL [Online] EMBL; 6 November 2001 (2001-11-06) OTA T ET AL.: "Human polypeptide, SEQ ID NO: 3953." Database accession no. AAM93859 XP002210466 the whole document	1-5,12, 13,56,82
P,X	-& EP 1 130 094 A (HELIX RES INST) 5 September 2001 (2001-09-05) abstract claims 8-22	1-5,12, 13,56,82

T	DATABASE EMBL [Online] EMBL; 6 November 2001 (2001-11-06) OTA T ET AL.: "Human full-length cDNA, SEQ ID NO: 2969." Database accession no. AAK94306 XP002210467 the whole document	1-5,12, 13,56,82
P,X	-& EP 1 130 094 A (HELIX RES INST) 5 September 2001 (2001-09-05) abstract claims 8-22	1-5,12, 13,56,82

T	DATABASE EMBL [Online] embl; 5 December 2001 (2001-12-05) VEKONY N ET AL.: "Homo sapiens cationic amino acid transporter (SLC7A3) mRNA, complete cds." Database accession no. AF320612 XP002210468 the whole document	1-5,12, 13,56,82
T	-& VEKONY N ET AL.: "Human cationic amino acid transporter hCAT-3 is preferentially expressed in peripheral tissues" BIOCHEMISTRY, vol. 40, no. 41, 16 October 2001 (2001-10-16), pages 12387-12394, XP002193513 the whole document	1-5,12, 13,56,82

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/28938

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. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 19, 22, 25, 33, 35, 36 and 39 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [X] Claims Nos.:
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. [X] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-55 (all partially); 56, 82 (completely)

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

International Application No. PCT/US 01/28938

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Present claims 21, 22, 24 and 25 relate to a compound defined by reference to a desirable characteristic or property, namely that it is an agonist or an antagonist of a polypeptide of claim 1.
The claims cover all compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the compounds by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible. Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to compounds which are antibody molecules.

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

International Application No. PCT/US 01/28938

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

1. Claims: 1-55 (all partially); 56, 82 (both completely)

An isolated polypeptide comprising an amino acid sequence according to SEQ ID NO:1, and homologs or variants thereof. An isolated polynucleotide encoding a polypeptide as said, e.g., according to SEQ ID NO:27, and hybridizing or complementary polynucleotides, and fragments thereof. A recombinant polynucleotide comprising a promoter sequence linked to a polynucleotide as said, a cell transformed therewith, and a transgenic organism comprising said recombinant polynucleotide. A method of producing a polypeptide as said. An antibody binding to a polypeptide as said, a method of preparing an antibody as said, a composition comprising an antibody as said, and the use thereof, e.g., in a diagnostic test, in detecting or in purifying a polypeptide as said. A method of detecting a polynucleotide as said. A pharmaceutical composition comprising a polypeptide as said, and the use thereof. A method of screening for a compound for effectiveness as an agonist or as an antagonist of a polypeptide as said, a composition comprising said agonist or antagonist, and the use thereof. A method of screening for a compound that specifically binds to a polypeptide as said, that modulates the activity of a polypeptide as said, or that alters the expression of polynucleotide as said. A method of assessing toxicity of a test compound comprising detecting the amount of the presence of a polynucleotide as said. A microarray or an array comprising a polynucleotide as said or a fragment thereof, and the use thereof.

2. Claims: 1-55 (all partially); 57, 83 (both completely)

As invention 1, but concerning SEQ ID NOs 2 and 28.

3. Claims: 1-55 (all partially); 58, 84 (both completely)

As invention 1, but concerning SEQ ID NOs 3 and 29.

4. Claims: 1-55 (all partially); 59, 85 (both completely)

As invention 1, but concerning SEQ ID NOs 4 and 30.

5. Claims: 1-55 (all partially); 60, 86 (both completely)

As invention 1, but concerning SEQ ID NOs 5 and 31.

6. Claims: 1-55 (all partially); 61, 87 (both completely)

International Application No. PCT/US 01/28938

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

As invention 1, but concerning SEQ ID NOs 6 and 32.

7. Claims: 1-55 (all partially); 62, 88 (both completely)

As invention 1, but concerning SEQ ID NOs 7 and 33.

8. Claims: 1-55 (all partially); 63, 89 (both completely)

As invention 1, but concerning SEQ ID NOs 8 and 34.

9. Claims: 1-55 (all partially); 64, 90 (both completely)

As invention 1, but concerning SEQ ID NOs 9 and 35.

10. Claims: 1-55 (all partially); 65, 91 (both completely)

As invention 1, but concerning SEQ ID NOs 10 and 36.

11. Claims: 1-55 (all partially); 66, 92 (both completely)

As invention 1, but concerning SEQ ID NOs 11 and 37.

12. Claims: 1-55 (all partially); 67, 93 (both completely)

As invention 1, but concerning SEQ ID NOs 12 and 38.

13. Claims: 1-55 (all partially); 68, 94 (both completely)

As invention 1, but concerning SEQ ID NOs 13 and 39.

14. Claims: 1-55 (all partially); 69, 95 (both completely)

As invention 1, but concerning SEQ ID NOs 14 and 40.

15. Claims: 1-55 (all partially); 70, 96 (both completely)

As invention 1, but concerning SEQ ID NOs 15 and 41.

16. Claims: 1-55 (all partially); 71, 97 (both completely)

As invention 1, but concerning SEQ ID NOs 16 and 42.

17. Claims: 1-55 (all partially); 72, 98 (both completely)

International Application No. PCT/US 01/28938

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

As invention 1, but concerning SEQ ID NOs 17 and 43.

18. Claims: 1-55 (all partially); 73, 99 (both completely)
As invention 1, but concerning SEQ ID NOs 18 and 44.

19. Claims: 1-55 (all partially); 74, 100 (both completely)
As invention 1, but concerning SEQ ID NOs 19 and 45.

20. Claims: 1-55 (all partially); 75, 101 (both completely)
As invention 1, but concerning SEQ ID NOs 20 and 46.

21. Claims: 1-55 (all partially); 76, 102 (both completely)
As invention 1, but concerning SEQ ID NOs 21 and 47.

22. Claims: 1-55 (all partially); 77, 103 (both completely)
As invention 1, but concerning SEQ ID NOs 22 and 48.

23. Claims: 1-55 (all partially); 78, 104 (both completely)
As invention 1, but concerning SEQ ID NOs 23 and 49.

24. Claims: 1-55 (all partially); 79, 105 (both completely)
As invention 1, but concerning SEQ ID NOs 24 and 50.

25. Claims: 1-55 (all partially); 80, 106 (both completely)
As invention 1, but concerning SEQ ID NOs 25 and 51.

26. Claims: 1-55 (all partially); 81, 107 (both completely)
As invention 1, but concerning SEQ ID NOs 26 and 52.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 01/28938

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0012711 A	09-03-2000	AU 6137699 A	21-03-2000
		CA 2341148 A1	09-03-2000
		WO 0012711 A2	09-03-2000
		EP 1117781 A2	25-07-2001
US 5989825 A	23-11-1999	US 5882926 A	16-03-1999
		US 5658782 A	19-08-1997
		US 6284505 B1	04-09-2001
		US 6458571 B1	01-10-2002
		US 6074828 A	13-06-2000
		US 5919628 A	06-07-1999
		US 6100085 A	08-08-2000
		US 5932424 A	03-08-1999
		US 6090560 A	18-07-2000
		US 6020479 A	01-02-2000
		US 6060307 A	09-05-2000
		US 6333400 B1	25-12-2001
		US 5919699 A	06-07-1999
		US 5776774 A	07-07-1998
		US 6451993 B1	17-09-2002
		US 5912171 A	15-06-1999
		US 5840516 A	24-11-1998
EP 1074617 A	07-02-2001	AU 6180800 A	19-02-2001
		AU 6180900 A	19-02-2001
		AU 6181000 A	19-02-2001
		AU 6181100 A	19-02-2001
		AU 6181200 A	19-02-2001
		AU 6181300 A	19-02-2001
		AU 6181400 A	19-02-2001
		AU 6181500 A	19-02-2001
		AU 6181600 A	19-02-2001
		AU 6315800 A	19-02-2001
		EP 1074617 A2	07-02-2001
		EP 1205549 A1	15-05-2002
		WO 0109315 A1	08-02-2001
		WO 0109345 A1	08-02-2001
		WO 0109316 A1	08-02-2001
		WO 0109349 A1	08-02-2001
		WO 0109317 A1	08-02-2001
		WO 0109318 A1	08-02-2001
		WO 0109319 A1	08-02-2001
		WO 0109346 A1	08-02-2001
		WO 0109320 A1	08-02-2001
		WO 0109321 A1	08-02-2001
		WO 0109322 A1	08-02-2001
		WO 0109323 A1	08-02-2001
		JP 2002171977 A	18-06-2002
		JP 2002191363 A	09-07-2002
EP 1130094 A	05-09-2001	EP 1130094 A2	05-09-2001
		JP 2002017375 A	22-01-2002

Form PCT/ISA/210 (patent family annex) (July 1998)

フロントページの続き

(51)Int.Cl.[7]	FI		テーマコード（参考）
Ａ６１Ｐ 25/00	Ａ６１Ｐ 35/00		４Ｃ０８４
Ａ６１Ｐ 35/00	Ａ６１Ｐ 37/00		４Ｈ０４５
Ａ６１Ｐ 37/00	Ｃ０７Ｋ 1/22		
Ｃ０７Ｋ 1/22	Ｃ０７Ｋ 14/47		
Ｃ０７Ｋ 14/47	Ｃ０７Ｋ 16/18		
Ｃ０７Ｋ 16/18	Ｃ０７Ｋ 16/46		
Ｃ０７Ｋ 16/46	Ｃ０７Ｋ 19/00		
Ｃ０７Ｋ 19/00	Ｃ１２Ｎ 1/15		
Ｃ１２Ｎ 1/15	Ｃ１２Ｎ 1/19		
Ｃ１２Ｎ 1/19	Ｃ１２Ｎ 1/21		
Ｃ１２Ｎ 1/21	Ｃ１２Ｐ 21/02	Ｃ	
Ｃ１２Ｎ 5/10	Ｃ１２Ｐ 21/08		
Ｃ１２Ｎ 15/02	Ｃ１２Ｑ 1/68	Ａ	
Ｃ１２Ｐ 21/02	Ｇ０１Ｎ 33/15	Ｚ	
Ｃ１２Ｐ 21/08	Ｇ０１Ｎ 33/50	Ｚ	
Ｃ１２Ｑ 1/68	Ｇ０１Ｎ 33/53	Ｄ	
Ｇ０１Ｎ 33/15	Ｇ０１Ｎ 33/53	Ｍ	
Ｇ０１Ｎ 33/50	Ｇ０１Ｎ 33/566		
Ｇ０１Ｎ 33/53	Ｇ０１Ｎ 37/00	１０２	
Ｇ０１Ｎ 33/566	Ｃ１２Ｎ 5/00	Ａ	
Ｇ０１Ｎ 37/00	Ｃ１２Ｎ 15/00	Ｆ	
	Ａ６１Ｋ 37/02		
	Ｃ１２Ｎ 15/00	Ｃ	

(31)優先権主張番号 60/239,057
(32)優先日 平成12年10月5日(2000.10.5)
(33)優先権主張国 米国(US)
(31)優先権主張番号 60/240,540
(32)優先日 平成12年10月13日(2000.10.13)
(33)優先権主張国 米国(US)
(31)優先権主張番号 60/241,700
(32)優先日 平成12年10月18日(2000.10.18)
(33)優先権主張国 米国(US)

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

(72)発明者 ユエ、ヘンリー
アメリカ合衆国カリフォルニア州９４０８７・サニーベイル・ルイスアベニュー ８２６
(72)発明者 ラル、プリーティ・ジー
アメリカ合衆国カリフォルニア州９５０５６・サンタクララ・ピーオーボックス ５１４２
(72)発明者 チョーラ、ナリンダー・ケイ
アメリカ合衆国カリフォルニア州９４５８７・ユニオンシティ・＃７１２・ユニオンスクエア ３３

(72)発明者 ボーグン、マライア・アール
アメリカ合衆国カリフォルニア州９４５７７・サンレアンドロ・サンティアゴロード １４２４４
(72)発明者 ワレン、ブリジット、エイ
アメリカ合衆国カリフォルニア州９５０１４・クーペルティーノ・＃２・パーウッドドライブ １０１３０
(72)発明者 リー、サリー
アメリカ合衆国カリフォルニア州９４０８６・サニーベイル・＃４２５・イーストイーブリン ８２５
(72)発明者 サンジャンワラ、マデュー・エス
アメリカ合衆国カリフォルニア州９４０２４・ロスアルトス・シルビアコート ２１０
(72)発明者 ヤオ、モニーク・ジー
アメリカ合衆国インディアナ州４６０３３・カーメル・ウッドゲートドライブ・ １１８９
(72)発明者 ランクマール、ジャヤラクシミ
アメリカ合衆国カリフォルニア州９４５５５・フレモント・メイバードサークル ３４３５９
(72)発明者 ソーントン、マイケル
アメリカ合衆国カリフォルニア州９４０６２－２６１２・ウッドサイド・メッドウェイロード ９
(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９４０２５・メンロパーク・＃１・ローブルアベニュー ８３７
(72)発明者 ポリッキー、ジェニファー・エル
アメリカ合衆国カリフォルニア州９５１１８・サンノゼ・ジャービスコート １５１１
(72)発明者 エリオット、ビッキー・エス
アメリカ合衆国カリフォルニア州９５１２１・サンノゼ・ポルトンプレイスウェイ ３７７０
(72)発明者 アービズ、チャンドラ
アメリカ合衆国カリフォルニア州９５１２５・サンノゼ・モロッコドライブ １７０６
(72)発明者 ラーマン・ブリジット・イー
アメリカ合衆国イリノイ州６０６３７・シカゴ・＃３ビー・サウスドーチェスターアベニュー ５８０１
(72)発明者 ブラン、クリストファー・エム
アメリカ合衆国カリフォルニア州９４０４０・マウンテンビュー・＃１２６・サウスレングストーフアベニュー ５７５
(72)発明者 ナイニ・アミール
アメリカ合衆国カリフォルニア州９４６１８・オークランド・スパイグラスヒル ５１
(72)発明者 ハファリア、エープリル・ジェイ・エイ
アメリカ合衆国カリフォルニア州９５０５４・サンタクララ・コーレデプリマベーラ ２２２７
(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 スー、ユーミング
アメリカ合衆国カリフォルニア州９４０４０・マウンテンビュー・ウォルナットドライブ １７３９
(72)発明者 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 アイソン、クレイグ・エイチ
アメリカ合衆国カリフォルニア州９５１１８・サンノゼ・ウェザーズフィールドウェイ １２４２
(72)発明者 グリフィン、ジェニファー・エイ
アメリカ合衆国カリフォルニア州９４５５５・フレモント・メローウェイ ３３６９１
(72)発明者 レディ、ルーパ・エム
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 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

Ｆターム(参考) 2G045 AA34 AA35 AA40 BA11 BB50 DA12 DA13 DA14 DA36 FB02
FB03
4B024 AA01 AA11 BA80 CA04 CA11 DA01 DA02 DA05 DA11 EA02
EA04 GA01 GA11 HA03 HA08 HA11
4B063 QA01 QA18 QQ02 QQ08 QQ42 QQ53 QQ61 QR08 QR32 QR42
QR50 QR56 QR72 QR77 QR82 QS24 QS25 QS28 QS34 QS36
QX02
4B064 AG01 AG26 AG27 CA02 CA05 CA10 CA11 CA19 CA20 CC24
CE12 CE13 DA03 DA13
4B065 AA01X AA57X AA87X AA93Y AB01 AB02 BA02 BA08 CA24 CA25
CA44 CA46
4C084 AA02 AA06 AA07 BA01 BA02 BA22 CA53 NA14 ZA022 ZA942
ZB072 ZB262 ZC212
4H045 AA10 AA11 AA20 AA30 BA09 CA40 DA50 DA75 DA76 EA20
EA50 FA72 FA74 GA26

专利名称(译)	转运蛋白和离子通道		
公开(公告)号	JP2004537254A	公开(公告)日	2004-12-16
申请号	JP2002526934	申请日	2001-09-14
[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	リーアーンスティーンエイ ユエヘンリー ラルプリーティジー チョーラナリンダーケイ ボーグンマライアアール ワレンブリジットエイ リーサリー サンジャンワラマデューエス ヤオモニークジー ランクマールジャヤラクシミ ソーントンマイケル ガンディーアミーナアール ポリッキージェニファーエル エリオットビッキーエス アービズチャンドラ ラーマンブリジットイー ブランクリストファーエム ナイニアミール ハファリアエープリルジェイエイ ニュエンダニエルビー スーユーミング リュデュングアイナエム アイソンクレイグエイチ グリフィンジェニファーエイ レディルーパエム バーフォードニール		
发明人	リー、アーンスティーン·エイ ユエ、ヘンリー ラル、プリーティ·ジー チョーラ、ナリンダー·ケイ ボーグン、マライア·アール ワレン、ブリジット、エイ リー、サリー サンジャンワラ、マデュー·エス ヤオ、モニーク·ジー ランクマール、ジャヤラクシミ ソーントン、マイケル ガンディー、アミーナ·アール ポリッキー、ジェニファー·エル エリオット、ビッキー·エス アービズ、チャンドラ ラーマン·ブリジット·イー ブラン、クリストファー·エム ナイニ·アミール		

	ハファリア、エープリル·ジェイ·エイ ニュエン、ダニエル·ビー スー、ユーミング リュ、デュング·アイナ·エム アイソン、クレイグ·エイチ グリフィン、ジェニファー·エイ レディ、ルーパ·エム バーフォード、ニール
IPC分类号	G01N33/50 A61K38/00 A61K45/00 A61P3/00 A61P21/00 A61P25/00 A61P35/00 A61P37/00 C07K1/22 C07K14/47 C07K14/705 C07K16/18 C07K16/46 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/02 C12N15/09 C12N15/12 C12P21/02 C12P21/08 C12Q1/68 G01N33/15 G01N33/53 G01N33/566 G01N37/00
CPC分类号	A61P3/00 A61P21/00 A61P25/00 A61P35/00 A61P37/00 C07K14/705
FI分类号	C12N15/00.ZNA.A A61K45/00 A61P3/00 A61P21/00 A61P25/00 A61P35/00 A61P37/00 C07K1/22 C07K14/47 C07K16/18 C07K16/46 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 G01N37/00.102 C12N5/00.A C12N15/00.F A61K37/02 C12N15/00.C
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/AA40 2G045/BA11 2G045/BB50 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA80 4B024/CA04 4B024/CA11 4B024/DA01 4B024/DA02 4B024/DA05 4B024/DA11 4B024/EA02 4B024/EA04 4B024/GA01 4B024/GA11 4B024/HA03 4B024/HA08 4B024/HA11 4B063/QA01 4B063/QA18 4B063/QQ02 4B063/QQ08 4B063/QQ42 4B063/QQ53 4B063/QQ61 4B063/QR08 4B063/QR32 4B063/QR42 4B063/QR50 4B063/QR56 4B063/QR72 4B063/QR77 4B063/QR82 4B063/QS24 4B063/QS25 4B063/QS28 4B063/QS34 4B063/QS36 4B063/QX02 4B064/AG01 4B064/AG26 4B064/AG27 4B064/CA02 4B064/CA05 4B064/CA10 4B064/CA11 4B064/CA19 4B064/CA20 4B064/CC24 4B064/CE12 4B064/CE13 4B064/DA03 4B064/DA13 4B065/AA01X 4B065/AA57X 4B065/AA87X 4B065/AA93Y 4B065/AB01 4B065/AB02 4B065/BA02 4B065/BA08 4B065/CA24 4B065/CA25 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA06 4C084/AA07 4C084/BA01 4C084/BA02 4C084/BA22 4C084/CA53 4C084/NA14 4C084/ZA022 4C084/ZA942 4C084/ZB072 4C084/ZB262 4C084/ZC212 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA09 4H045/CA40 4H045/DA50 4H045/DA75 4H045/DA76 4H045/EA20 4H045/EA50 4H045/FA72 4H045/FA74 4H045/GA26
优先权	60/232685 2000-09-15 US 60/234842 2000-09-22 US 60/236882 2000-09-29 US 60/239057 2000-10-05 US 60/240540 2000-10-13 US 60/241700 2000-10-18 US
外部链接	Espacenet

摘要(译)

本发明提供了鉴定和编码TRICH的人转运蛋白和离子通道（TRICH）以及多核苷酸。 本发明还提供表达载体，宿主细胞，抗体，激动剂和拮抗剂。 本发明还提供了用于诊断，治疗或预防与TRICH异常表达有关的疾病的方法。

の例を示している（実施例４および実施例５を参照）。

接頭コード	解析タイプおよびまたはプログラム例
GNN, GFG, ENST	たとえばGENSCAN (Stanford University, CA, USA) またはFGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK)を使ったゲノム配列からのエキソン予測。
GBI	ゲノム配列の手動編集解析。
FL	縫合（ステッチ）またはストレッチされたゲノム配列（実施例５を参照)。
INCY	ゲノムへのEST配列マッピングに由来する完全長転写およびエキソン予測。ゲノム位置とEST組成物データを結合して、エキソンと生じる転写を予想する。