

(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2004-500032

(P2004-500032A)

(43) 公表日 平成16年1月8日(2004.1.8)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
C 1 2 N 15/09	C 1 2 N 15/00	Z N A A
A 6 1 K 38/00	A 6 1 K 45/00	2 G O 4 5
A 6 1 K 45/00	A 6 1 P 1/04	4 B O 2 4
A 6 1 P 1/04	A 6 1 P 1/16	4 B O 6 3
A 6 1 P 1/16	A 6 1 P 3/10	4 B O 6 4
	審査請求 未請求 予備審査請求 有	4 B O 6 5
		(全 158 頁) 最終頁に続く

(21) 出願番号	特願2001-511182 (P2001-511182)	(71) 出願人	301005050
(86) (22) 出願日	平成12年7月12日 (2000. 7. 12)		インサイト・ゲノミックス・インコーポレイテッド
(85) 翻訳文提出日	平成14年1月11日 (2002. 1. 11)		アメリカ合衆国カリフォルニア州94304・パロアルト・ポータードライブ 3160
(86) 国際出願番号	PCT/US2000/019036	(74) 代理人	100089266
(87) 国際公開番号	W02001/005969		弁理士 大島 陽一
(87) 国際公開日	平成13年1月25日 (2001. 1. 25)	(72) 発明者	ラル、ブリーティ
(31) 優先権主張番号	60/143, 816		アメリカ合衆国カリフォルニア州95054・サンタクララ・ラスドライブ 2382
(32) 優先日	平成11年7月14日 (1999. 7. 14)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 電子輸送タンパク質

(57) 【要約】

本発明は、ヒト電子輸送タンパク質 (E T R N) と、E T R N を同定及びコードするポリヌクレオチドとを提供する。本発明はまた、発現ベクター及び宿主細胞、抗体、アゴニスト、アンタゴニストを提供する。更に、本発明は、E T R N の発現に関連する疾患の診断・治療・予防方法を提供する。

【特許請求の範囲】

【請求項 1】

単離されたポリペプチドであって、

- (a) SEQ ID NO: 1 乃至 5 を含む一群から選択されたアミノ酸配列と、
- (b) SEQ ID NO: 1 乃至 5 からなる一群から選択されたアミノ酸配列と少なくとも 90% の配列同一性を有する天然のアミノ酸配列と、
- (c) SEQ ID NO: 1 乃至 5 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、
- d) SEQ ID NO: 1 乃至 5 からなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むことを特徴とする単離されたポリペプチド。

10

【請求項 2】

SEQ ID NO: 1 乃至 5 からなる一群から選択された請求項 1 の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 4】

請求項 2 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 5】

SEQ ID NO: 6 乃至 10 からなる一群から選択された請求項 4 の単離されたポリヌクレオチド。

20

【請求項 6】

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

【請求項 7】

請求項 6 の組換えポリヌクレオチドで形質転換された細胞。

【請求項 8】

請求項 6 の組換えポリヌクレオチドを含む遺伝子組換え生物。

【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、

30

- (a) 前記ポリペプチドの発現に好適な条件下で、請求項 1 のポリペプチドをコードするポリヌクレオチドに機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養するステップと、
- (b) そのように発現したポリペプチドを回収するステップとを含むことを特徴とする請求項 1 のポリペプチドの生産方法。

【請求項 10】

請求項 1 のポリペプチドに特異的に結合する単離された抗体。

【請求項 11】

単離されたポリヌクレオチドであって、

- (a) SEQ ID NO: 6 乃至 10 からなる一群から選択されたポリヌクレオチド配列と、
- (b) SEQ ID NO: 6 乃至 10 からなる一群から選択されたポリヌクレオチド配列と少なくとも 70% の配列同一性を有する天然のポリヌクレオチド配列と、
- (c) 前記 (a) に相補的なポリヌクレオチド配列と、
- (d) 前記 (b) に相補的なポリヌクレオチド配列と、
- (e) 前記 (a) 乃至 (d) の RNA 等価物とで構成される一群から選択されたポリヌクレオチド配列を含む単離されたポリヌクレオチド。

40

【請求項 12】

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

50

【請求項 13】

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

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

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

10

【請求項 14】

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

【請求項 15】

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

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

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

20

【請求項 16】

有効量の請求項 1 のポリペプチド及び医薬的に容認できる賦形剤を含む医薬組成物。

【請求項 17】

前記ポリペプチドが、SEQ ID NO: 1 乃至 5 からなる一群から選択されたアミノ酸配列を含むポリペプチドであることを特徴とする請求項 16 の医薬組成物。

【請求項 18】

機能的 ETRN (電子輸送タンパク質である精製されたポリペプチド) の発現の低下に関連する疾患やその症状の治療方法であって、請求項 16 の医薬組成物をそのような治療が必要な患者に投与することを含むことを特徴とする治療方法。

30

【請求項 19】

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

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝露するステップと、

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

【請求項 20】

請求項 19 のスクリーニング方法によって同定されたアゴニスト化合物及び医薬的に容認できる賦形剤を含む医薬組成物。

40

【請求項 21】

機能的 ETRN の発現の低下に関連する疾患やその症状の治療方法であって、請求項 20 の医薬組成物をそのような治療が必要な患者に投与することを含むことを特徴とする治療方法。

【請求項 22】

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

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝露するステップと、

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

50

【請求項 2 3】

請求項 2 2 のスクリーニング方法によって同定されたアンタゴニスト化合物及び医薬的に容認できる賦形剤を含む医薬組成物。

【請求項 2 4】

機能的 E T R N の過剰な発現に関連する疾患やその症状の治療方法であって、請求項 2 3 の医薬組成物をそのような治療が必要な患者に投与することを含むことを特徴とする治療方法。

【請求項 2 5】

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

10

【請求項 2 6】

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

20

【請求項 2 7】

請求項 5 の配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、
(a) 前記標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、
(b) 前記標的ポリヌクレオチドの発現の変化を検出するステップとを含むことを特徴とするスクリーニング方法。

30

【請求項 2 8】

試験化合物の毒性を評価する方法であって、
(a) 核酸を含む生体サンプルを前記試験化合物で処置するステップと、
(b) 処置した前記生体サンプルの核酸を、請求項 1 1 のポリヌクレオチドの連続する少なくとも 2 0 個のヌクレオチドを含むプローブと、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ハイブリダイズさせるステップであって、標的ポリヌクレオチドが請求項 1 1 のポリヌクレオチドのポリヌクレオチド配列またはその断片を含む、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測定するステップと、
(d) 前記処置した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処置の生体サンプルにおけるハイブリダイゼーション複合体の収量とを比較するステップとを含み、
前記処置した生体サンプルにおけるハイブリダイゼーション複合体の収量の差異が試験化合物の毒性を示唆することを特徴とする試験化合物の毒性評価方法。

40

【発明の詳細な説明】**【0 0 0 1】****(技術分野)**

本発明は、電子輸送タンパク質の核酸配列及びアミノ酸配列に関し、また、細胞増殖異常症、生殖障害、および免疫反応の疾患の診断・治療・予防におけるこれらの配列の利用に関する。

50

【0002】

(発明の背景)

チトクロムなどの電子輸送体は、 $NADH$ または $FADH_2$ から電子を受け取り、他の電子輸送体にそれらを渡す。ユビキノンを除くほとんどの電子輸送タンパク質は、内膜タンパク質に結合するフラビン、ヘム、 FeS クラスター、および銅などの補欠分子族である。例えば、アドレノドキシンは、 $NADPH$ (ドレノドキシレダクターゼおよびチトクロム $P450$)と複合体を形成する。チトクロムは、鉄原子と強く結合したものを含むポルフィリン環からなるヘム補欠分子族である。電子輸送反応は、細胞内のエネルギー生産において極めて重要な役割を果たしている。

【0003】

エネルギーは、グルコースおよび脂肪酸の酸化によって生産される。グルコースは最初、細胞質でピルビン酸に変換される。脂肪酸およびピルビン酸はミトコンドリアに輸送されて CO_2 に完全酸化される。その際、酵素によって、 $NADH$ および $FADH_2$ から酸素に電子が輸送され、 ADP および P_i から ATP (酸化的リン酸化)が合成される。

【0004】

ピルビン酸は、ミトコンドリア内に輸送されてアセチル CoA に変換され、クエン酸回路で酸化される。これには、ピルビン酸デヒドロゲナーゼ成分、ジヒドロリポイルトランスアセチラーゼ(*dihydrolipoyl transacetylase*)、およびジヒドロリポイルデヒドロゲナーゼ(*dihydrolipoyl dehydrogenase*)が関与する。クエン酸回路に関与する酵素には、クエン酸合成酵素、アコニターゼ類、イソクエン酸デヒドロゲナーゼ、*transsuccinylases*を含むケトグルタル酸デヒドロゲナーゼ複合体、スクシニル CoA シンセターゼ、コハク酸デヒドロゲナーゼ、フマラーゼ類、リンゴ酸デヒドロゲナーゼがある。アセチル CoA が CO_2 に酸化される時には、 $NADH$ 、 $FADH_2$ 、および GTP の形成が伴う。酸化的リン酸化において、脱水素酵素類による $NADH$ および $FADH_2$ から酸素への電子の輸送の際、ミトコンドリア内膜に存在する F_0F_1 ATP 分解酵素複合体による ADP および P_i からの ATP の合成が伴う。電子の輸送および ATP 合成に関与する酵素複合体には、 F_0F_1 ATP 分解酵素複合体、ユビキノン(CoQ)-チトクロム c 還元酵素、ユビキノ還元酵素、チトクロム b 、チトクロム c_1 、 FeS タンパク質、およびチトクロム c 酸化酵素がある。

【0005】

ATP の合成には、リン酸輸送体および $ATP-DTP$ アンチポートタンパク質を含む膜輸送酵素が必要である。 ATP 結合カセット(ABC)スーパーファミリーも、ミトコンドリア輸送族に属することが示唆された(Hogue, D.L. 他(1999) J. Mol. Biol. 285:379-389)。褐色脂肪脱共役タンパク質が酸化エネルギーを熱として放出し、感染や外傷の発熱反応に関与しているものと思われる(Cannon, B. 他(1998) Ann. NY Acad. Sci. 856:171-187)。

【0006】

新たな複数の電子輸送タンパク質及びそれらをコードするポリヌクレオチドの発見によって、細胞増殖異常症、生殖障害、および免疫反応の疾患の診断・治療・予防において有用な新規の組成物を提供することにより、当分野におけるニーズが満たされる。

【0007】

(発明の要約)

本発明は、総称して「 $ETRN$ 」、個別にはそれぞれ「 $ETRN-1$ 」、「 $ETRN-2$ 」、「 $ETRN-3$ 」、「 $ETRN-4$ 」、および「 $ETRN-5$ 」と呼ぶ電子輸送タンパク質である精製されたポリペプチドを提供する。本発明の一実施態様では、(a) $SEQ ID NO: 1$ 乃至5($SEQ ID NO: 1-5$)からなる一群から選択されたアミノ酸配列と、(b) $SEQ ID NO: 1-5$ からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) $SEQ ID NO: 1$

10

20

30

40

50

- 5 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1 - 5 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含む単離されたポリペプチドを提供する。別法では、SEQ ID NO: 1 - 5 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。

【0008】

更に本発明は、(a) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドをコードする単離されたポリヌクレオチドを提供する。別法では、このポリヌクレオチドは、SEQ ID NO: 1 - 5 からなる一群から選択されたポリペプチドをコードする。別法では、このポリヌクレオチドは、SEQ ID NO: 6 - 10 からなる一群から選択される。

10

【0009】

更に、本発明は、(a) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドをコードするポリヌクレオチドと機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドを提供する。別法では、本発明は、この組換えポリヌクレオチドで形質転換された細胞を提供する。更なる別法では、本発明は、この組換えポリヌクレオチドを含む遺伝子組換え生物を提供する。

20

【0010】

また、本発明は、(a) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1 - 5 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドの生産方法を提供する。この方法は、(a) このポリペプチドの発現に好適な条件下で、このポリペプチドをコードするポリヌクレオチドと機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養するステップと、(b) このように発現したポリペプチドを回収するステップとを含む。

30

【0011】

更に、本発明は、(a) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1 - 5 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドに特異的に結合する単離された抗体を提供する。

40

【0012】

更に、本発明は、(a) SEQ ID NO: 6 - 10 からなる一群から選択されたポリヌクレオチド配列と、(b) SEQ ID NO: 6 - 10 からなる一群から選択されたポリヌクレオチド配列と70%以上の配列同一性を有する天然のポリヌクレオチド配列と、(c) 前記(a)に相補的なポリヌクレオチド配列と、(d) 前記(b)に相補的なポリヌクレオチド配列と、(e) 前記(a)乃至(d)のRNA等化物とで構成される一群から

50

選択されたポリヌクレオチド配列を含む単離されたポリヌクレオチドを提供する。別法では、このポリヌクレオチドは、少なくとも60個の連続するヌクレオチドを含む。

【0013】

更に本発明は、(a) SEQ ID NO: 6-10 からなる一群から選択されたポリヌクレオチド配列と、(b) SEQ ID NO: 6-10 からなる一群から選択されたポリヌクレオチド配列と70%以上の配列同一性を有する天然のポリヌクレオチド配列と、(c) 前記(a)に相補的なポリヌクレオチド配列と、(d) 前記(b)に相補的なポリヌクレオチド配列と、(e) 前記(a)乃至(d)のRNA等化物とで構成される一群から選択されたポリヌクレオチド配列を含むポリヌクレオチド配列を有する標的ポリヌクレオチドをサンプルにおいて検出する方法を提供する。この方法は、(a) 前記サンプル内の標的ポリヌクレオチドと相補的な配列を構成する少なくとも20個の連続するヌクレオチドを含むプローブと前記サンプルをハイブリダイズさせるステップであって、前記プローブと前記標的ポリヌクレオチドまたはその断片とでハイブリダイゼーション複合体が形成される条件下で、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズする、該ステップと、(b) 前記ハイブリダイゼーション複合体の存在するか否かを検出し、存在する場合には随意選択でその収量を測定するステップとを含む。別法では、前記プローブは、少なくとも60個の連続するヌクレオチドを含む。

10

【0014】

更に本発明は、(a) SEQ ID NO: 6-10 からなる一群から選択されたポリヌクレオチド配列と、(b) SEQ ID NO: 6-10 からなる一群から選択されたポリヌクレオチド配列と70%以上の配列同一性を有する天然のポリヌクレオチド配列と、(c) 前記(a)に相補的なポリヌクレオチド配列と、(d) 前記(b)に相補的なポリヌクレオチド配列と、(e) 前記(a)乃至(d)のRNA等化物とで構成される一群から選択されたポリヌクレオチド配列を含むポリヌクレオチド配列を有する標的ポリヌクレオチドをサンプルにおいて検出する方法を提供する。この方法は、(a) ポリメラーゼ連鎖反応増幅を用いて、前記標的ポリヌクレオチドまたはその断片を増幅するステップと、(b) 増幅された前記標的ポリヌクレオチドまたはその断片が存在するか否かを検出し、存在する場合には随意選択でその収量を測定するステップとを含む。

20

【0015】

更に本発明は、(a) SEQ ID NO: 1-5 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-5 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-5 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-5 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含む効果的な量のポリペプチド及び好適な医薬用賦形剤を含む医薬組成物を提供する。一実施例では、SEQ ID NO: 1-5 からなる一群から選択されたアミノ酸配列を含む医薬組成物を提供する。更に、本発明は、患者にこの医薬組成物を投与することを含む、機能的ETRNの発現の低下に関連した疾患やその症状の治療方法を提供する。

30

【0016】

更に本発明は、(a) SEQ ID NO: 1-5 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1-5 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1-5 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1-5 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドのアゴニストとして効果的な化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを含むサンプルを化合物に曝露するステップと、(b) このサンプルのアゴニスト活性を検出するステップとを含む。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む医薬組成物を提供する。更なる別法では、本発明は、この医

40

50

薬組成物の患者への投与を含む、機能的 E T R N 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。

【0017】

更に、本発明は、(a) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と 90% 以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1 - 5 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドのアンタゴニストとして効果的な化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを含むサンプルを化合物に曝露するステップと、(b) このサンプルのアンタゴニスト活性を検出するステップとを含む。別法では、本発明は、この方法によって同定されたアンタゴニスト化合物と好適な医薬用賦形剤とを含む医薬組成物を提供する。更なる別法では、本発明は、この医薬組成物の患者への投与を含む、機能的 E T R N の過剰な発現に関連した疾患やその症状の治療方法を提供する。

10

【0018】

更に本発明は、(a) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と 90% 以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1 - 5 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドに特異的に結合する化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを好適な条件下で少なくとも 1 つの化合物と結合させるステップと、(b) このポリペプチドとこの試験化合物との結合を検出して、このポリペプチドと特異的に結合する化合物を同定するステップとを含む。

20

【0019】

更に本発明は、(a) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と、(b) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列と 90% 以上の配列同一性を有する天然のアミノ酸配列と、(c) SEQ ID NO: 1 - 5 からなる一群から選択されたアミノ酸配列の生物学的に活性な断片と、(d) SEQ ID NO: 1 - 5 とからなる一群から選択されたアミノ酸配列の免疫原性断片とで構成される一群から選択されたアミノ酸配列を含むポリペプチドの活性を調節する化合物をスクリーニングする方法を提供する。このスクリーニング方法は、(a) このポリペプチドを、その活性が許容される条件下で少なくとも 1 つの化合物と結合させるステップと、(b) この試験化合物の存在下でのこのポリペプチドの活性を評価するステップと、(c) この試験化合物の存在下でのこのポリペプチドの活性と、この試験化合物の不在下でのこのポリペプチドの活性とを比較するステップとを含み、この試験化合物の存在下でのこのポリペプチドの活性の変化が、このポリペプチドの活性を調節する化合物の存在を示唆するという特徴を有する。

30

40

【0020】

更に本発明は、SEQ ID NO: 6 - 10 からなる一群から選択された配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、(a) この標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、(b) この標的ポリヌクレオチドの発現の変化を検出するステップとを含む、該スクリーニング方法を提供する。

【0021】

本発明はさらに、試験化合物の毒性を評価する方法を提供する。この方法は、(a) 核酸を含む生体サンプルを前記試験化合物で処置するステップと、(b) 処置した前記生体サンプルの核酸をプローブとハイブリダイズするステップと、(c) ハイブリダイゼーショ

50

ン複合体の収量を測定するステップと、(d)前記処置した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処置の生体サンプルにおけるハイブリダイゼーション複合体の収量とを比較するステップとを含み、前記処置した生体サンプルにおけるハイブリダイゼーション複合体の収量の差異が試験化合物の毒性を示唆する。この方法における前記プローブは、(1)SEQ ID NO: 6乃至10からなる一群から選択されたポリヌクレオチド配列と、(2)SEQ ID NO: 6乃至10からなる一群から選択されたポリヌクレオチド配列と少なくとも70%の配列同一性を有する天然のポリヌクレオチド配列と、(3)前記(1)に相補的なポリヌクレオチド配列と、(4)前記(2)に相補的なポリヌクレオチド配列と、(5)前記(1)乃至(4)のRNA等価物とで構成される一群から選択されたポリヌクレオチド配列を含むポリヌクレオチドの連続する少なくとも20個のヌクレオチドを含む。また、前記ハイブリダイゼーションは、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われる。また、前記標的ポリヌクレオチドが、(1)SEQ ID NO: 6乃至10からなる一群から選択されたポリヌクレオチド配列と、(2)SEQ ID NO: 6乃至10からなる一群から選択されたポリヌクレオチド配列と少なくとも90%の配列同一性を有する天然のポリヌクレオチド配列と、(3)前記(1)に相補的なポリヌクレオチド配列と、(4)前記(2)に相補的なポリヌクレオチド配列と、(5)前記(1)乃至(5)のRNA等価物とを含む。代替的に前記標的ポリヌクレオチドは前記ポリヌクレオチド配列の断片である。

10

【0022】

20

(本発明の記載について)

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

【0023】

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

30

【0024】

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

【0025】

40

(定義)

用語「ETRN」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ETRNのアミノ酸配列を指す。

【0026】

用語「アゴニスト」は、ETRNの生物学的活性を強化したり、模倣する分子を指す。このアゴニストは、ETRNに直接相互作用するか、或いはETRNが関与する生物学的経路の成分と作用して、ETRN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。

【0027】

50

用語「アレル変異配列」は、E T R Nをコードする遺伝子の別の形を指す。アレル変異配列は、核酸配列における少なくとも1つの変異によって生じ、変異mRNA若しくは変異ポリペプチドになり、これらの構造や機能は変わる場合もあれば変わらない場合もある。ある遺伝子は、天然型のアレル変異配列が存在しないもの、1つ或いは多数存在するものがある。一般にアレル変異配列を生じる変異は、ヌクレオチドの自然な欠失、付加、或いは置換による。これらの各変異は、単独或いは他の変異と同時に起こり、所定の配列内で一回或いはそれ以上生じる。

【0028】

E T R N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、E T R Nと同じポリペプチド或いはE T R N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、E T R Nをコードするポリヌクレオチド配列の正常な染色体の遺伝子座ではない位置でのアレル変異配列との不適當或いは予期しないハイブリダイゼーション、並びにE T R Nをコードするポリヌクレオチドの特定のオリゴヌクレオチドプローブを用いて容易に検出可能な或いは検出困難な多形性を含む。コードされたタンパク質も変異され得り、サイレント変化を生じE T R N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E T R Nの活性が保持される範囲で、残基の極性、電荷、溶解度、疎水性、親水性、及び/または両親媒性についての類似性に基づいて成され得る。例えば、負に荷電したアミノ酸にはアスパラギン酸及びグルタミン酸が含まれ、正に荷電したアミノ酸にはリシン及びアルギニンが含まれ得る。類似の親水性の値をもち極性非荷電側鎖を有するアミノ酸には、アスパラギン、グルタミン、セリン、トレオニンが含まれ得る。類似の親水性の値をもち非荷電側鎖を有するアミノ酸には、ロイシン、イソロイシン、バリン、グリシン、アラニン、フェニルアラニン及びチロシンが含まれ得る。

【0029】

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

【0030】

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

【0031】

用語「アンタゴニスト」は、E T R N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、E T R Nに直接相互作用するか、或いはE T R Nが関与する生物学的経路の成分と作用して、E T R N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。

【0032】

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

【0033】

用語「抗原決定基」は、特定の抗体と接触する分子の領域(即ちエピトープ)を指す。タンパク質或いはタンパク質の断片が、宿主動物を免疫化するのに用いられるとき、このタンパク質の種々の領域は、抗原決定基(タンパク質上の特定の領域或いは三次元構造体)

10

20

30

40

50

に特異的に結合する抗体の産生を誘発し得る。抗原決定基は、抗体と結合するために無傷の抗原（即ち、免疫応答を引き出すために用いられる免疫原）と競合し得る。

【0034】

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

10

【0035】

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

20

【0036】

用語「相補的」は、塩基対合によってアニールする2つの一本鎖核酸配列間の関係を指す。例えば、配列「5'AGT3'」が相補的な配列「3'TCA5'」と対をなす。

【0037】

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

30

【0038】

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

40

【0039】

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

元の残基	保存的な置換
Ala	Gly, Ser
Arg	His, Lys

50

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

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

【0040】

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

【0041】

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

【0042】

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

【0043】

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

【0044】

SEQ ID NO: 6-10の断片は、例えば、この断片を得たゲノム内の他の配列とは異なる、SEQ ID NO: 6-10を明確に同定する固有のポリヌクレオチド配列の領

10

20

30

40

50

域を含む。SEQ ID NO: 6 - 10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: 6 - 10を関連ポリヌクレオチド配列から区別する類似の方法に有用である。ある断片と一致するSEQ ID NO: 6 - 10の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に測定できる。

【0045】

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

【0046】

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

【0047】

SEQ ID NO: 1 - 5のある断片は、SEQ ID NO: 6 - 10のある断片によってコードされる。SEQ ID NO: 1 - 5のある断片は、特異的にSEQ ID NO: 1 - 5を同定する固有のアミノ酸配列の領域を含む。例えば、SEQ ID NO: 1 - 5のある断片は、特異的にSEQ ID NO: 1 - 5を認識する抗体の作製用の免疫原性ペプチドとして有用である。ある断片と一致するSEQ ID NO: 1 - 5の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に測定できる。

【0048】

用語「類似性」は相補性の程度を表す。これには、部分的類似性と完全な類似性とがある。用語「同一性」を「類似性」とも言える。同一の配列と標的の核酸とのハイブリダイゼーションが少なくとも部分的に阻止される部分的に相補的な配列は、「実質的に類似」と呼ばれる。完全に相補的な配列と標的の配列とのハイブリダイゼーションの阻止は、緩いストリンジェントな条件の下、ハイブリダイゼーションアッセイ（サザンブロッティング法、溶液ハイブリダイゼーション等）を用いて検査される。実質的に類似の配列またはハイブリダイゼーションプローブは、緩いストリンジェントな条件の下、完全に類似（同一）の配列と標的の配列との結合に対して競合して抑制する。これは、緩いストリンジェントな条件下では非特異的な結合が許容されるということではなく、緩いストリンジェントな条件では、2つの配列の互いへの結合が特異的（即ち、選択的）に相互作用しなければならぬ。部分的な相補性ともいえない（例えば、30%未満の類似性または同一性）第2の標的配列を用いて、非特異的結合が存在しないことの検査が可能である。非特異的結合が存在しない場合は、実質的に類似配列またはプローブが第2の非相補的標的配列とハイブリダイズしない。

【0049】

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

【0050】

ポリヌクレオチド配列間の同一性のパーセントは、MEGALIGN version 3.12e配列アラインメントプログラムに組込まれるCLUSTAL Vアルゴリズムのデフォルトパラメータを用いて決定可能である。このプログラムはLASERGENEソフトウェアパッケージの一部であり、分子生物学分析プログラム一式（DNA STAR, Madison WI）である。このCLUSTAL Vは、Higgins, D. G. 及び P. M. Sharp (1989) CABIOS 5: 151 - 153、Higgins, D. G. 他 (1992) CABIOS 8: 189 - 191に記載されている。ポリヌクレオチド配列の対のアライメントの場合、デフォルトパラメーターは、K t

10

20

30

40

50

uple = 2、gap penalty = 5、window = 4、「diagonals saved」= 4と設定する。「重み付けされた」残基重み付け表が、デフォルトとして選択された。同一性のパーセントは、アラインメントされたポリヌクレオチド配列間の「類似性のパーセント」としてCLUSTAL Vによって報告される。

【0051】

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

【0052】

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

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

【0053】

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

【0054】

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

10

20

30

40

50

【0055】

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

10

【0056】

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

【0057】

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

20

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

【0058】

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

30

【0059】

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

【0060】

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

40

50

【0061】

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

【0062】

本発明のポリヌクレオチド間の高いストリンジェンシーのハイブリダイゼーションでは、約0.2 x SSC及び約1%のSDSの存在の下、約68 で1時間の洗浄過程を含む。別法では、65、60、55、42 の温度で行う。SSCの濃度は、約0.1%のSDSが存在の下、約0.1~2 x SSCの範囲である。通常は、ブロッキング試薬を用いて非特異的なハイブリダイゼーションを阻止する。このようなブロッキング試薬には、例えば、約100~200 µg/mlの切断され変性したサケ精子DNAが含まれる。約35~50% v/vの濃度のホルムアミドなどの有機溶剤が、例えば、RNAとDNAのハイブリダイゼーションなどの特定の場合に用いることができる。これらの洗浄条件の有用な改変は、当業者には周知である。特に高いストリンジェントな条件でのハイブリダイゼーションは、ヌクレオチド間の進化における類似性を示唆し得る。このような類似性は、それらのヌクレオチド及びコードされたポリペプチドが類似の役割を果たしていることを強く示唆する。

【0063】

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

【0064】

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

【0065】

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

【0066】

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

【0067】

用語「エレメント」または「アレイエレメント」は、マイクロアレイ上に固有の指定された位置を有する、ポリヌクレオチド、ポリペプチドまたはその他の化合物を指す。

【0068】

用語「変調」は、ETRNの活性の変化を指す。例えば、変調によって、ETRN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。

【0069】

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

10

20

30

40

50

A)、任意のDNA様物質、及びRNA様物質である。

【0070】

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

【0071】

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

10

【0072】

ETRN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、ETRNの酵素環境に依存し、細胞の種類によって異なり得る。

【0073】

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

20

【0074】

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

30

【0075】

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

40

50

【0076】

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

【0077】

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

【0078】

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

【0079】

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

【0080】

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

剤、基質、補助因子、インヒビター、磁気粒子及びその他の当分野で既知の成分が含まれる。

【0081】

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

【0082】

用語「サンプル」は、その最も広い意味で用いられている。ETRNをコードする核酸若しくはその断片、ETRN自体を含むと推定されるサンプルには、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、溶液中に存在する又は基板に固定されたゲノムDNA、RNA、cDNAと、組織又は組織プリント等も含まれ得る。

10

【0083】

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

20

【0084】

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

【0085】

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

【0086】

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

30

【0087】

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

【0088】

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

40

【0089】

本明細書における「遺伝子組換え生物」とは、当分野で周知の遺伝子組換え技術などを用

50

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

10

【0090】

特定の核酸配列の「変異配列」とは、デフォルトパラメーター設定の「BLAST 2 Sequences」ツール Version 2.0.9 (May - 07 - 1999) を用いる *blastn* によって、ある核酸配列のある長さに対する該特定の核酸配列の同一性が、少なくとも40%と決定された核酸配列のことである。このような核酸の対は、ある長さにおいて、例えば、少なくとも50%または60%、70%、80%、85%、90%、95%、98%、或いはそれ以上の同一性を示し得る。ある変異配列は、例えば、「アレル」変異配列 (上述) または「スプライス」変異配列、「種」変異配列、「多型」変異配列と表すことができる。スプライス変異配列は基準分子と同一性が極めて高い可能性があるが、mRNA プロセッシング中のエキソンの択一的スプライシングによってポリヌクレオチドの数が多くなったり、少なくなったりする。対応するポリペプチドは、基準分子に存在する追加の機能ドメインを有したり、基準分子に存在するドメインが欠落したりし得る。種変異配列は、種によって異なるポリヌクレオチド配列である。得られるポリペプチドは、互いに高いアミノ酸同一性を有する。多型変異配列は、所定の種と種における特定の遺伝子のポリヌクレオチド配列が異なる。多型変異配列はまた、ポリヌクレオチド配列の1つのヌクレオチドが異なる「1ヌクレオチド多型」(SNP) も含み得る。SNPの存在は、例えば、或る集団、病態、病態の性向を示唆し得る。

20

【0091】

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

30

【0092】

(発明)

本発明は、新規のヒト電子輸送タンパク質 (ETRN) 及び ETRN をコードするポリヌクレオチドの発見に基づいた、細胞増殖異常症、生殖障害、および免疫反応の疾患の診断、治療、及び予防におけるそれらの組成物の使用に関する。

40

【0093】

表1は、ETRNをコードする完全長のヌクレオチド配列の構築に用いたインサイト社クローンを示す。列1及び列2はそれぞれ、ポリペプチド配列及びヌクレオチド配列の配列番号 (SEQ ID NO) を示す。列3は、各ETRNをコードする核酸が同定された *Incyte* クローンのクローンIDを示し、列4は、それらのクローンが単離されたcDNAライブラリを示す。列5は、*Incyte* クローン及びそれらに対応するcDNAライブラリを示す。cDNAライブラリが示されていないインサイト社クローンは、プールされたcDNAライブラリに由来する。列5に示されているインサイト社クローンは、各ETRNのコンセンサスヌクレオチド配列の構築に用いられ、ハイブリダイゼーション技術における断片として有用である。

50

【0094】

表2の各列は、本発明の各ポリペプチドの様々な特性を示す。列1は配列番号(SEQ ID NO)、列2は各ポリペプチドにおけるアミノ酸残基の数、列3は潜在的なリン酸化部位、列4は潜在的なグリコシル化部位、列5はシグネチャ(signature)配列及びモチーフを有するアミノ酸残基、列6は、BLAST分析によって同定された相同配列、及び相当する引用を示す。また、引用することを持って本明細書のの一部とする。列7は、分析方法、場合によってはその分析方法が適用できる検索可能なデータベースを示す。列7の分析方法は、配列相同性及びタンパク質モチーフによって各ポリペプチドを特長つけるために用いられた。

【0095】

表3の列は、ETRNをコードするヌクレオチド配列に関連した組織特異性及び疾患、異常症、症状を示している。表3の列1は、ヌクレオチドの配列番号(SEQ ID NO)を示している。列2は、列1のヌクレオチド配列の断片を示している。これらの断片は、例えば、SEQ ID NO: 6 - 10を同定し、SEQ ID NO: 6 - 10と関連するポリヌクレオチド配列とを区別する、ハイブリダイゼーション若しくは増幅の技術において有用である。これらの断片によってコードされるポリペプチドは、例えば、免疫原性ペプチドとして有用である。列3は、ETRNを発現する組織名、及びETRNを発現する全組織におけるその割合を示す。列4は、ETRNを発現する組織に関連する疾患若しくは異常症、症状、並びにETRNを発現する全組織におけるそれらの割合を示す。列5は、各cDNAライブラリのサブクローニングに用いたベクターを示す。

10

20

【0096】

表4の各列は、ETRNをコードするcDNAのクローンが単離されたcDNAライブラリの作製に用いられた組織についての説明である。列1は、ヌクレオチドのSEQ ID NOを示し、列2はそれらのクローンが単離されたcDNAライブラリを示し、列3は列2のcDNAライブラリに対応する組織の由来及び詳細を示す。

【0097】

SEQ ID NO: 6は、第1染色体の143.1 ~ 146.6センチモルガン範囲内にマッピングされる。この範囲には癌に関連するRAS遺伝子も含まれる。

【0098】

SEQ ID NO: 9は、第4染色体の51.1 ~ 60.0センチモルガン及び第9染色体の65.0 ~ 78.4センチモルガンの範囲内にマッピングされる。この第4染色体の51.1 ~ 60.0センチモルガンの範囲には、細胞増殖に関連する遺伝子も含まれる。また、この第9染色体の65.0 ~ 78.4センチモルガンの範囲内には、自己免疫疾患に関連する遺伝子も含まれる。

30

【0099】

本発明はまた、ETRNの変異体も含む。好適なETRNの変異体は、ETRN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ETRN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。

【0100】

本発明はまた、ETRN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、ETRNをコードするSEQ ID NO: 6 - 10からなる一群から選択された配列を含むポリヌクレオチド配列を提供する。配列表に示したSEQ ID NO: 6 - 10のポリヌクレオチド配列は、窒素系塩基のチミンがウラシルに置換され、糖鎖の背骨がデオキシリボースではなくリボースからなる等価RNA配列を含む。

40

【0101】

本発明はまた、ETRN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、ETRN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

50

有する。本発明の特定の実施形態は、SEQ ID NO: 6 - 10 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: 6 - 10 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、ETRN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。

【0102】

遺伝暗号の縮重により作り出され得るETRNをコードする種々のポリヌクレオチド配列には、既知の自然発生する任意の遺伝子のポリヌクレオチド配列と最小の類似性しか有しないものも含まれることを、当業者は理解するであろう。したがって本発明には、可能なコドン選択に基づいた組み合わせの選択によって作り出され得る可能なポリヌクレオチド配列の変異の全てが含まれ得る。これらの組み合わせは、天然のETRN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。

10

【0103】

ETRN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ETRN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含むETRN或いはその誘導体をコードするヌクレオチド配列を作るとは有利となり得る。特定の

20

【0104】

本発明はまた、ETRN及びその誘導体をコードするDNA配列又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、ETRNまたはその任意の断片をコードする配列の中に突

30

【0105】

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: 6 - 10 及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる(例えば、Wahl, G. M. 及びS. L. Berger (1987) *Methods Enzymol.* 152: 399 - 407; and Kimmel, A. R. (1987) *Methods Enzymol.* 152: 507 - 511. を参照)。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。

【0106】

当分野で周知のDNAのシーケンシング方法を用いて、本発明の何れの実施例も実行可能である。この方法には、例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (PE Biosystems, Foster City CA)、熱安定性T7ポリメラーゼ (Amersham, Pharmacia Biotech Piscataway NJ)、或いはELONGASE増幅システム (Life Technologies, Gaithersburg MD) にみられるような校正エキソヌクレアーゼとポリメラーゼとの組み合わせなどの酵素が用いられる。好ましくは、MICROLAB 2200 液体転移システム (Hamilton, Reno, NV)、PTC 200 Thermal Cycler 200 (MJ Research, Watertown MA) 及びA

40

50

BI CATALYST 800 (PE Biosystems) などの装置を用いて配列の準備を自動化する。次に、ABI 373 或いは 377 DNA シークエンシングシステム (PE Biosystems)、MEGABACE 1000 DNA シークエンシングシステム (Molecular Dynamics, Sunnyvale CA) または当分野で周知の他の方法を用いてシークエンシングを行う。得られた配列を当分野で周知の様々なアルゴリズムを用いて分析する (例えば、Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856 - 853. を参照)。

【0107】

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、ETRN をコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば制限部位 PCR 法を利用する 1 つの方法では、一般的なプライマー及び入れ子プライマー (nested primer) を用いてクローニングベクター内のゲノム DNA から未知の配列を増幅する (例えば、Sarkar, G. (1993) PCR Methods Applic 2: 318 - 322 を参照)。逆 PCR 法を用いる別法では、広範な方向に伸長して環状化した鋳型から未知の配列を増幅するプライマーを用いる。この鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限断片に由来する (例えば、Triglia, T. 等 (1988) Nucleic Acids Res 16: 8186 を参照)。キャプチャ PCR 法を用いる第 3 の方法は、ヒト及び酵母菌人工染色体 DNA の既知の配列に隣接する DNA 断片の PCR 増幅を含む (例えば、Lagerstrom, M. 他 (1991) PCR Methods Applic 1: 111 - 119 を参照)。この方法では、多数の制限酵素による消化及びライゲーションを用いて、PCR を行う前に未知の配列の領域の中に組換え二本鎖配列を挿入することが可能である。また、当分野で周知の別の方法を用いて未知の配列を得ることも可能である。(例えば、Parker, J.D. 他 (1991) Nucleic Acids Res. 19: 3055 - 3060 を参照)。更に、PCR、ネスト化プライマー、PROMOTER FINDER ライブラリ (Clontech, Palo Alto CA) を用いれば、ゲノム DNA 内の歩行が可能である。この方法ではライブラリをスクリーニングする必要がなく、イントロン/エキソン接合部を探すのに有用である。全ての PCR 法をベースにした方法では、プライマーは、市販の OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) 或いは別の好適なプログラムなどを用いて、長さが 22 ~ 30 ヌクレオチド、GC 含量が 50% 以上、約 68 ~ 72 の温度で鋳型に対してアニリングするよう設計される。

【0108】

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

【0109】

市販のキャピラリー電気泳動システムを用いて、シークエンシングまたは PCR 産物のヌクレオチド配列のサイズの分析、または確認が可能である。詳しくは、キャピラリーシークエンシングには、電気泳動による分離のための流動性ポリマー、及び 4 つの異なるヌクレオチドに特異的なレーザーで活性化される蛍光色素、放出された波長の検出に利用する CCD カメラを使用することが可能である。出力/光強度は、適切なソフトウェア (例えば、GENOTYPER 及び SEQUENCE NAVIGATOR、PE Biosy

s t e m s) を用いて電気信号に変換され、サンプルのローディングからコンピュータ分析までのプロセス及び電子データ表示がコンピュータ制御可能である。キャピラリー電気泳動法は、特定のサンプルに少量しか存在しない場合もあるDNAの小片のシーケンシングに特に適している。

【0110】

本発明の別の実施例では、ETRN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ETRN、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ETRNのクローン化及び発現に利用可能である。

10

【0111】

種々の目的でETRNをコードする配列を変えるために、当分野で一般的に知られている方法を用いて、本発明のヌクレオチド配列を組換えることができる。この目的には、遺伝子産物のクローン化、プロセッシング及び/または発現の調節が含まれるが、これらに限定されるものではない。ランダムな断片によるDNAの混合や遺伝子断片と合成オリゴヌクレオチドのPCR再組み立てを用いて、ヌクレオチド配列の組換えが可能である。例えば、オリゴヌクレオチド媒介性定方向突然変異誘発を利用して、新しい制限部位を生成する突然変異の導入、グリコシル化パターンの変更、コドン選択の変更、スプライスバリアントの作製等が可能である。

【0112】

本発明のヌクレオチドを、MOLECULAR BREEDING (Maxygen Inc., Santa Clara CA; 米国特許第5,837,458号; Chang, C.-C. 他(1999) Nat. Biotechnol. 17:793-797; Christians, F.C. 他(1999) Nat. Biotechnol. 17:259-264; Cramer, A. 他(1996) Nat. Biotechnol. 14:315-319)などのDNAシャフリング技術を用いてシャフリングして、ETRN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ETRNの生物学的特性を変更或いは改良することができる。DNAシャフリングは、PCR法による遺伝子断片の組換えで遺伝子変異体のライブラリを作製するプロセスである。次に、このライブラリを、目的の特性を有する遺伝子変異体を同定するために選択或いはスクリーニングする。これらの好ましい変異体をプールし、DNAシャフリング及び選択/スクリーニングを繰り返す。従って、人工的な育種及び急速な分子の進化によって多様な遺伝子が作られる。例えば、ランダムな位置に変異がある1つの遺伝子の断片を、目的の特性が最適化するまで、組換え及びスクリーニング、シャフリングを実施することもできる。別法では、所定の遺伝子の断片を、同じ或いは異なった種の同じ遺伝子ファミリーの相同な遺伝子の断片で組換え、それによってプロトコルに従った調節可能な方法で、多数の天然遺伝子の遺伝子多様性を最大にすることができる。

20

30

【0113】

別の実施例によれば、ETRN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. 等(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ETRN自体またはその断片を合成することが可能である。例えば、ペプチド合成は種々の固相技術を用いて実行可能である(例えば、Creighton, T. (1984) Proteins. Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; Roberge, J.Y. 等(1995) Science 269:202-204を参照)。また、合成の自動化は例えばABI 431Aペプチドシンセサイザー(PE Biosystems)を用いて達成し得る。更にETRNのアミノ酸配列または任意のその一部は、直接的な合成の際の変

40

50

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。

【0114】

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

【0115】

生物学的に活性なETRNを発現させるために、ETRN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ETRNをコードするポリヌクレオチド配列におけるエンハンサー、構成型及び発現誘導型のプロモーター、5'及び3'の非翻訳領域などの調節配列が含まれる。このようなエレメントは、その長さ及び特異性が様々である。特定の開始シグナルによって、ETRNをコードする配列のより効果的な翻訳を達成することが可能である。このようなシグナルには、ATG開始コドン及びコザック配列などの近傍の配列が含まれる。ETRNをコードする配列及びその開始コドン、上流の調節配列が好適な発現ベクターに挿入された場合は、更なる転写調節シグナルや翻訳調節シグナルは必要なくなるであろう。しかしながら、コーディング配列或いはその断片のみが挿入された場合は、インフレームのATG開始コドンを含む外来性の翻訳調節シグナルが発現ベクターに含まれなければならない。外来性の翻訳要素及び開始コドンは、自然及び合成の様々なものから得ることが可能である。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である。（例えば、Scharf, D. 他 (1994) *Results Probl. Cell Differ.* 201-18-162.を参照）。

【0116】

当業者に周知の方法を用いて、ETRNをコードする配列、好適な転写及び翻訳調節エレメントを含む発現ベクターを作製することが可能である。これらの方法には、in vitro組換えDNA技術、合成技術、及びin vivo遺伝子組換え技術が含まれる。（例えば、Sambrook, J. 他. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, 4章及び8章, 及び16-17章; 及び Ausubel, F. M. 他. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9章及び13章1-4章を参照）。

【0117】

種々の発現ベクター/宿主系を利用して、ETRNをコードする配列の保持及び発現が可能である。これらには、限定するものではないが、組換えバクテリオファージ、プラスミド、またはコスミドDNA発現ベクターで形質転換された細菌などの微生物や、酵母菌発現ベクターで形質転換された酵母菌や、ウイルス発現ベクター（例えば、バキュロウイルス）に感染した昆虫細胞系や、ウイルス発現ベクター（例えば、カリフラワーモザイクウイルス、CaMV; タバコモザイクウイルス、TMV）または細菌発現ベクター（例えば、TiまたはpBR322プラスミド）で形質転換された植物細胞系や、動物細胞系などが含まれる（前出のSambrook、前出のAusubel、Van Heeke, G. and S. M. Schuster (1989) *J. Biol. Chem.* 264: 5503-5509、Bitter, G. A. 他 (1987) *Methods Enzymol.* 153: 516-544; Scorer, C. A. ら (1994) *Bio/Technology* 12: 181-184; Engelhard, E. K

10

20

30

40

50

他 (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227、Sandig, V. 他 (1996) Hum. Gene Ther. 7:1937-1945、タカマツ, N. (1987) EMBOJ. 6:307-311、Coruzzi, G. 他 (1984) EMBOJ. 3:1671-1680、Broglie, R. 他 (1984) Science 224:838-843、Winter, J. 他 (1991) Results Probl. Cell Differ. 17:85-105、『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, pp.191-196、Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659、Harrington, J. J. 他 (1997) Nat. Genet. 15:345-355等を参照)。レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ輸送することができる (Di Nicola, M. 他 (1998) Cancer Gen. Ther. 5(6):350-356、Yu, M. 他 (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344、Buller, R. M. 他 (1985) Nature 317(6040):813-815; McGregor, D. P. 他 (1994) Mol. Immunol. 31(3):219-226、Verma, I. M. 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。

【0118】

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

【0119】

ETRNの発現に酵母の発現系の使用が可能である。因子やアルコールオキシダーゼやPGHプロモーターなどの構成型或いは誘導型のプロモーターを含む多種のベクターが、酵母菌サッカロミセス-セレビジエまたは *Pichia pastoris* に使用可能である。更に、このようなベクターは、発現したタンパク質の分泌か細胞内への保持のどちらかを誘導し、安定した増殖のために宿主ゲノムの中に外来配列を組み込む。(例えば、上記の Ausubel.; 及び Bitter, G. A. 他 (1987) Methods Enzymol. 153:51-794; Scorer, C. A. 他 (1994) Bio/Technology 12:1-181-184. を参照)

植物系もETRNの発現に使用可能である。ETRNをコードする配列の転写は、例えば、CaMV由来の35S及び19Sプロモーターなどのウイルスプロモーターが単独で、或いはTMV (例えば、Coruzzi, 前出、Broglie, 前出、Winter

、前出を参照)由来のオメガリーダー配列と組み合わせて促進される。これらの作製物は、直接のDNA形質転換或いは病原体を介したトランスフェクションによって、植物細胞の中に導入可能である。(例えば、The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill NY, pp. 191 - 196を参照)。

【0120】

哺乳動物細胞では、多種のウイルスベースの発現系が利用され得る。アデノウイルスが発現ベクターとして用いられる場合、後発プロモーター及び3連リーダー配列からなるアデノウイルス転写物/翻訳複合体にETRNをコードする配列を結合し得る。ウイルスのゲノムの非必須のE1またはE3領域への挿入により、感染した宿主細胞にETRNを発現する生ウイルスを得ることが可能である(Logan, J. 及びShenk, T. (1984) Proc. Natl. Acad. Sci. 81: 3655 - 3659を参照)。さらに、ラウス肉腫ウイルス(RSV)エンハンサーなどの転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させることが可能である。タンパク質を高レベルで発現させるために、SV40またはEBVを基にしたベクターを用いることが可能である。

10

【0121】

ヒト人工染色体(HAC)を用いて、プラスミドで発現しそれに含まれているものより大きなDNAの断片を供給可能である。治療のために約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)。

20

【0122】

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ETRN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、ETRNをコードする配列を株化細胞に形質転換することが可能である。このような発現ベクターは、ウイルス起源の複製及び/または内在性の発現要素や、同じ或いは別のベクターの上の選択マーカー遺伝子を含む。ベクターの導入の後、細胞を選択培地に移す前に、強化培地で約1~2日の間増殖させる。選択マーカーの目的は選択的な媒介物に対する抵抗性を与えるとともに、その存在により導入された配列を確実に発現する細胞の増殖及び回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に好適な組織培養技術を用いて増殖可能である。

30

【0123】

任意の数の選択系を用いて、形質転換された細胞系を回収することが可能である。選択系には、以下のものに限定はしないが、単純ヘルペスウイルスチミジンキナーゼ遺伝子及びアデニンホスホリボシルトランスフェラーゼ遺伝子が含まれ、それぞれtk^r又はapr^r細胞において使用される。(例えば、Wigler, M. 他(1977) Cell 11: 223 - 232; 及びLowy, I. 他(1980) Cell 22: 817 - 823を参照)。また代謝拮抗物質、抗生物質或いは除草剤への耐性を選択のベースとして用いることができる。例えばdhfrはメトトレキセートに対する耐性を与え、neoはアミノグリコシッドネオマイシン及びG-418に対する耐性を与え、als或いはpatはクロルスルフロン(cETRNsulfuron)、ホスフィンオトリシンアセチルトランスフェラーゼ(phosphinotricin acetyltransferase)に対する耐性を与える(例えば、Wigler, M. 他(1980) Proc. Natl. Acad. Sci. 77: 3567 - 3570; Colberre-Garapin, F. 他(1981) J. Mol. Biol. 150: 1 - 14を参照)。さらに選択に利用できる遺伝子、例えば、代謝のために細胞が必要なものを変えるtrpB及びhisDが文献に記載されている(例えば、Hartman, S. C. 及びR. C. Mulligan(1988) Proc. Natl. Acad. Sci. 85: 8047 - 51を参照)。アニトシアニン、緑色蛍光タンパク質(GFP

40

50

; Clontech)、グルクロニダーゼ及びその基質GUS、ルシフェラーゼ及びその基質ルシフェリンなどの可視マーカーが用いられる。緑色蛍光タンパク質(GFP)(Clontech, Palo Alto, CA)も使用できる。これらのマーカーを用いて、トランスフォーマントを特定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することが可能である(例えば、Rhodes, C. A. 他(1995) *Methods Mol. Biol.* 55:121-131を参照)。

【0124】

マーカー遺伝子の発現の存在/不在によって目的の遺伝子の存在が示されても、その遺伝子の存在及び発現の確認が必要な場合もある。例えば、ETRN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、ETRNをコードする配列を含む形質転換された細胞は、マーカー遺伝子機能の欠落により特定可能である。または、1つのプロモーターの制御下でマーカー遺伝子がETRN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。

10

【0125】

一般に、ETRNをコードする核酸配列を含み、ETRN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。

20

【0126】

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ETRNの発現の検出及び計測のための免疫学的方法は、当分野で周知である。このような技法には、酵素に結合したイムノソルベントアッセイ(ELISA)、ラジオイムノアッセイ(RIA)、蛍光標示式細胞分取器(FACS)などがある。ETRN上の2つの非干渉エピトープに反応するモノクローナル抗体を用いた、2部位のモノクローナルベースイムノアッセイ(two-site, monoclonal-based immunoassay)が好ましいが、競合の結合アッセイも用いることもできる。これらのアッセイ及びその他のアッセイは、当分野では十分に知られている。(例えば、Hampton, R. 他.(1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Sect. IV; Coligan, J. E. 他 Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York, NY; 及びPound, J. D. (1990) Immunochemical Protocols, Humana Press, Totowa NJ)。

30

【0127】

種々の標識技術及び結合技術が当業者には周知であり、様々な核酸アッセイおよびアミノ酸アッセイに用いられ得る。ETRN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、ETRNをコードする配列、またはその任意の断片をmRNAプローブを生成するためのベクターにクローニングすることも可能である。当分野では周知であり市販されているこのようなベクターを、T7, T3, またはSP6などの好適なRNAポリメラーゼ及び標識されたヌクレオチドの追加によって、in vitroでのRNAプローブの合成に用いることができる。これらの方法は、例えば、Amersham Pharmacia Biotech及びPromega(Madison WI)、U.S. Biochemical Corp(Cle

40

50

v e l a n d O H) が市販する種々のキットを用いて行うことができる。容易な検出のために用い得る好適なレポーター分子或いは標識には、基質、コファクター、インヒビター、磁気粒子、及び放射性核種、酵素、蛍光剤、化学発光剤、色素産生剤などが含まれる。

【 0 1 2 8 】

E T R N をコードするヌクレオチド配列で形質転換された宿主細胞は、細胞培地でのこのタンパク質の発現及び回収に好適な条件下で培養される。形質転換された細胞から産生されたタンパク質が分泌されるか細胞内に留まるかは、使用されるその配列及び/またはそのベクターによる。E T R N 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E T R N 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。

10

【 0 1 2 9 】

更に、挿入した配列の発現調節能力または発現したタンパク質を所望の形にプロセッシングする能力によって宿主細胞株が選択される。このようなポリペプチドの修飾には、アセチル化、カルボキシル化、グリコシル化、リン酸化、脂質化 (l i p i d a t i o n)、及びアシル化が含まれるが、これらに限定されるものではない。タンパク質の「プロ」または「プロ」形を切断する翻訳後のプロセッシングを利用して、標的タンパク質、折りたたみ及び/または活性を特定することが可能である。翻訳後の活性のための特定の細胞装置及び特徴のある機構をもつ種類の宿主細胞 (例えば、C H O、H e L a、M D C K、M E K 2 9 3、W I 3 8) が A m e r i c a n T y p e C u l t u r e C o l l e c t i o n (A T C C ; B e t h e s d a , M D) より入手可能であり、外来のタンパク質の正しい修飾及びプロセッシングを確実にするために選択される。

20

【 0 1 3 0 】

本発明の別の実施例では、E T R N 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E T R N タンパク質が、E T R N の活性のインヒビターに対するペプチドライブラリのスクリーニングを促進し得る。また、異種タンパク質部分及び異種ペプチド部分が、市販の親和性基質を用いて融合タンパク質の精製を促進し得る。このような部分には、グルタチオンSトランスフェラーゼ (G S T)、マルトース結合タンパク質 (M B P)、チオレドキシン (T r x)、カルモジュリン結合ペプチド (C B P)、6 - H i s、F L A G、c - m c、赤血球凝集素 (H A) が含まれるが、これらに限定されるものではない。G S T 及びM B P、T r x、C B P、6 - H i s によって、固定されたグルタチオン、マルトース、フェニルアルシン酸化物 (p h e n y l a r s i n e o x i d e)、カルモジュリン、金属キレート樹脂のそれぞれで同族の融合タンパク質の精製が可能となる。F L A G、c - m c、及び赤血球凝集素 (H A) によって、これらのエピトープ標識を特異的に認識する市販のモノクローナル抗体及びポリクローナル抗体を用いた融合タンパク質の免疫親和性の精製ができる。また、E T R N 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、E T R N が精製の後に異種部分から切断され得る。融合タンパク質の発現と精製の方法は、A u s u b e l . (1 9 9 5、前出 c h 1 0) . に記載されている。市販されている様々なキットを用いて、融合タンパク質の発現及び精製を促進できる。

30

40

【 0 1 3 1 】

本発明の別の実施例では、T N T ウサギ網状赤血球可溶化液またはコムギ胚芽抽出系 (P r o m e g a) を用いて *i n v i t r o* で放射能標識したE T R N の合成が可能である。これらの系は、T 7 またはT 3、S P 6 プロモーターと機能的に結合したタンパク質をコードする配列の転写と翻訳をつなげる。翻訳は、例えば、³ ⁵ S メチオニンである放射能標識されたアミノ酸前駆体の存在の下で起こる。

【 0 1 3 2 】

本発明のE T R N またはその断片を用いて、E T R N に特異結合する化合物をスクリーニ

50

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、ETRN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。

【0133】

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ETRN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、ETRN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ETRNを発現する好適な細胞の作製が含まれる。好適な細胞には、哺乳動物、酵母、大腸菌からの細胞が含まれる。ETRNを発現する細胞またはETRNを含有する細胞膜断片を試験化合物と接触させて、ETRN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。

10

【0134】

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

20

【0135】

本発明のETRNまたはその断片を用いて、ETRN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、ETRNが少なくとも1つの試験化合物と結合する、ETRNの活性が許容される条件下でアッセイを実施し、試験化合物の存在下でのETRNの活性が試験化合物不在下でのETRNの活性と比較する。試験化合物の存在下でのETRNの活性の変化は、ETRN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ETRNの活性に適した条件下でETRN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、ETRN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。

30

【0136】

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、ETRNまたはその哺乳動物相同体をコードするポリヌクレオチドを「ロックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの作製に有用である（米国特許第5,175,383号及び第5,767,337号等を参照）。例えば129/SvJ細胞株等のマウスES細胞は初期のマウス胚に由来し、培地で増殖させることができる。このES細胞は、ネオマイシンホスホトランスフェラーゼ遺伝子（neo: Capecchi, M. R. (1989) Science 244:1288-1292）等のマーカー遺伝子で破壊した目的の遺伝子を含むベクターで形質転換する。このベクターは、相同組換えにより宿主ゲノムの対応する領域に組み込まれる。別法では、Cre-loxP系を用いて相同組換えを行い、組織特異的または発生段階特異的に目的遺伝子をロックアウトする（Marth, J. D. (1996) Clin. Invest. 97:1999-2002; Wagner, K. U. 他（1997）N

40

50

ucleic Acids Res. 25:4323-4330)。形質転換したES細胞を同定し、例えばC57BL/6マウス系等から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を決め、これを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、潜在的な治療薬や毒性薬剤で検査することができる。

【0137】

ETRN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)。

【0138】

ETRN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、ETRNをコードするポリヌクレオチドの或る領域を動物ES細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について研究し、潜在的な医薬品を用いて処置し、ヒトの疾患の治療に関する情報を得る。別法では、例えばETRNを乳汁内に分泌するなどETRNを過剰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)。

【0139】

(治療)

ETRNのある領域と電子輸送タンパク質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、ETRNの発現は、生殖組織や細胞増殖組織に密接に関連する。従って、ETRNは、細胞増殖異常症、生殖障害、および免疫反応の疾患においてある役割を果たすと考えられる。ETRN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、ETRNの発現または活性を低下させることが望ましい。また、ETRN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、ETRNの発現または活性を増大させることが望ましい。

【0140】

従って、一実施例において、ETRN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ETRNまたはその断片や誘導体を投与することが可能である。限定するものではないが、このような疾患の例には細胞増殖異常が含まれ、その中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病(MCTD)、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺含まれ、また生殖障害が含まれ、その中には、プロラクチン産生異常と、卵管病、排卵異常、及び子宮内膜症、発情期異常、月経周期異常、多嚢胞卵巣症候群、卵巣過剰刺激症候群、子宮内膜癌及び卵巣癌、子宮筋腫、自己免疫異常、異所性妊娠、及び奇形発生を含む不妊症と、乳癌、線維嚢胞性乳腺症、及び乳漏症と、精子形成異常、生理学上の精子異常、精巣癌、前立腺癌、良性の前立腺過形成、前立腺炎、ペーロニー病、インポテンス、男性乳房癌及び女性化乳房とが含まれ、また免疫反応の疾患も含まれ、その中には後天性免疫不全症候群(AIDS)及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィ(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、心筋又は心膜炎症、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、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、体外循環と、ウイルス感染症及び細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫の感染症が含まれる。

【0141】

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 E T R N 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、E T R N 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。

10

【0142】

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 E T R N 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 E T R N を含む医薬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。

【0143】

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 E T R N 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、E T R N 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。

【0144】

更なる実施例では、E T R N 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 E T R N 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、上記した細胞増殖異常症、生殖障害、および免疫反応の疾患が含まれる。一実施態様では、E T R N 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 E T R N を発現する細胞または組織に薬剤を運ぶターゲティング或いは運搬機構として間接的に用いられ得る。

20

【0145】

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 E T R N 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、E T R N 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。

【0146】

別の実施例では、本発明の任意のタンパク質、アンタゴニスト、抗体、アゴニスト、相補的な配列、ベクターを別の好適な治療薬と組み合わせることもできる。当業者は、従来の医薬原理にしたがって併用療法で用いる好適な治療薬を選択可能である。治療薬との組み合わせにより、上に列記した種々の疾患の治療または予防に相乗効果をもたらし得る。この方法を用いて少ない量の各薬剤で医薬効果をあげることが可能であり、広範囲な副作用の可能性を低減し得る。

30

【0147】

E T R N 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 E T R N 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 E T R N と特異的に結合するものを同定が可能である。E T R N 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、F a b フラグメント、及び F a b 発現ライブラリによって作られたフラグメントが含まれる。但し、これらに限定されるものではない。治療用には、中和抗体（即ち、二量体の形成を阻害するもの）が特に好ましい。

40

【0148】

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、E T R N または任意の断片、または免疫原性の特性を備えるそのオリゴペプチドの注入によって免疫化され得る。宿主の種に応じて、種々のアジュバントを用いて免疫応答を高めることもできる。このようなアジュバントにはフロイントアジュバント、水

50

酸化アルミニウムなどのミネラルゲルアジュバント、リゾレシチン、プルロニックポリオール、ポリアニオン、ペプチド、油性乳剤、キーホールリンペットヘモシニアン、及びジニトロフェノールなどの界面活性剤が含まれるが、これらに限定されるものではない。ヒトに用いられるアジュバントの中では、BCG (*bacilli Calmette - Guerin*) 及び *Corynebacterium parvum* が特に好ましい。

【0149】

ETRNに対する抗体を誘発するために用いられるオリゴペプチド、ペプチド、または断片は、少なくとも約5個のアミノ酸からなり、一般的には約10個以上のアミノ酸からなるものが好ましい。これらのオリゴペプチド或いはペプチド、またはそれらの断片は、天然のタンパク質のアミノ酸配列の一部と同一であることが望ましい。ETRN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。

10

【0150】

ETRNに対するモノクローナル抗体は、培地内の連続した細胞株によって、抗体分子を産生する任意の技術を用いて作製することが可能である。これらの技術には、ハイブリドーマ技術、ヒトB細胞ハイブリドーマ技術、及びEBV-ハイブリドーマ技術が含まれるが、これらに限定されるものではない(例えば、Kohler, G. 等. (1975) *Nature* 256:495-497; Kozbor, D. 等. (1985) *J. Immunol. Methods* 81-8-42; Cote, R.J. 等. (1983) *Proc. Natl. Acad. Sci.* 80:2026-2030; Cole, S.P. 等. (1984) *Mol. Cell Biol.* 62:109-120を参照)。

20

【0151】

更に、「キメラ抗体」作製のために発達したヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの技術が、好適な抗原特異性及び生物学的活性を備える分子を得るために用いられる(例えば、Morrisson, S.L. 他. (1984) *Proc. Natl. Acad. Sci.* 81-4851-4855; Neuberger, M.S. 他. (1984) *Nature* 312:604-608; Takeda, S. 等. (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、ETRN特異性一本鎖抗体を生成する。関連する特異性を備えるが別のイディオタイプの組成の抗体は、ランダムな組み合わせの免疫グロブリンライブラリから鎖混合によって生成することもできる(例えば、Burton D.R. (1991) *Proc. Natl. Acad. Sci.* 88:11120-3を参照)。

30

【0152】

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

40

【0153】

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

【0154】

50

種々のイムノアッセイを用いてスクリーニングし、所望の特異性を有する抗体を同定する。隔離された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる競合的な結合、または免疫放射線活性のための数々のプロトコルが、当分野では周知である。通常このようなイムノアッセイには、ETRN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ETRN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)。

【0155】

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、ETRNに対する抗体の親和性を評価する。親和性を結合定数 K_a で表すが、この K_a は、平衡状態の下でETRN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ETRNエピトープに対して親和性が不均一なポリクローナル抗体医薬の K_a は、ETRN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ETRNエピトープに単一特異的なモノクローナル抗体医薬の K_a は、親和性の真の測定値を表す。 K_a 値が $10^9 \sim 10^{12}$ L/molの高親和性抗体医薬は、ETRN抗体複合体が激しい操作に耐えなければならないイムノアッセイに用いるのが好ましい。 K_a 値が $10^6 \sim 10^7$ L/molの低親和性抗体医薬は、ETRN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)。

【0156】

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

【0157】

本発明の別の実施例では、ETRN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、ETRN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、ETRN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。

【0158】

治療に用いる場合、アンチセンス配列を好適な標的細胞に導入するのに好適な任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコードする細胞配列の少なくとも一部に相補的な配列を発現する発現プラスミドの形で細胞内に送達することができる(例えば、Slater, J. E. 他(1998) J. Allergy Clin. Immunol. 102(3): 469-475; and Scanlon, K. J. 他(1995) 9(13): 1288-1296.を参照)。また、アンチセンス配列は、例えばレトロウイルスやアデノ関連ウイルスベクター等のウイルスベクターを用いて細胞内に導入することもできる(例えば、Miller, A. D. (1990) Blood 76: 271; Ausubel, 前出; Ucker, W. and W. Walther (1994) Pharmacol. Ther

10

20

30

40

50

. 63 (3) : 323 - 347 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、及び当分野で周知のその他の系が含まれる (Rossi, J. J. (1995) Br. Med. Bull. 51 (1) : 217 - 225; Boado, R. J. 他 (1998) J. Pharm. Sci. 87 (11) : 1308 - 1315; and Morris, M. C. 他 (1997) Nucleic Acids Res. 25 (14) : 2730 - 2736. を参照)。

【0159】

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

【0160】

本発明の更なる実施例では、ETRNの欠損による疾患や異常症は、ETRN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ETRN欠損細胞に導入することによって治療する。in vivo 或いは ex vitro の細胞に用いる機械的な導入技術には、(i) 個々の細胞内へのDNAのマイクロインジェクション、(ii) 金粒子の打ち込み、(iii) リポソーム仲介性トランスフェクション、(iv) 受容体仲介性遺伝子導入、及び (v) DNAトランスポソン (Morgan, R. A. and W. F. Anderson (1993) Annu. Rev. Biochem. 62 : 191 - 217; Ivics, Z. (1997) Cell 91 : 501 - 510; Boulay, J. L. and H. Recipon (1998) Curr. Opin. Biotechnol. 9 : 445 - 450) の使用が含まれる。

【0161】

ETRNの発現に影響を及ぼし得る発現ベクターには、限定するものではないが、PCDNA 3.1、EPITAG、PRCCMV2、PREP、PVAXベクター (Invitrogen, Carlsbad CA)、PCMV-SCRIPT、PCMV-TAG、PEGSH/PERV (Stratagene, La Jolla CA)、PTET

- OFF、PTET-ON、PTRE2、PTRE2-LUC、PTK-HYG (Clontech, Palo Alto CA) が含まれる。ETRNを発現させるために、(i) 恒常的に活性なプロモーター (例えば、サイトメガロウイルス (CMV)、ラウス肉腫ウイルス (RSV)、SV40ウイルス、チミジンキナーゼ (TK)、若しくは - アクチン遺伝子等)、(ii) 誘導性プロモーター (例えば、市販されている T-REX プラスミド (Invitrogen) に含まれている、テトラサイクリン調節性プロモーター (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. U.S.A. 89:5547-5551; Gossen, M. 他 (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456))、エクジソン誘導性プロモーター (市販されているプラスミド PVGRXR 及び PIND に含まれている: Invitrogen)、FK506/ラパマイシン誘導性プロモーター、または RU486/ミフェプリスト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 ETRN 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。

【0162】

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

【0163】

本発明の別の実施例では、ETRN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 ETRN をコードするポリヌクレオチドと、(ii) 好適な RNA パッケージングシグナルと、(iii) 追加のレトロウイルス・シス作用性 RNA 配列及び効率的なベクターの増殖に必要なコーディング配列を伴う Rev 応答性エレメント (RRE) とからなるレトロウイルスベクターを作製して治療することができる。レトロウイルスベクター (例えば、PFB 及び PFBNEO) は Stratagene 社から入手可能であり、公表データ (Riviere, I. 他. (1995) Proc. Natl. Acad. Sci. U.S.A. 92:6733-6737) に基づいている。上記データを引用することをもって本明細書の一部とする。このベクターは、VSVg (Armentano, D. 他 (1987) J. Virol. 61:1647-1650; Bender, M.A. 他 (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. 他 (1998) J. Virol. 72:8463-8471; Zufferey, R. 他 (1998) J. Virol. 72:9873-9880) 等の乱交雑エンベロープタンパク質若しくは標的細胞上の受容体に対する親和性を有するエンベロープ遺伝子を発現する好適なベクター産生細胞系 (VPCL) において増殖される。RIGG に付与された米国特許第 5,910,434 号 (「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」) において、レトロウイルスパッケージング細胞系を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、ある細胞集団 (例えば、CD4⁺T細胞) の形質導入、並びに形質導入した細胞を患者に戻

す方法は、遺伝子治療の分野では周知であり、多数の文献に記載されている (Ranga, U. 他 (1997) *J. Virol.* 71:7020-7029; Bauer, G. 他 (1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. 他 (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290)。

【0164】

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、ETRNの発現に関連する1
 10 あるいは複数の遺伝子異常を有する細胞にETRNをコードするポリヌクレオチドを送達する。アデノウイルス系ベクターの作製及びパッケージングは当分野では周知である。複製欠損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を脾臓の無損傷の脾島の中に導入するために可変性であることが証明された (Csete, M.E. 他 (1995) *Transplantation* 27:263-268)。使用できる可能性のあるアデノウイルスベクターが、米国特許第5,707,618号 (「Adenovirus vectors for gene therapy」) に記載されており、引用することをもって本明細書の一部とする。アデノウイルスベクターについてはまた、Antinozzi, P.A. 他 (1999) *Annu. Rev. Nutr.* 19:511-544; and Verma, I.M. and N. Somia (1997) *Nature* 18:389:239-242を参照し、引用することをもって
 20 本明細書の一部とする。

【0165】

別法では、ヘルペス系遺伝子治療の送達系を用いて、ETRNの発現に関連する1
 10 あるいは複数の遺伝子異常を有する標的細胞にETRN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ETRNを導入する際に特に重要である。ヘルペス系ベクターの作製及びパッケージングは当分野では周知である。複製適格性の単純疱疹ウイルス (HSV) I型系のベクターは、霊長類の眼にレポーター遺伝子を送達するために用いられてきた (Liu, X. 他 (1999) *Exp. Eye Res.* 169:385-395)。HSV-1ウイルスベクターの作製は、DeLucaに付与された米国特許第5,804,413号 (*Herpes simplex virus swains for gene transfer*) に記載されており、引用することをもって本明細書の一部とする。米国特許第5,804,413号には、ヒト遺伝子治療を含む目的のために、好適なプロモーターのコントロールの下で、細胞に導入される少なくとも1つの内在性遺伝子を含むゲノムからなる組換えHSV
 30 d92についての記載がある。また上記特許には、ICP4、ICP27及びICP22のために除去される組換えHSV株の作製及び使用方法が開示されている。HSVベクターについては、Goins, W.F. 他 (1999) *J. Virol.* 73:519-532 and Xu, H. 他 (1994) *Dev. Biol.* 163:152-161を参照し、引用することをもって本明細書の一部とする。クローニングされたヘルペスウイルス配列の操作や、巨大ヘルペスウイルスのゲノムの異なった部分を含む多
 40 数のプラスミドをトランスフェクトした後の組換えウイルスの継代、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は当分野で周知の技術である。

【0166】

別法では、ウイルス (正の一本鎖RNAウイルス) ベクターを用いてETRNをコードするポリヌクレオチドを標的細胞に送達する。プロトタイプのウイルスであるセムリキ森林熱ウイルス (*Semliki Forest Virus, SFV*) の生物学的な研究が広範に行われ、遺伝子伝達ベクター (*gene transfer vector*) がSFVゲノムに基づいていることが分かった (Garoff, H. and K. - J. Li (1998) *Cun. Opin. Biotech.* 9:464-469)。
 ウイルスRNAの複製中に、通常はウイルスカプシドタンパク質をコードするサブゲノム
 50

RNAが作り出される。このサブゲノムRNAが完全長のゲノムRNAより高いレベルで複製されるため、酵素活性（例えばプロテアーゼ及びポリメラーゼ）を有するウイルスタンパク質に対してカプシドタンパク質が過剰に産生される。同様に、ETRNをコードする配列をウイルスゲノムのカプシドをコードする領域に導入することによって、ベクター導入細胞において多数のETRNをコードするRNAが産生され、高いレベルでETRNが合成される。通常はウイルス感染は2～3日以内の細胞溶解に関係するが、シンドビスウイルス(SIN)の変異体を有するハムスターの正常な腎細胞(BHK-21)の持続的な感染を確立する能力は、ウイルスの溶解性の複製が遺伝子治療に適用できるように好適に変更することが可能であることを示唆している(Dryga, S.A. 他. (1997) *Virology* 228:74-83)。様々な宿主にウイルスを導入できることから、様々なタイプの細胞にETRNを導入することができる。ある集団における細胞のサブセットの特定の形質導入には、形質導入する前に細胞のソーティングを必要とする場合がある。ウイルスの感染性cDNAクローンの操作、ウイルスcDNA及びRNAのトランスフェクション、並びにウイルスの感染方法は当分野で周知である。

10

20

30

40

50

【0167】

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

【0168】

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

【0169】

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

【0170】

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

【0171】

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

【0172】

本発明の更なる実施例は、ETRNをコードするポリヌクレオチドの発現の変化に有効な化合物をスクリーニングする方法を含む。特定のポリヌクレオチドの発現の変化に有効な化合物には、限定するものではないが、特定のポリヌクレオチド配列と相互作用可能な非高分子化学物質、オリゴヌクレオチド、アンチセンスオリゴヌクレオチド、三重らせん形成オリゴヌクレオチド、転写因子やその他のポリペプチド転写調節因子が含まれる。有効な化合物は、ポリヌクレオチド発現のインヒビター或いはエンハンサーとして作用し、ポリヌクレオチドの発現を変化させ得る。従って、ETRN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、ETRN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、ETRN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、ETRN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。

【0173】

特定のポリヌクレオチドの発現の変化の有効性を調べるために、少なくとも1個から複数個の試験化合物をスクリーニングすることができる。試験化合物は、有効な化合物の化学修飾を含む当分野で周知の任意の方法で得ることができる。このような方法は、ポリヌクレオチドの発現を変化させる場合、一般に市販されている或いは専売の天然または非天然の化合物ライブラリから選択する場合、標的ポリヌクレオチドの化学的及び/または構造的な特性に基づいて化合物を合理的にデザインする場合、更に組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合に有効である。ETRNをコードするポリヌクレオチドを含むサンプルは、少なくとも1つの試験化合物に曝露して得る。サンプルには、例えば無傷細胞、透過処理した細胞、*in vitro*細胞遊離系または再構成生化学系が含まれ得る。ETRN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、ETRNをコードするポリヌクレオチドの配列に相補的なヌクレオチド配列を有するプローブを用いたハイブリダイゼーションにより、特定のヌクレオチドの発現を検出する。ハイブリダイゼーションの収量を定量し、その値が1或いは複数の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較における基準となり得る。試験化合物に曝露されるポリヌクレオチドの発現の変化が検出される場合は、ポリヌクレオチドの発現の変化に試験化合物が有効であることを示している。特定のポリヌクレオチドの発現の変化に有効な化合物を調べるために、例えば *Schizosaccharomyces pombe* 遺伝子発現系 (Atkins, D. 他 (1999) 米国特許第5,932,435号、Arndt, G.M. 他 (2000) *Nucleic Acids Res.* 28:E15) または HeLa 細胞等のヒト細胞株 (Clarke, M.L. 他 (2000) *Biochem. Biophys. Res. Commun.* 268:8-13) を用いてスクリーニングする。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性を調べるための、各オリゴヌクレオチド (デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、及び修飾オリゴヌクレオチド) の組み合わせライブラリのスクリーニングを含む (Bruice, T.W. 他 (1997) 米国特許第5,686,242号、Bruice, T.W. 他 (2000) 米国特許第6,022,691号)。

【0174】

10

20

30

40

50

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

【0175】

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

10

【0176】

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

【0177】

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

20

【0178】

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

30

【0179】

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

【0180】

医薬組成物の特殊な形状は、ETRN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調製される。例えば、細胞不透過性高分子を含むリポソーム製剤は、細胞融合及び高分子の細胞内輸送を促進し得る。別法では、ETRNまたはその断片をHIV Tat-1タンパク質の陽イオンN末端部に結合することもできる。このようにして作製された融合タンパク質は、マウスモデル系の脳を含む全ての組織の細胞に形質導入されることが確認されている（Schwarze, S. R. 他 (1999) *Science* 285: 1569 - 1572）。

40

【0181】

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

50

【0182】

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

10

【0183】

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

20

【0184】

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

【0185】

(診断)

別の実施例では、ETRNに特異的に結合する抗体が、ETRNの発現によって特徴付けられる疾患の診断、またはETRNやETRNのアゴニストまたはアンタゴニスト、インヒビターで治療を受けている患者をモニターするためのアッセイに用いられる。診断に有用な抗体は、治療のところで記載した方法と同じ方法で製剤される。ETRN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ETRNを検出する方法が含まれる。これらの抗体は、修飾をして或いはしないで使用され、レポーター分子の共有結合性或いは非共有結合性の接着によって標識化され得る。当分野で周知の種々のレポーター分子が用いられるが、その内の幾つかは上記した。

30

【0186】

ETRN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ETRN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ETRN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ETRNに対する抗体とを結合させることによって決定する。標準的な複合体形成の量は、測光法(photometric)などの種々の方法で定量され得る。被験者のETRNの発現の量、制御及び疾患、生検組織からのサンプルが基準値と比較される。基準値と被験者との間の偏差が、診断の指標となる。

40

【0187】

別の実施例によれば、ETRNをコードするポリヌクレオチドを診断のために用いることもできる。用いられるポリヌクレオチドには、オリゴヌクレオチド配列、相補的なRNA及びDNA分子、及びPNAが含まれる。このポリヌクレオチドを用いて、疾患と関連し

50

得る E T R N 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、E T R N の存在の有無、更に過剰な発現を調べ、治療中の E T R N 値の調節を監視する。

【0188】

ある実施形態では、E T R N または近縁の分子をコードする遺伝子配列を含むポリヌクレオチド配列を検出可能な P C R プローブを用いたハイブリダイゼーションによって、E T R N 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 E T R N をコードする自然界の配列のみを同定するかどうか、或いはアレルや関連配列コードする自然界の配列のみを同定するかどうかによって決まるであろう。

10

【0189】

プローブはまた、関連する配列の検出に利用され、E T R N をコードする任意の配列と少なくとも 50% の配列同一性を有し得る。目的の本発明のハイブリダイゼーションプローブには、D N A あるいは R N A が可能であり、S E Q I D N O : 6 - 10 の配列、或いは E T R N 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。

【0190】

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

20

【0191】

E T R N をコードするポリヌクレオチド配列を用いて、E T R N 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、このような疾患の例には細胞増殖異常が含まれ、その中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病 (M C T D)、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺含まれ、また生殖障害が含まれ、その中には、プロラクチン産生異常と、卵管病、排卵異常、及び子宮内膜症、発情期異常、月経周期異常、多嚢胞卵巣症候群、卵巣過剰刺激症候群、子宮内膜癌及び卵巣癌、子宮筋腫、自己免疫異常、異所性妊娠、及び奇形発生を含む不妊症と、乳癌、線維嚢胞性乳腺症、及び乳漏症と、精子形成異常、生理学上の精子異常、精巣癌、前立腺癌、良性の前立腺過形成、前立腺炎、ペーロニー病、インポテンス、男性乳房癌及び女性化乳房とが含まれ、また免疫反応の疾患も含まれ、その中には後天性免疫不全症候群 (A I D S) 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 (A P E C E D)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時的リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋又は心膜炎、骨関節炎、骨粗しょう症、膵炎、多発性筋炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、血小

30

40

50

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、体外循環と、ウイルス感染症及び細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫の感染症が含まれる。ETRNをコードするポリヌクレオチド配列は、サザン法やノーザン法、ドットプロット法、或いはその他の膜系の技術、PCR法、ディップスティック(dipstick)、ピン(pin)、ELISA式アッセイ、及び変異ETRNの発現を検出するために患者から採取した体液或いは組織を利用するマイクロアレイに使用することが可能である。このような質的或いは量的方法は、当分野では周知である。

【0192】

ある実施態様では、ETRN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。ETRNをコードするヌクレオチド配列は、標準的な方法で標識化され、ハイブリダイゼーション複合体の形成に好適な条件の下、患者から採取した体液或いは組織のサンプルに加えることができるであろう。好適な培養期間の後、サンプルを洗浄し、シグナルを定量して基準値と比較する。患者のサンプルのシグナルの量が、制御サンプルと較べて著しく変わっている場合は、サンプル内のETRNをコードするヌクレオチド配列の変異レベルにより、関連する疾患の存在が明らかになる。このようなアッセイを用いて、動物実験、臨床試験、或いは個人の患者の治療を監視における、特定の治療効果を推定することが可能である。

10

【0193】

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

20

【0194】

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

30

【0195】

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

【0196】

ETRNをコードする配列から設計されたオリゴヌクレオチドのさらなる診断への利用には、PCRの利用が含まれ得る。このようなオリゴマーは、化学的な合成、酵素を用いた生成、或いは*in vitro*で生成され得る。オリゴマーは、好ましくはETRNをコードするポリヌクレオチドの断片、或いはETRNをコードするポリヌクレオチドと相補的なポリヌクレオチドの断片を含み、最適な条件の下、特定の遺伝子や条件を識別するために利用される。また、オリゴマーは、やや緩いストリンジェントな条件の下、近縁のDNA或いはRNA配列の検出及び/または定量のため用いることが可能である。

40

【0197】

或る実施態様において、ETRNをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いて、一塩基多型(SNP)を検出し得る。SNPは、ヒトの先天性または後天性遺伝病の原因となる場合が多いヌクレオチドの置換、挿入及び欠失である。限定するものではないが、SNPの検出方法には、一本鎖立体構造多型(SSCP)及び蛍光SSCP(fSSCP)法が含まれる。SSCPでは、ETRNをコードするポリ

50

ヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いたポリメラーゼ連鎖反応 (PCR) で DNA を増幅する。この DNA は、例えば病変或いは正常な組織、生検サンプル、体液等に由来し得る。この DNA 内の SNP は、一本鎖形状の PCR 産物の 2 次及び 3 次構造に差異を生じさせる。この差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。fSCCP では、オリゴヌクレオチドプライマーを蛍光標識することによって、DNA シークエンシング装置などのハイスループット機器でアンプリマー (amplimer) の検出をすることが可能になる。更に、インシリコ SNP (in silico SNP: isSNP) と呼ばれる配列データベース分析法は、共通のコンセンサス配列の構築に用いられる個々の重複する DNA 断片の配列を比較することによって、多型を同定することができる。これらのコンピュータベースの方法は、DNA 配列クロマトグラムの自動分析及び統計モデルを用いたシークエンシングエラーや研究室での DNA の調整に起因する配列のばらつきを排除する。別法では、例えばハイスループットの MASSARRAY システム (Sequenom, Inc., San Diego CA) を用いた質量分析により SNP を検出し、特徴付ける。

10

20

30

40

50

【0198】

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

【0199】

更に別の実施例では、本明細書に記載した任意のポリヌクレオチド配列に由来するオリゴヌクレオチドまたはより長い断片を、マイクロアレイにおける標的として用いることができる。マイクロアレイを転写イメージング技術に用いて、多数の遺伝子の相対発現レベルを同時にモニタリングすることができる。これについては、Seilhamer, J. J. 他に付与された米国特許第 5,840,484 号 (名称「Comparative Gene Transcript Analysis」) に記載されており、この引用を以って本明細書の一部とする。マイクロアレイはまた、遺伝子変異、突然変異及び多型の同定に用いることができる。この情報を用いて、遺伝子機能を決定し、疾患の遺伝的根拠を解明し、疾患を診断し、遺伝子発現に関連する疾病の進行/後退をモニタリングし、疾患の治療における治療薬の開発や活性のモニタリングを行うことができる。特に、患者にとって最適かつ有効な治療法を選択するために、この情報を用いて患者の薬理ゲノムプロファイルを作成することができる。例えば、患者の薬理ゲノムプロファイルに基づいて、患者に対して極めて効果的でありながら副作用を殆ど示さない治療薬を選択することができる。

【0200】

別の実施例では、ETRN に特異的な抗体、ETRN またはその断片をマイクロアレイ上でエレメントとして用いることができる。マイクロアレイを用いて、上記のようにタンパク質間相互作用、薬剤-標的相互作用及び遺伝子発現プロファイルをモニタリング及び測定することが可能である。

【0201】

或る実施例は、或る組織または細胞型の転写イメージを生成する本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞型により遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所定の条件下で所定の時間に発現した遺伝子の数及び相対存在量を定量することにより分析される (Seilliamer 他、米国特許第 5,840,484 号の "Comparative Gene Transcript Analysis" を参照。この特許に言及することを以って本明細書の一部

とする)。従って、特定の組織または細胞型の転写物全体または逆転写物全体に本発明のポリヌクレオチドまたはその相補配列をハイブリダイズすることにより、転写イメージが生成され得る。或る実施例では、本発明のポリヌクレオチドまたはその相補配列がマイクロアレイ上に複数のエレメントのサブセットを構成するハイスループット型でハイブリダイゼーションさせる。結果として得られる転写イメージは、遺伝子活性のプロフィールとなり得る。

【0202】

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

10

【0203】

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

20

30

【0204】

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

40

【0205】

別の実施例は、本発明のポリペプチド配列を用いて組織または細胞型のプロテオームを分析することに関連する。「プロテオーム」という用語は、特定のある組織または細胞型におけるタンパク質発現の包括的パターンを指す。プロテオームを構成する各タンパク質は、更に個々に分析することができる。プロテオーム発現パターン即ちプロフィールは、所定の条件下で所定の時間に発現したタンパク質の数及びそれらの相対的な存在量を定量することにより分析する。従って、ある細胞のプロテオームのプロフィールは、特定の組織または細胞型のポリペプチドを分離及び分析することにより作成し得る。或る実施例では

50

、このような分離は2次元ゲル電気泳動によって行う。この2次元ゲル電気泳動法では、まず、1次元の等電点電気泳動によりサンプルからタンパク質を分離し、次に、2次元のドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に従って分離する（前出のSteiner and Anderson）。これらのタンパク質は、通常クーマシーブルーまたはシルバーまたは蛍光染色などの染色剤を用いてゲルを染色して、分散した個別の位置にあるスポットとしてゲル中で可視化される。各タンパク質スポットの光学密度は、通常サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処置済みまたは未処置のいずれかの生体サンプルから得られる等位置にあるタンパク質スポットの光学密度を比較し、処置に関連するタンパク質スポット密度の変化を調べる。スポット内のタンパク質は、例えば化学的または酵素的に切断した後、質量分析する標準的な方法を用いて部分的にシーケンシングする。スポット内のタンパク質の同一性は、好適には少なくとも5個の連続するアミノ酸残基であるその部分的な配列を、本発明のポリペプチド配列と比較することにより決定し得る。場合によっては、決定的なタンパク質同定のための更なる配列が得られる。

10

【0206】

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

20

【0207】

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

30

【0208】

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

40

【0209】

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

【0210】

50

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

【0211】

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

【0212】

in situ 蛍光ハイブリダイゼーション(FISH)は、他の物理的及び遺伝子地図データと相関し得る(例えば、Heinz-Ulrich, 他による(1995) in Meyers, 前出, pp. 965-968を参照)。遺伝子地図データの例は、種々の科学誌あるいはOnline Mendelian Inheritance in Man(OMIM)のワールドワイドウェブのサイトで見付けることができる。物理的な染色体地図上のETRN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。

【0213】

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

る染色体位置の違いを検出することもある。

【0214】

本発明の別の実施例では、ETRN、その触媒作用断片或いは免疫原断片またはそのオリゴペプチドを、種々の任意の薬剤スクリーニング技術における化合物のライブラリのスクリーニングに用いることができる。このようなスクリーニングに用いる断片は、溶液に遊離、固体支持物に固定、細胞の表面上に保持、或いは細胞内に存在する。ETRN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。

【0215】

薬剤スクリーニングに用いる別の方法は、目的のタンパク質に対して、好適な結合親和性を有する化合物のスクリーニング処理能力を高めるために用いられる（例えば、Geysen, 他による(1984) PCT出願番号 WO 84/03564 を参照）。この方法では、相当な数の異なる小さな試験用化合物が、プラスチックピン或いは他の基板の上に合成される。試験用化合物は、ETRN、或いはその断片と反応してから洗浄される。次に、結合されたETRNが、当分野で周知の方法で検出される。精製されたETRNはまた、前記した薬剤をスクリーニングする技術に用いられるプレート上で直接被覆することもできる。別法では、非中和抗体を用いて、ペプチドを捕らえ、固体支持物に固定することもできる。

10

【0216】

別の実施例では、ETRNと結合可能な中和抗体がETRN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、ETRN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。

20

【0217】

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

【0218】

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

30

【0219】

前述した及び以下に記載した全ての特許出願、特許、刊行物、特に米国特許出願通し番号60/140,109に言及することをもって本明細書の一部とする。

【0220】

(実施例)

1 cDNAライブラリの作製

RNAは、Clontech社から購入、或いは表4に列記した組織から単離した。まず、この組織の一部をホモジナイズしてグアニジニウムイソチオシアネート溶液に溶解する一方、この組織の別の一部をホモジナイズしてフェノールに溶解するか、或いはTRIzol (Life Technologies)、グアニジニウムイソチオシアネート及びフェノールの単相溶液などの好適な変性剤の混合液に溶解した。この溶解物を塩化セシウムにおいて遠心分離によって、或いはクロロホルムで抽出した。イソプロパノール或いは酢酸ナトリウムのどちらかとエタノール、或いは別の方法でこの溶解物からRNAを沈殿させた。

40

【0221】

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

50

RE mRNA精製キット(Ambion, Austin TX)などの別のRNA単離キットを用いて組織溶解物から直接単離した。

【0222】

ある場合には、Stratagene社にRNAを提供し、Stratagene社が対応するcDNAライブラリを作製した。そうでない場合は、UNIZAPベクターシステム(Stratagene)またはSUPERSCRIPTプラスミドシステム(Life Technologies)を用いて当分野で周知の推奨方法または類似の方法でcDNAを合成してcDNAライブラリを作製した。(例えば、Ausubel, 1997, 前出, ユニット5.1-6.6を参照)。逆転写は、オリゴd(T)またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖cDNAに結合させてから、好適な1つの制限酵素或いは複数の制限酵素でcDNAを消化した。殆どのライブラリでは、SEPHACRYL S 1000またはSEPHAROSE CL 2B、SEPHAROSE CL 4Bカラムクロマトグラフィー(Amersham Pharmacia Biotech)、アガロースゲル電気泳動法によってcDNAの大きさ(300~1000bp)を選択した。PBLUESCRIPTプラスミド(Stratagene)またはpSPORT1プラスミド(Life Technologies)、pcDNA2.1プラスミド(Invitrogen Carlsbad CA)、pINCYプラスミド(Incyte Pharmaceuticals, Palo Alto CA)などの好適なプラスミドのポリリンカーの適合性制限酵素部位にcDNAを結合させた。この組換えプラスミドを、Stratagene社のXL1-Blue, XL1-BlueMRF、SOLR、またはLife Technologies社のDH5 またはDH10B、ELECTROMAX DH10Bを含むコンピテント大腸菌細胞に導入し組み込んだ。

10

20

【0223】

2 cDNAクローンの単離

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

30

【0224】

別法では、ハイスループットの直接結合PCR法によって宿主細胞溶解物からプラスミドDNAを増幅した。(Rao, V.B. (1994) Anal. Biochem. 216:1-14)。宿主細胞の溶解及び熱サイクリング過程を単一反応混合液で行った。サンプルを処理してから384-ウェルプレートに移して保管し、増幅したプラスミドDNAの濃度をPICOGREEN色素(Molecular Probes, Eugene OR)及びFluoroskan II蛍光スキャナー(Labsystems Oy, Helsinki, Finland)を用いて蛍光定量的に測定した。

40

【0225】

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において回収したインサイト社cDNAは、以下に示すようにシークエンシングした。cDNAのシークエンシング反応は、標準的な方法で、或いはABI CATALYST 800 (PE Biosystems) thermal cyclerまたはPTC-200 thermal cycler (MJ Research)と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の

50

組み合わせなどのハイスループット装置で行った。cDNAのシーケンシング反応の準備には、Amersham Pharmacia Biotech社の試薬、またはABI PRISM BIGDYE Terminator cycle sequencing ready reactionキット(PE Biosystems)などのABIシーケンシングキットに含まれる試薬を用いた。cDNAのシーケンシング反応の電気泳動的な分離及び標識したポリヌクレオチドの検出には、MEGABACE 1000 DNAシーケンシングシステム(Molecular Dynamics)、標準ABIプロトコル及び塩基対呼び出しソフトウェアを用いるABI PRISM 373または377シーケンシングシステム(PE Biosystems)、当分野で周知のその他の配列解析システムを用いた。cDNA配列の読み枠は、標準的な方法(Ausubel, 1997, 前出, unit 7.7)を用いて決定した。cDNA配列の幾つかを選択して、本実施例の6に記載した方法で配列を伸長した。

【0226】

cDNAのシーケンシングから得たポリヌクレオチド配列の構築及び解析は、当分野の技術者に周知のアルゴリズムを利用したソフトウェアを組合せて行った。表5は、利用したツール、ソフトウェア、アルゴリズム、それらの説明、引用文献、閾値パラメーターの概要を示す。表5の列1は用いたツール及びプログラム、アルゴリズム、列2はそれらの簡単な説明、列3は引用することで本明細書の一部とした引用文献、列4の記載部分は2つの配列の一致度の評価に用いたスコア及び確率値、他のパラメータを示す(確率値が高ければ高いほど配列間の相同性が高くなる)。配列の解析には、MACDNASIS PROソフトウェア(Hitachi Software Engineering, S. San Francisco CA)及びLASERGENEソフトウェア(DNA STAR)を用いた。

【0227】

ポリヌクレオチド配列の確証は、BLAST及び動的計画法、ジヌクレオチド最近接分析に基づいたプログラム及びアルゴリズムを用いて、ベクター及びリンカー、ポリA配列を取り除き、あいまいな塩基対をマスクすることで行った。次に、BLAST及びFASTA、BLIMPSに基づいたプログラムを用いて、公共のデータベースであるGenBankの霊長類及びげっ歯類、哺乳類、脊椎動物、真核生物のデータベースやBLOCKS、PRINTS、DOMO、PRODOM及びPFAMなどのデータベースから選択した配列に対してこれらの配列を問合わせて注釈を得た。Phred及びPhrap、Consedに基づいたプログラムを用いて完全長のポリヌクレオチド配列の中にこれらの配列を構築して、BLAST及びFASTA、BLIMPSに基づいたプログラムでオープンリーディングフレームのためにスクリーニングした。完全長のポリヌクレオチド配列を翻訳して対応する完全長のアミノ酸配列を引き出し、GenBankデータベース(上記)及びSwissProt、BLOCKS、PRINTS、DOMO、PRODOM及びPrositeなどのデータベース、またはPFAMなどの隠れマルコフモデル(HMM)に基づいたタンパク質ファミリーデータベースに対して問い合わせてこれらの完全長の配列を分析した。HMMは、確率を利用して遺伝子ファミリーのコンセンサス一次構造を解析する(例えば、Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365を参照)。

【0228】

完全長のポリヌクレオチド配列及びアミノ酸配列の構築及び分析に用いる上記のプログラムは、SEQ ID NO: 6-10からのポリヌクレオチド配列の断片の同定にも使用できる。約20~4000個までのヌクレオチドの断片はハイブリダイゼーション及び増幅に有用であり、上記の発明で説明した。

【0229】

4 ポリヌクレオチド発現の分析

ノーザン分析は、遺伝子の転写物の存在を検出するために用いられる実験用技術であり、特定の細胞種或いは組織からのRNAが結合されている膜への標識されたヌクレオチド配

列のハイブリダイゼーションを伴う（例えば、Sambrook, 前出, 7章; 及び Ausubel, F.M., 他, 前出, 4章及び16章を参照）。

【0230】

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

【0231】

【数1】

(BLAST スコア×配列一致率)

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

【0232】

として定義される積スコアである。積スコアは、0～100の標準化された値であり、以下のように求める。BLASTスコアにヌクレオチド配列の一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除する。高スコアのセグメントの対(HSP)において一致する各塩基に+5のスコアを割り当て、各不適性塩基対に-4を割り当てることにより、BLASTスコアを計算する。2つの配列は、2以上のHSPを共有し得る(ギャップにより離隔される)。2以上のHSPがある場合には、最高BLASTスコアの塩基対を用いて積スコアを計算する。積スコアは、BLASTアラインメントの断片的重複と質とのバランスを表す。例えば積スコア100は、比較した2つの配列の短い方の長さ全体にわたって100%一致する場合にのみ得られる。積スコア70は、100%の一致で一端が70%重畳しているか、或いは88%一致で他端が100%重畳しているかの何れかの場合である。積スコア50は、100%の一致で一端が50%重畳しているか、或いは79%の一致で他端が100%重畳しているかの何れかの場合である。

【0233】

ノーザン分析の結果は、ETRNをコードする転写物が発生したライブラリの分布割合として報告される。分析には、器官/組織及び疾患によるcDNAライブラリの分類も含まれる。器官/組織のカテゴリーには、心血管、皮膚、発生、内分泌、胃腸、造血/免疫、筋骨格、神経、生殖、泌尿器が含まれる。疾患のカテゴリーには、癌、炎症、外傷、細胞増殖、神経、プール(pool)が含まれる。カテゴリー別に、目的の配列を発現するライブラリの数を数えて、それを全ての範囲のライブラリの数で除した。各組織に特異的に発現する割合(パーセント)と各疾患で発現する割合を表3に示した。

【0234】

5 ETRN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: 6-10を構築するために用いたcDNA配列を、BLAST及びSmith-Watermanアルゴリズムを用いて、インサイト社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: 6-10と一致するこれらのデータベースの配列を、Phrap(表5)などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。Stanford Human Genome Center (SHGC)、Whitehead Institute for Genome Research (WIGR) 及びGenethonなどの公共の情報源から入手できる放射線ハイブリッド(radiation hybrid)及び遺伝子マッピングのデータを用いて、クラスター化した配列がすでにマッピングされているかを調べる。クラスターにマッピングされた配列が含まれている場合は、そのクラスターの全ての配列(特定のSEQ ID NOを含む)をそのマッピング位置に割り当てた。

10

20

30

40

50

【0235】

SEQ ID NO: 6 及び SEQ ID NO: 9 の遺伝子地図の位置は、ヒト染色体の区間即ち範囲として本明細書の発明の部分に記載した。SEQ ID NO: 9 では、2 つ以上のマッピング位置を記載したが、これは、SEQ ID NO: 9 と類似性を有する完全には一致していないが既にマッピングされた配列が、それぞれのクラスターに組み入れられたことを示唆するものである。センチモルガンで示したマッピング位置の範囲は、染色体の短腕 (p) の末端から測定した (センチモルガン (cM) は、同一染色体上の遺伝子間の乗換え率に基づいた距離を表す単位である。平均すると、1 cM はヒトの染色体の 1 メガベースに概ね等しいが、組換え率の高い部分と低い部分があるため、大きく変化し得る)。距離 cM は、配列がそれぞれのクラスターに含まれている放射線ハイブリッドマーカーの境界を検出できる Genethon によってマッピングされた遺伝子マーカーに基づいている。示された区間内に位置する公共の配列及びインサイト社の配列に関連する疾患も、本明細書の該当する発明の部分に記載した。

10

【0236】

6 ETRN をコードするポリヌクレオチドの伸長

SEQ ID NO: 6 - 10 の完全長の核酸配列は、完全長分子の好適な断片から設計したオリゴヌクレオチドプライマーを用いてその完全長分子の好適な断片を伸長して作製した。一方のプライマーは既知の断片の 5' の伸長を開始するために合成し、他方のプライマーは既知の断片の 3' の伸長を開始するために合成した。開始プライマーは、OLIGO 4.06 ソフトウェア (National Biosciences) 或いは他の適切なプログラムを用いて、約 22 個から約 30 個のヌクレオチドの長さで約 50% 以上の GC 含量を有し、かつ約 68 ~ 72 の温度で標的配列にアニールするように設計した。ヘアピン構造及びプライマー-プライマー二量体が生じないようにヌクレオチドを伸長した。

20

【0237】

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

【0238】

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

30

ステップ 1 94 で 3 分間

ステップ 2 94 で 15 秒

ステップ 3 60 で 1 分間

ステップ 4 68 で 2 分間

40

ステップ 5 ステップ 2、3、及び 4 を 20 回繰り返す

ステップ 6 68 で 5 分間

ステップ 7 4 で保管

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

ステップ 1 94 で 3 分間

ステップ 2 94 で 15 秒

ステップ 3 57 で 1 分間

ステップ 4 68 で 2 分間

ステップ 5 ステップ 2、3、及び 4 を 20 回繰り返す

ステップ 6 68 で 5 分間

50

ステップ7 4 で保管。

【0239】

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PICO GREEN定量試薬(0.25% (v/v) PICO GREEN; Molecular Probes, Eugene OR)を不透明な蛍光光度計プレート(Coming Costar, Acton MA)の各ウェルに分配してDNAが試薬と結合できるようにして測定する。このプレートをFluoroskan II (Labsystems Oy, Helsinki, Finland)でスキャンして、サンプルの蛍光を計測してDNAの濃度を定量化する。反応混合物の5~10µlのアリコットを1%のアガロースミニゲル上での電気泳動によって解析し、何れの反応物が配列を伸長することに成功したかを決定する。 10

【0240】

伸長したヌクレオチドを脱塩及び濃縮してから384ウェルプレートに移し、CviJIコレラウイルスエンドヌクラーゼ(Molecular Biology Research, Madison WI)で消化し、pUC18ベクター(Amersham Pharmacia Biotech)に再連結する前に音波処理またはせん断を行った。ショットガンシーケンシングのために、消化したヌクレオチドを低濃度(0.6~0.8%)のアガロースゲル上に分離して断片を切断し、寒天をAgar ACE (Promega)で消化した。T4リガーゼ(New England Biolabs, Beverly MA)を用いて伸長したクローンをpUC18ベクター(Amersham Pharmacia Biotech)に再連結し、Pfu DNAポリメラーゼ(Stratagene)で制限部位の伸び出しを処理してコンピテント大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りとってLB/2Xカルベニシリン培養液の384ウェルプレートに37 で一晩培養した。 20

【0241】

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 30

ステップ3 60 で1分間

ステップ4 72 で2分間

ステップ5 ステップ2、3、及び4を29回繰り返す

ステップ6 72 で5分間

ステップ7 4 で保管。

上記したようにPICO GREEN試薬(Molecular Probes)でDNAを定量化した。DNA回収率の悪いサンプルは、上記した条件で再び増幅した。サンプルを20%のジメチルサルホサイド(dimethylsulphoxide)(1:2, v/v)で希釈し、DYENAMIC DIRECTキット(Amersham Pharmacia Biotech)またはABI PRISM BIGDYE Terminator cycle sequencing ready reactionキット(PE Biosystems)を用いてシーケンシングした。 40

【0242】

同様に上述の手順で、SEQ ID NO: 6-10のヌクレオチド配列を利用し、この伸長のために設計したオリゴヌクレオチドと好適なゲノムライブラリを用いて5 調節配列を得た。

【0243】

7 個々のハイブリダイゼーションプローブの標識化及び使用法

SEQ ID NO: 6-10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 50

るオリゴヌクレオチドの標識について特に記すが、より大きなcDNAフラグメントの場合でも基本的に同じ手順を用いる。オリゴヌクレオチドを、OLIGO 4.06ソフトウェア(National Bioscience)のような最新式のソフトウェアを用いてデザインし、50 pmolの各オリゴマーと、250 μ Ciの[³²P]アデノシン三リン酸(Amersham, Chicago, IL)及びT4ポリヌクレオチドキナーゼ(DuPont NEN, Boston MA)とを組み合わせることで用いることにより標識する。標識されたオリゴヌクレオチドを、SEPHADEX G-25超精細排除デキストランビードカラム(Amersham Pharmacia Biotech)を用いて実質的に精製する。毎分10⁷カウントの標識されたプローブを含むアリコットを、次のエンドヌクレアーゼ、Ase I、Bgl II、Eco RI、Pst I、Xba I 10
 あるいはPvu II(DuPont NEN)の1つを用いて切断したヒトゲノムDNAの典型的な膜ベースのハイブリダイゼーション解析において用いる。

【0244】

各切断物からのDNAを、0.7%アガロースゲル上で分画して、ナイロン製メンブラン(Nytran Plus, Schleicher & Schuell, Durham NH)に転写する。ハイブリダイゼーションは40℃で16時間かけて行う。非特異的シグナルを取り除くため、例えば、最大0.1Mクエン酸ナトリウム食塩水及び0.5%ドデシル硫酸ナトリウムの条件の下、プロットを順次室温にて洗浄する。ハイブリダイゼーションパターンをオートラジオグラフィー或いは別のイメージ化手段で視覚化して比較する。 20

【0245】

8. マイクロアレイ

マイクロアレイ上のアレイエレメントの連結または合成は、フォトリソグラフィ、ピエゾプリント(インクジェットプリンター、前出のBaldeschweiler等を参照)、機械的マイクロスポッティング技術及びこれらから派生したものをを用いて達成することが可能である。上記各技術において基板は、均一な非多孔性の固体とするべきである(Schena(1999), 前出)。推奨する基板には、シリコン、シリカ、スライドガラス、ガラスチップ及びシリコンウエハがある。別法では、ドットプロット法またはスロットプロット法に類似のアレイを利用して、熱や紫外線、または化学的或いは機械的な結合手段で基板の表面にエレメントを配置して結合させることができる。通常のアレイは利用可能な方法や機械を用いて作製でき、任意の適正な数のエレメントを含めることができる(Schena, M. 他(1995) Science 270: 467-470、Shalon, D. 他(1996) Genome Res. 6: 639-645、Marshall, A. and J. Hodgson(1998) Nat. Biotechnol. 16: 27-31.を参照)。 30

【0246】

完全長cDNA、発現遺伝子配列断片(EST)、或いはそれらの断片やオリゴマーが、マイクロアレイのエレメントとなり得る。ハイブリダイゼーションに好適な断片やオリゴマーを、LASERGENEソフトウェア(DNA STAR)などの当分野で周知のソフトウェアを用いて選択することが可能である。このアレイエレメントを、生体サンプル中のポリヌクレオチドとハイブリダイズさせる。生体サンプル中のポリヌクレオチドは、検出を容易にするために蛍光標識またはその他の分子タグに結合する。ハイブリダイゼーションの後、生体サンプルからハイブリダイズしなかったヌクレオチドを除去し、蛍光スクリーニングを用いて各アレイエレメントにおけるハイブリダイゼーションを検出する。別法では、レーザー脱離及び質量スペクトロメトリーを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上のエレメントにハイブリダイズする各ポリヌクレオチドの相補性の程度及び相対的存在量は、算定することができる。一実施例におけるマイクロアレイの調整及び使用について、以下に詳述する。 40

【0247】

組織または細胞サンプルの準備

10

20

30

40

50

グアニジウムチオシアネート法を用いて組織サンプルから全RNAを単離し、オリゴ(dT)セルロース法を用いてポリ(A)⁺RNAを精製する。各ポリ(A)⁺RNAサンプルは、MMLV逆転写酵素、0.05 pg/μlのオリゴ(dT)プライマー(21mer)、1×第1鎖緩衝液、0.03単位/μlのRNアーゼインヒビター、500 μM dATP、500 μM dGTP、500 μM dTTP、40 μM dCTP、40 μM dCTP-Cy3(BDS)またはdCTP-Cy5(Amersham Pharmacia Biotech)を用いて逆転写する。この逆転写反応は、GEMBRIGHTキット(Incycyte)を用いて、200 ngのポリ(A)⁺RNAを含む25 ml容量で行う。特異的なコントロールポリ(A)⁺RNAは、in vitro転写により非コーディング酵母ゲノムDNAから合成する。370 で2時間インキュベートした後、各反応サンプル(一方はCy3標識、他方はCy5標識)は、2.5 mlの0.5 M水酸化ナトリウムで処理し、850 で20分間インキュベートし、反応を停止させてRNAを変性する。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピナラム(CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA)を用いて精製する。結合後、2つの反応サンプルを、1 mlのグリコーゲン(1 mg/ml)、60 mlの酢酸ナトリウム及び300 mlの100%エタノールを用いてエタノール沈殿させる。サンプルは次に、Speed VAC(Savant Instruments Inc., Holbrook NY)を用いて乾燥して仕上げ、14 μl 5×SSC/0.2% SDS中で再懸濁する。

10

20

【0248】

マイクロアレイの準備

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

【0249】

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

30

【0250】

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

40

【0251】

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

50

【0252】

ハイブリダイゼーション

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

10

【0253】

検出

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

20

【0254】

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

30

【0255】

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

40

【0256】

光電子増倍管の出力は、IBMコンパチブルPCコンピュータにインストールされた12ビットRTI-835Hアナログ-デジタル(AID)変換ボード(Analog Devices, Inc., Norwood MA)を用いてデジタル化される。デジタル化されたデータは、リニア20色変換を用いてシグナル強度が青色(低シグナル)から赤色(高シグナル)までの擬似カラー範囲にマッピングされるイメージとして表示される。データはまた、定量的に分析される。2つの異なる蛍光体を同時に励起して測定する場合には、各蛍光体の発光スペクトルを用いて、先ずデータは蛍光体間の光学的漏話(重複発光スペクトルに起因する)に対して補正される。

【0257】

50

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

【0258】

9 相補的ポリヌクレオチド

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

【0259】

10 ETRN の発現

ETRN 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 ETRN 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、これらに限定されるものではない。組換えベクターを、BL21 (DE3) などの好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピル-Dチオガラクトピラノシド (IPTG) で誘発されると ETRN を発現する。真核細胞での ETRN 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、ETRN をコードする cDNA と置換する。ウイルスの感染力は維持され、強いポリヘドリンプロモーターによって高いレベルの cDNA の転写が行われる。組換えバキュロウイルスは、多くの場合は Spodoptera frugiperda (Sf9) 昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる。(例えば、Engelhard, E. K. 他 (1994) Proc. Natl. Acad. Sci. USA 91: 3224 - 3227; Sandig, V. 他 (1996) Hum. Gene Ther. 7: 1937 - 1945. を参照)。

【0260】

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

10

20

30

40

50

れている。これらの方法で精製した E T R N を直接用いて以下の実施例 1 1 及び 1 5 のアッセイを行うことができる。

【0261】

1 1 E T R N の活性の実証

E T R N の活性は、N A D P の酸化または還元によって実証することができる。基質として、A s n - G a l , b i o c y t i d i n e 、またはユビキノン - 1 0 などを用い得る。反応混合液は、1 - 2 m g / m l H O R P 、および 1 5 m M の基質と、0 . 1 M のリン酸緩衝液 (p H 7 . 1) 中の 2 . 4 m M N A D (P) H (酸化の場合) 、または 0 . 1 M の N a ₂ H P O ₄ 緩衝液 (p H 7 . 4) 中の 2 . 0 m M N A D (P) H (還元の場合) を含み合計 0 . 1 m l である。当分野で周知の方法に従って、N A D と共に F A D を含め得る。吸収率の変化を、記録式分光光度計で測定する。N A D (P) H の量は、初めに存在する基質の量に化学量論的に等しく、A ₃₄₀ における変化 (A ₃₄₀ = 6 6 2 0 [N A D H]) は、生成された N A D (P) H の量を直接示す。E T R N の活性は、本アッセイの N A D (P) H の存在量に比例する。E T R N の活性は、3 4 0 n m における N A D (P) H 補酵素の吸光係数に比例する。例えば、3 4 0 n m における N A D (P) H 補酵素の吸光係数の上昇は酸化活性を示し、この吸光係数の低下は還元活性を示す (D a l z i e l , K . (1 9 6 3) J . B i o l . C h e m . 2 3 8 : 2 8 5 0 - 2 8 5 8) 。

10

【0262】

1 2 機能のアッセイ

E T R N 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 E T R N をコードする配列の発現によって評価する。c D N A を、c D N A を高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにサブクローニングする。このようなベクターには、p C M V S P O R T ^{T M} (L i f e T e c h n o l o g i e s .) 及び p C R 3 . 1 (I n v i t r o g e n , C a r l s b a d , C A) が含まれ、どちらもサイトメガロウイルスプロモーターを含んでいる。5 ~ 1 0 μ g の組換えベクターを、例えば内皮由来か造血由来のヒト細胞株にリポソーム製剤或いは電気穿孔法によって一時的に形質移入する。更に、標識タンパク質をコードする配列を含む 1 ~ 2 μ g のプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入された細胞と形質移入されていない細胞とを区別できる。また、標識タンパク質の発現によって、c D N A の組換えベクターからの発現を正確に予想できる。このような標識タンパク質には、緑色蛍光タンパク質 (G F P ; C l o n t e c h) 、及び C D 6 4 または C D 6 4 - G F P 融合タンパク質が含まれる。レーザー光学に基づいた技術を利用した自動流動細胞計測法 (F C M) を用いて、G F P または C D 6 4 - G F P を発現する形質移入された細胞を同定し、その細胞のアポトーシス状態や他の細胞特性を評価する。また、F C M で、先行した或いは同時の細胞死の現象を診断する蛍光分子の取り込みを検出して計量する。これらの現象には、プロピジウムヨウ化物での D N A の染色によって計測される核 D N A 内容物の変化と、プロモデオキシウリジンの取り込み量の低下によって計測される D N A 合成の下方調節と、特異的な抗体との反応性によって計測される細胞表面及び細胞内のタンパク質の発現の変化と、蛍光複合アネキシン V タンパク質の細胞表面への結合によって計測される原形質膜組成の変化とが含まれる。流動細胞計測法は、O r m e r o d , M . G . による (1 9 9 4) F l o w C y t o m e t r y O x f o r d , N e w Y o r k , N Y . に記載されている。

20

30

40

【0263】

遺伝子発現における E T R N の影響は、E T R N をコードする配列と C D 6 4 または C D 6 4 - G F P 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。C D 6 4 または C D 6 4 - G F P 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 (I g G) の保存された領域と結合する。形質転換された細胞と形質転換されない細胞とは、ヒト I g G が C D 6 4 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 (D Y N A L . L a k e S u c c e s s . N Y)

50

。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。ETRN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。

【0264】

1.3 ETRNに特異的な抗体の産生

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ETRN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。

【0265】

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

【0266】

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

【0267】

1.4 特異的な抗体を用いる天然ETRNの精製

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

【0268】

ETRNを含む培養液をイムノアフィニティークラムに通し、ETRN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ETRN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塩イオンのようなカオトロピックイオンで)溶出させ、ETRNを回収する。

【0269】

1.5 ETRNと相互作用する分子の同定

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

【0270】

別法では、ETRNと相互作用する分子を、Fields, S. 及びO. Song (1

10

20

30

40

50

989, Nature 340:245-246)に記載の酵母2-ハイブリッドシステム(yeast two-hybrid system)やMATCHMAKERシステム(Clontech)などの2-ハイブリッドシステムに基づいた市販のキットを用いて分析する。

【0271】

ETRN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)。

10

【0272】

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

【0273】

(表の簡単な説明)

表1は、ETRNをコードする完全長の配列を作り出すために用いた、ポリペプチド配列及びヌクレオチド配列の配列番号(SEQ ID NO)、クローン識別番号、cDNAライブラリ、及びcDNA断片を示す。

20

【0274】

表2は、潜在モチーフ及び相同配列を含む各ポリペプチド配列の特徴、並びにETRNの解析に用いた方法、アルゴリズム、及び検索可能なデータベースを示す。

【0275】

表3は、各核酸配列の選択された断片と、ノーザン分析によって決定された各核酸配列の組織特異的発現パターンと、これらの組織に関連した疾患、異常症及び症状と、各DNAがクローニングされたベクターとを示す。

【0276】

表4は、ETRNをコードするcDNAクローンを単離したcDNAライブラリの作製に用いた組織を示す。

30

【0277】

表5は、本発明のポリヌクレオチド及びポリペプチドの分析に用いたツール、プログラム、及びアルゴリズム、並びにその説明、引用文献、閾値パラメーターを示す。

【0278】

【表1】

表1

ペプチド SEQ ID NO:	ヌクレオチド SEQ ID NO:	クローンID	ライブラリ	断片
1	6	1319812	BLADNOT04	904583R2 (COLNNOT07), 1319812F6 (BLADNOT04), 1319812HI (BLADNOT04), 1526227T1 (UCMCL5T01), 1559639H1 (SPLNNOT04), 1922786R6 (BRSTTUT01), 1968261R6 (BRSTNOT04), 3270967H1 (BRAINOT20), 3505563H1 (ADRENOT11)
2	7	2052013	LIVRFET02	1717630F6 (UCMNCOT02), 2052013H1 (LIVRFET02)
3	8	2693187	LUNGNOT23	1421691F6 (KIDNNOT09), 1540176RI (SINTTUT01), 1998452R6 (BRSTTUT03), 3292442F6 (BONRFET01)
4	9	4315138	BRAFNOT01	263103HI (HNT2AGT01), 1428228F6 (SINTBST01), 1813156F6 (PROSTUT12), 1994911R6 (BRSTTUT03), 2450737F6 (ENDANOT01), 2450737T6 (ENDANOT01), 2457623F6 (ENDANOT01), 2893095F6 (LUNGFET04), 2993723H1 (KIDNFET02), 3697155H1 (SININOT05), 3731896H1 (SMCCNON03), 4315138HI (BRAFNOT01), 4758279H1 (BRAMNOT01)
5	10	5135362	OVARDIT04	878706T1 (THYRNOT02), 1421057F1 (KIDNNOT09), 1456570F6 (COLNFET02), 3128634F6 (LUNGTUT12), 5135362H1 (OVARDIT04)

10

20

30

【 0 2 7 9 】
【 0 2 8 0 】
【 表 2 】

表2

ペプチド SEQ ID NO:	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	シグネチャ (signature) 配列	相同配列	分析方法及び データベース
1	263	S3 S182 T186 S222 T232 T255	N38 N185 N201	シグナルペプチド: M1-Q20 シグナルペプチド: L11-Y30 M43-Y59 L120-F139, F199-Y216 チトクロム b 561モチーフ: G17-Q238	g1651201 チトクロム b 561	Motifs SPSCAN HMME BLAST_PRODUM BLAST_DOMO BLAST_GenBank
2	141	S3 S38 T137		ABCトランスポーター モチーフ: L56-V70 グロビンシグネチャ: M1-R141	g4455876 α D-グロビン	Motifs HMME_PPFAM BLIMPS_PRINTS BLAST_GenBank
3	376	S7 T28 S30 S82 T167 S302 T357 T359 S68 S111 T219 T232 S327 T332 T335 T361 T366	N76 N203	ATP/GTP結合部位: G48-T56 ATP/GTP結合コバラミン シグネチャモチーフ: V73-S222	g3152705 コバラミン合成酵素様 胎盤タンパク質 (Cobalamin biosynthase-like placental protein)	Motifs BLAST_PRODUM BLAST_DOMO BLAST_GenBank
4	162	T92 T81		FADシグネチャモチーフ: R9-Q152	g840671 FADシグネチャモチーフ	Motifs BLAST_PRODUM BLAST_DOMO BLAST_GenBank
5	265	S108 S117 T81 S244		ユビキノリン 合成タンパク質 モチーフ: Y38-T255	g2240198 Coq4b ユビキノリン 合成タンパク質	Motifs BLAST_PRODUM BLAST_GenBank

10

20

30

【 0 2 8 1 】

【 表 3 】

表 3

ヌクレオチド SEQ ID NO:	選択断片	発現組織 (割合)	疾患または症状 (割合)	ペクター
6	489-533	生殖 (0.244) 造血/免疫 (0.178) 神経 (0.122)	癌 (0.489) 炎症/外傷 (0.289) 細胞増殖 (0.156)	PINCY
7	55-99	免疫 (0.636) 造血/免疫 (0.182) 神経 (0.091) 生殖 (0.091)	細胞増殖 (0.636) 癌 (0.273)	PINCY
8	569-613	生殖 (0.223) 神経 (0.175) 胃腸 (0.147)	癌 (0.422) 炎症/外傷 (0.347) 胎原 (0.194)	PINCY
9	325-369	神経 (0.228) 胃腸 (0.215) 心血管 (0.165) 生殖 (0.165)	癌 (0.430) 炎症/外傷 (0.354) 細胞増殖 (0.177)	PINCY
10	395-439	生殖 (0.315) 神経 (0.146) 胃腸 (0.135)	癌 (0.506) 炎症/外傷 (0.337) 細胞増殖 (0.157)	PINCY

【 0 2 8 2 】

【 表 4 】

表 4

ヌクレオチド SEQ ID NO	ライブラリ	ライブラリの説明
6	BLADNOT04	このライブラリは、銃自殺で死亡した28歳の白人男性から採取した膀胱組織を用いて作製した。
7	LIVRFET02	このライブラリは、妊娠 20 週目に死亡した白人胎児（女児）から採取した肝臓組織から単離した RNA を用いて作製した。
8	LUNGNOT23	このライブラリは、年齢 58 歳の白人男性の左肺葉から採取された肺組織から単離された RNA を用いて作製した。関連腫瘍組織は、転移性のグレード 3（4 段階中）の骨肉腫を示していた。患者の既往症には、軟組織癌、続発性肺癌、前立腺癌、出血を伴った急性十二指腸潰瘍があった。家族歴には、前立腺癌、乳癌、及び急性白血病があった。
9	BRAFNOT01	このライブラリは、35 歳白人男性の脳から除去した小脳（およびその近接領域）から単離した RNA を用いて作製した。
10	OVARDIT04	このライブラリは、22 歳白人女性の左卵巣切除の際に採取した左卵巣病変組織から単離した RNA を用いて作製した。病理学的には、線維組織及び毛髪の凝集からなる左卵巣の成熟嚢胞性奇形腫（類皮嚢胞）であった。患者の病歴には、脾臓膜の破壊、単核細胞症及びクバラミディア（21 歳で治療）があった。

10

20

30

【 0 2 8 3 】

【 表 5 】

【 0 2 8 4 】
【 表 6 】

プログラム名	説明	引用文献	パラメーター閾値
ABI FACTURA	核酸配列においてベクター配列を除去して不足の塩基をマスクするプログラム。	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finder は、アミノ酸または核酸配列の比較及び注釈付け (annotation) に有用である。	Perkin-Elmer 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。	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	Basic Local Alignment Search Tool は、アミノ酸及び核酸配列の配列類似性検索に有用であり、blastp 及び blastn、blastx、tblastn、tblastx の5つのファンクショナルがある。	Altschul, S.F.他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: 確率値=1.0E-8 以下 完全長配列: 確率値=1.0E-10 以下
FASTA	Pearson 及び Lipman アルゴリズムは、問合わせの配列と同種の配列群との類似性を検索する。FASTA は、fasta 及び tfasta、fastx、tfastx、ssearch の少なくとも5つのファンクショナルを含む。	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; Smith, T.F. and 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、一致長さ=200 塩基以上、fastx E 値=1.0E-8 以下 完全長配列: fastx スコア=100 以上
BLIMPS	Blocks IMPROVED Searcher は、BLOCKS 及び PRINTS、DOMO、PRODOM、PFAM データベースにおける配列に対して、ある配列の一致性を調べ、遺伝子ファミリー及び配列同一性、構造的フィンガープリント領域を探索する。	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:5665-72, 1991. J.C. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; Artwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	スコア=1000 以上 スコア/強度=0.75 以上 該当する場合、確率値=1.0E-3 以下
HMMER	PFAM などのタンパク質ファミリーコンセンサス配列の隠れマルコフモデル(HMM)に基づいたデータベースに対して問合せ配列を検索するアルゴリズム。	Krogh, A. 他 (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. 他 (1988) Nucleic Acids Res. 26:320-322.	スコア=10-50 ビット、各タンパク質ファミリーによって異なる。

10

20

30

40

【 0 2 8 5 】

プログラム名	説明	参考文献	パラメーター値
ProfileScan	Prosite で定義された配列パターンと一致するタンパク質配列における構造及び配列のモチーフを検索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4:61-66; Gribskov, 他 (1989) Methods Enzymol. 183:146-159; Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221	標準化された質のスコアを特定の Prosite モチーフに対する GCG 指定 "HIGH" 値 通常、スコア=1.4-2.1
Phred	高い感度及び確率で自動配列決定機のトレースを調べる塩基読み出しアルゴリズムである。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的なインプレメンテーションに基づく SWAT や CrossMatch を含む Phils Revised Assembly プログラムであって、配列相同性の検索及び DNA 配列の構築に有用である。	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; Green, P., University of Washington, Seattle, WA.	スコア=120 以上 一致長さ=56 以上
Consed	Phrap で構築したものの表示及び編集をするためのグラフィックツールである。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキヤンして分泌シグナルペプチドの存在を調べる重み付けマトリクス解析プログラムである。	Nielson, H. 他 (1997) Protein Engineering 10: 1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	スコア=3.5 以上
Motifs	Prosite で定義された配列と一致したパターンについてアミノ酸配列を検索するプログラムである。	Bairoch 他、前出; Wisconsin Package Program Manual, version 9, Page M51-59, 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
25 January 2001 (25.01.2001)

PCT

(10) International Publication Number
WO 01/05969 A2

(51) International Patent Classification: C12N 15/12,
5/10, C07K 1447, 16/18, C12Q 1/68, G01N 33/50, A61K
38/17, A01K 67/027

Sunnyvale, CA 94087 (US); BAUGHN, Mariah, R.
[US/US]; 14244 Santiago Road, San Leandro, CA 94577
(US); LU, Dyang, Aina, M. [US/US]; 53 Park Belmont
Place, San Jose, CA 95136 (US).

(21) International Application Number: PCT/US00/19036

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

(22) International Filing Date: 12 July 2000 (12.07.2000)

(25) Filing Language: English

(81) Designated States (national): AE, AL, AM, AT, AU, AZ,
BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, 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, MD,
MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD,
SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ,
VN, YU, ZA, ZW.

(26) Publication Language: English

(30) Priority Data: 60/143,816 14 July 1999 (14.07.1999) US

(63) Related by continuation (CON) or continuation-in-part
(CIP) to earlier application: 60/143,816 (CIP)
US Filed on 14 July 1999 (14.07.1999)

(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, SF), OAPI patent (BF, BI, CF, CG,
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).



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

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

(72) Inventors; and
(75) Inventors/Applicants (for US only): LAL, Preeti
[IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US);
TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose,
CA 95118 (US); YUE, Henry [US/US]; 826 Lois Avenue,

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

WO 01/05969 A2

(54) Title: ELECTRON TRANSFER PROTEINS

(57) Abstract: The invention provides human electron transfer proteins (ETRN) and polynucleotides which identify and encode ETRN. 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 expression of ETRN.

WO 01/05969

PCT/US00/19036

ELECTRON TRANSFER PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of electron transfer proteins and
5 to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative disorders,
reproductive disorders, and disorders of the immune response.

BACKGROUND OF THE INVENTION

Electron carriers such as cytochromes accept electrons from NADH or FADH₂ and donate
10 them to other electron carriers. Most electron-transferring proteins, except ubiquinone, are prosthetic
groups such as flavins, heme, FeS clusters, and copper, bound to inner membrane proteins.
Adrenodoxin, for example, is an FeS protein that forms a complex with NADPH:adrenodoxin
reductase and cytochrome P450. Cytochromes contain a heme prosthetic group consisting of a
porphyrin ring containing a tightly bound iron atom. Electron transfer reactions play a crucial role in
15 cellular energy production.

Energy is produced by the oxidation of glucose and fatty acids. Glucose is initially converted
to pyruvate in the cytoplasm. Fatty acids and pyruvate are transported to the mitochondria for
complete oxidation to CO₂ coupled by enzymes to the transport of electrons from NADH and FADH₂
to oxygen and to the synthesis of ATP (oxidative phosphorylation) from ADP and P_i.

20 Pyruvate is transported into the mitochondria and converted to acetyl-CoA for oxidation via
the citric acid cycle, involving pyruvate dehydrogenase components, dihydrolipoyl transacetylase, and
dihydrolipoyl dehydrogenase. Enzymes involved in the citric acid cycle include: citrate synthetase,
aconitases, isocitrate dehydrogenase, alpha-ketoglutarate dehydrogenase complex including
transsuccinylases, succinyl CoA synthetase, succinate dehydrogenase, fumarases, and malate
25 dehydrogenase. Acetyl CoA is oxidized to CO₂ with concomitant formation of NADH, FADH₂, and
GTP. In oxidative phosphorylation, the transfer of electrons from NADH and FADH₂ to oxygen by
dehydrogenases is coupled to the synthesis of ATP from ADP and P_i by the F₀F₁ ATPase complex in
the mitochondrial inner membrane. Enzyme complexes responsible for electron transport and ATP
synthesis include the F₀F₁ ATPase complex, ubiquinone(CoQ)-cytochrome c reductase, ubiquinone
30 reductase, cytochrome b, cytochrome c₁, FeS protein, and cytochrome c oxidase.

ATP synthesis requires membrane transport enzymes including the phosphate transporter and
the ATP-ADP antiport protein. The ATP-binding cassette (ABC) superfamily has also been suggested
as belonging to the mitochondrial transport group (Hogue, D.L. et al. (1999) J. Mol. Biol. 285:379-
389). Brown fat uncoupling protein dissipates oxidative energy as heat, and may be involved the fever
35 response to infection and trauma (Cannon, B. et al. (1998) Ann. NY Acad. Sci. 856:171-187).

WO 01/05969

PCT/US00/19036

The discovery of new electron transfer proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative disorders, reproductive disorders, and disorders of the immune response.

5

SUMMARY OF THE INVENTION

The invention features purified polypeptides, electron transfer proteins, referred to collectively as "ETRN" and individually as "ETRN-1," "ETRN-2," "ETRN-3," "ETRN-4," and "ETRN-5." In one aspect, the invention provides an isolated polypeptide comprising an amino acid
10 sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group
15 consisting of SEQ ID NO:1-5. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-5.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at
20 least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-5. In another alternative, the
25 polynucleotide is selected from the group consisting of SEQ ID NO:6-10.

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

WO 01/05969

PCT/US00/19036

The invention also provides a method for producing a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. 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 comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5.

The invention further provides an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions 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

WO 01/05969

PCT/US00/19036

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 comprising a polynucleotide
5 sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target
10 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 pharmaceutical composition comprising an effective amount of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an
15 amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and a pharmaceutically
20 acceptable excipient. In one embodiment, the pharmaceutical composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional ETRN, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention also provides a method for screening a compound for effectiveness as an
25 agonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic
30 fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. 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 pharmaceutical 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
35 condition associated with decreased expression of functional ETRN, comprising administering to a

WO 01/05969

PCT/US00/19036

patient in need of such treatment the pharmaceutical composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. 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 pharmaceutical 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 ETRN, comprising administering to a patient in need of such treatment the pharmaceutical composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-5. 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

WO 01/05969

PCT/US00/19036

change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:6-10, 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 comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, ii) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, ii) a naturally occurring polynucleotide sequence having at least 90% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of the above polynucleotide sequence; 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 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding ETRN.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of ETRN.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases.

WO 01/05969

PCT/US00/19036

disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding ETRN were isolated.

5 Table 5 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

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

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

20 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
25 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

"ETRN" refers to the amino acid sequences of substantially purified ETRN obtained from
30 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 ETRN. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other
35 compound or composition which modulates the activity of ETRN either by directly interacting with ETRN or by acting on components of the biological pathway in which ETRN participates.

WO 01/05069

PCT/US00/19036

An "allelic variant" is an alternative form of the gene encoding ETRN. 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 ETRN include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as ETRN or a polypeptide with at least one functional characteristic of ETRN. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding ETRN, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding ETRN. 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 ETRN. 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 ETRN 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 ETRN. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of ETRN either by directly interacting with ETRN or by acting on components of the biological pathway in which ETRN

WO 01/05969

PCT/US00/19036

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 ETRN polypeptides can be prepared using intact polypeptides or using
5 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
10 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
15 on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA:
20 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
25 introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic"
30 refers to the capability of the natural, recombinant, or synthetic ETRN, 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,
35 3'-TCA-5'.

WO 01/05969

PCT/US00/19036

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 ETRN or fragments of ETRN may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate, SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

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

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

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

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide

WO 01/05969

PCT/US00/19036

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

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

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide sequence 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.

10 A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

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

15 A "fragment" is a unique portion of ETRN or the polynucleotide encoding ETRN 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, 20 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 25 specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:6-10 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:6-10, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:6-10 is useful, for 30 example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:6-10 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:6-10 and the region of SEQ ID NO:6-10 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:1-5 is encoded by a fragment of SEQ ID NO:6-10. A fragment of 35 SEQ ID NO:1-5 comprises a region of unique amino acid sequence that specifically identifies SEQ

WO 01/05969

PCT/US00/19036

ID NO:1-5. For example, a fragment of SEQ ID NO:1-5 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-5. The precise length of a fragment of SEQ ID NO:1-5 and the region of SEQ ID NO:1-5 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the
5 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
10 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
15 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
20 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
30 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
35 programs are commonly used with gap and other parameters set to default settings. For example, to

WO 01/05969

PCT/US00/19036

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

5 *Penalty for mismatch: -2*

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

10 *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.

15 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 3.12c 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.

35 Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise

WO 01/05969

PCT/US00/19036

comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (Apr-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

5 *Open Gap: 11 and Extension Gap: 1 penalties*

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

10 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
15 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.

20 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
25 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
30 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.

35 Generally, stringency of hybridization is expressed, in part, with reference to the temperature

WO 01/05969

PCT/US00/19036

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_{81} or R_{41} 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 ETRN 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 ETRN which is useful in any of the antibody production methods disclosed herein or known in the art.

WO 01/05969

PCT/US00/19036

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.

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

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.

10 "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.

20 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 ETRN 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 ETRN.

"Probe" refers to nucleic acid sequences encoding ETRN, 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.

30 "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).

35 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

WO 01/05969

PCT/US00/19036

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.

5 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
10 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
15 100 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
20 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
25 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
30 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.

35 A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence

WO 01/05969

PCT/US00/19036

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
5 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
10 vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription,
15 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
20 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 nucleic
25 acids encoding ETRN, or fragments thereof, or ETRN itself, 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
30 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.

35 The term "substantially purified" refers to nucleic acid or amino acid sequences that are

WO 01/05969

PCT/US00/19036

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

WO 01/05969

PCT/US00/19036

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

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

THE INVENTION

The invention is based on the discovery of new human electron transfer proteins (ETRN), the polynucleotides encoding ETRN, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative disorders, reproductive disorders, and disorders of the immune response.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding ETRN. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each ETRN were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The Incyte clones in column 5 were used to assemble the consensus nucleotide sequence of each ETRN and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention.

WO 01/05969

PCT/US00/19036

column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding ETRN. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:6-10 and to distinguish between SEQ ID NO:6-10 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express ETRN as a fraction of total tissues expressing ETRN. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing ETRN as a fraction of total tissues expressing ETRN. Column 5 lists the vectors used to subclone each cDNA library.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding ETRN were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:6 maps to chromosome 1 within the interval from 143.1 to 146.6 centiMorgans. This interval also contains a RAS gene associated with cancer.

SEQ ID NO:9 maps to chromosome 4 within the interval from 51.1 to 60.0 centiMorgans, and to chromosome 9 within the interval from 65.0 to 78.4 centiMorgans. The interval on chromosome 4 within the interval from 51.1 to 60.0 centiMorgans also contains a gene associated with cell proliferation. The interval on chromosome 9 within the interval from 65.0 to 78.4 centiMorgans also contains a gene associated with an autoimmune disorder.

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

The invention also encompasses polynucleotides which encode ETRN. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:6-10, which encodes ETRN. The polynucleotide sequences of SEQ ID NO:6-10, as presented in the Sequence Listing, embrace the equivalent RNA sequences,

WQ 01/05969

PCT/US00/19036

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 ETRN. 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 ETRN. 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:6-10 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:6-10. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of ETRN.

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

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

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

WO 01/05969

PCT/US00/19036

hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:6-10 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."

5
10
15
20
Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (PE Biosystems, Foster City CA), 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 300 thermal cycler (PE Biosystems). Sequencing is then carried out using either the
15
20
25
30
35
ABI 373 or 377 DNA sequencing system (PE Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) *Short Protocols in Molecular Biology*, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) *Molecular Biology and Biotechnology*, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding ETRN 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
25
30
35
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

WO 01/05969

PCT/US00/19036

(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, PE Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode ETRN may be cloned in recombinant DNA molecules that direct expression of ETRN, 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 ETRN.

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

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat.

WO 01/05969

PCT/US00/19036

Biotechnol. 17:259-264; and Cramer, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of ETRN, 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
5 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
10 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 ETRN may be synthesized, in whole or in part,
15 using chemical methods well known in the art. (See, e.g., Carothers, 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, ETRN 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
20 York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (PE Biosystems). Additionally, the amino acid sequence of ETRN, 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
25 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 ETRN, the nucleotide sequences encoding ETRN or
30 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 ETRN. Such elements may vary in their strength and specificity. Specific initiation signals
35 may also be used to achieve more efficient translation of sequences encoding ETRN. Such signals

WO 01/05969

PCT/US00/10036

include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding ETRN 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,

5 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.)

10 Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding ETRN 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

15 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding ETRN. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with

20 yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., T1 or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, *supra*; Ausubel, *supra*; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544;

25 Scorer, C.A. et al. (1994) BioTechnology 12:181-184; 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; Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp.

30 191-196; 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.

35 USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al.

WO 01/65969

PCT/US00/19036

(1994) Mol. Immunol. 31(3):219-226; and Verma, J.M. and N. Soma (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 ETRN. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding ETRN 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 ETRN 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 ETRN are needed, e.g. for the production of antibodies, vectors which direct high level expression of ETRN may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of ETRN. 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, *supra*; and Scorer, *supra*.)

Plant systems may also be used for expression of ETRN. Transcription of sequences encoding ETRN may be driven 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, *supra*; Broglie, *supra*; and Winter, *supra*.) 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 ETRN 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 ETRN 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-

WO 01/05969

PCT/US00/19036

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

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

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding ETRN is inserted within a marker gene sequence, transformed cells containing sequences encoding ETRN can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding ETRN under the control of a single

WO 01/05969

PCT/US00/19036

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 ETRN and that express ETRN 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 ETRN 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 ETRN 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.F. 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 ETRN include oligelabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding ETRN, 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 ETRN 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 ETRN may be designed to contain signal sequences which direct secretion of ETRN through a prokaryotic or eukaryotic cell membrane.

WO 01/05969

PCT/US00/19036

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.

10 In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding ETRN 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 ETRN protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of ETRN 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 ETRN encoding sequence and the heterologous protein sequence, so that ETRN may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Austubel (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 ETRN may be achieved *in vitro* using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

ETRN of the present invention or fragments thereof may be used to screen for compounds that specifically bind to ETRN. At least one and up to a plurality of test compounds may be screened for specific binding to ETRN. Examples of test compounds include antibodies, oligonucleotides,

WO 01/05969

PCT/US00/19036

proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of ETRN, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, Coligan, J.F. et al. (1991) Current Protocols in Immunology (2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which ETRN 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 ETRN, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing ETRN or cell membrane fractions which contain ETRN are then contacted with a test compound and binding, stimulation, or inhibition of activity of either ETRN 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 ETRN, either in solution or affixed to a solid support, and detecting the binding of ETRN 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.

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

WO 01/05969

PCT/US00/19036

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 ETRN 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 ETRN 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 ETRN 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 ETRN, e.g., by secreting ETRN 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 ETRN and electron transfer proteins. In addition, the expression of ETRN is closely associated with reproductive and cell proliferative tissue. Therefore, ETRN appears to play a role in cell proliferative disorders, reproductive disorders, and disorders of the immune response. In the treatment of disorders associated with increased ETRN expression or activity, it is desirable to decrease the expression or activity of ETRN. In the treatment of disorders associated with decreased ETRN expression or activity, it is desirable to increase the expression or activity of ETRN.

Therefore, in one embodiment, ETRN or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or

WO 01/05969

PCT/US00/19036

activity of ETRN. Examples of such disorders include, but are not limited to, 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; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; and a disorder of the immune response 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.

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

In a further embodiment, a pharmaceutical composition comprising a substantially purified ETRN 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 ETRN including, but not limited to, those provided above.

WO 01/05969

PCT/US00/19036

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

5 In a further embodiment, an antagonist of ETRN may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of ETRN. Examples of such disorders include, but are not limited to, those cell proliferative disorders, reproductive disorders, and disorders of the immune response described above. In one aspect, an antibody which specifically binds ETRN 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 ETRN.

10 In an additional embodiment, a vector expressing the complement of the polynucleotide encoding ETRN may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of ETRN 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.

20 An antagonist of ETRN may be produced using methods which are generally known in the art. In particular, purified ETRN may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind ETRN. Antibodies to ETRN 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.

30 For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with ETRN 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 (*Bacillus Calmette-Guérin*) and *Corynebacterium parvum* are especially preferable.

35 It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to ETRN have an amino acid sequence consisting of at least about 5 amino acids, and generally will

WO 01/05969

PCT/US00/19036

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 ETRN amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

5 Monoclonal antibodies to ETRN 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
10 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,
15 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 ETRN-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.)

20 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 ETRN may also be generated.
25 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.)

30 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 ETRN and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies
35 reactive to two non-interfering ETRN epitopes is generally used, but a competitive binding assay may

WO 01/05969

PCT/US00/19036

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 ETRN. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of ETRN-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 ETRN epitopes, represents the average affinity, or avidity, of the antibodies for ETRN. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular ETRN 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 ETRN-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 ETRN, 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 ETRN-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 ETRN, 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 ETRN. 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 ETRN. (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.,

WO 01/05969

PCT/US00/19036

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 ETRN may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and Somia, N. (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 ETRN expression or regulation causes disease, the expression of ETRN 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 ETRN are treated by constructing mammalian expression vectors encoding ETRN and introducing these vectors by mechanical means into ETRN-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J.-L. and H. Récipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

WO 01/05969

PCT/US00/19036

Expression vectors that may be effective for the expression of ETRN include, but are not limited to, the PCIDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX 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). ETRN may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551); Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; invitrogen); the FK506/rapamycin inducible promoter, or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, supra), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding ETRN 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 ETRN expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding ETRN 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, J. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by

WO 01/05969

PCT/US00/19036

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; 5 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 ETRN to cells which have one or more genetic abnormalities with respect to the expression of ETRN. The construction and packaging of adenovirus-based vectors are well 10 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 Number 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) 15 *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 ETRN to target cells which have one or more genetic abnormalities with respect to the expression of ETRN. The use of herpes simplex virus (HSV)-based vectors may be 20 especially valuable for introducing ETRN 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. 25 Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 30 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 35 ordinary skill in the art.

WO 01/05969

PCT/US00/19036

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding ETRN 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. Biotech.* 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 ETRN into the alphavirus genome in place of the capsid-coding region results in the production of a large number of ETRN-coding RNAs and the synthesis of high levels of ETRN 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 ETRN 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.F. et al. (1994) in Huber, B.E. and B.J. Carr, *Molecular and Immunologic Approaches*, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

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

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,

WO 01/05969

PCT/US00/19036

5 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.

10 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 ETRN. 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.

15 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, 20 cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

25 An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding ETRN. 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 ETRN 30 expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding ETRN may be therapeutically useful, and in the treatment of disorders associated with decreased ETRN expression or activity, a compound which specifically promotes expression of the polynucleotide encoding ETRN may be therapeutically useful.

35 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

WO 01/05969

PCT/US00/19036

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

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

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

An additional embodiment of the invention relates to the administration of a pharmaceutical composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and

WO 01/05969

PCT/US00/19036

proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of *Remington's Pharmaceutical Sciences* (Mack Publishing, Easton PA). Such pharmaceutical compositions may consist of ETRN, antibodies to ETRN, and mimetics, agonists, antagonists, or inhibitors of ETRN.

5 The pharmaceutical 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.

10 Pharmaceutical 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
15 (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.

Pharmaceutical 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
20 determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of pharmaceutical compositions may be prepared for direct intracellular delivery of macromolecules comprising ETRN or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, ETRN or a fragment thereof may be joined to a short
25 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,
30 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 ETRN or fragments thereof, antibodies of ETRN, and agonists, antagonists or inhibitors of ETRN, which
35 ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by

WO 01/05969

PCT/US00/19036

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

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

DIAGNOSTICS

In another embodiment, antibodies which specifically bind ETRN may be used for the diagnosis of disorders characterized by expression of ETRN, or in assays to monitor patients being treated with ETRN or agonists, antagonists, or inhibitors of ETRN. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for ETRN include methods which utilize the antibody and a label to detect ETRN 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 ETRN, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of ETRN expression. Normal or standard values for ETRN expression are established by combining body fluids or cell extracts

WO 01/05969

PCT/US00/19036

taken from normal mammalian subjects, for example, human subjects, with antibody to ETRN under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of ETRN expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values.

5 Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding ETRN 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 ETRN may be correlated
10 with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of ETRN, and to monitor regulation of ETRN levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding ETRN or closely related molecules may be used to identify nucleic acid sequences which encode ETRN. The specificity of the probe, whether it is
15 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 ETRN, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50%
20 sequence identity to any of the ETRN 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:6-10 or from genomic sequences including promoters, enhancers, and introns of the ETRN gene.

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

30 Polynucleotide sequences encoding ETRN may be used for the diagnosis of disorders associated with expression of ETRN. Examples of such disorders include, but are not limited to, 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
35 adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in

WO 01/05969

PCT/US00/19036

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; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis, cancer of the breast, fibrocystic breast disease, and galactorrhea; a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; and a disorder of the immune response 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. The polynucleotide sequences encoding ETRN 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 ETRN expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding ETRN may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding ETRN 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 ETRN in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate

WO 01/05969

PCT/US00/19036

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 ETRN, 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 ETRN, 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 ETRN 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 ETRN, or a fragment of a polynucleotide complementary to the polynucleotide encoding ETRN, 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 ETRN 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 ETRN are used to

WO 01/05969

PCT/US00/19036

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

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described in Seilhamer, J.J. et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, incorporated herein by reference. 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, antibodies specific for ETRN, or ETRN or fragments thereof 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.

WO 01/05969

PCT/US00/19036

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

WO 01/05969

PCT/US00/19036

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
5 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
10 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
15 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
20 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
25 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 ETRN to quantify the levels of ETRN expression. In one embodiment, the antibodies are used as elements on a microarray,
30 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; Mendoz, 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
35 array element.

WO 01/05969

PCT/US00/19036

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 ETRN 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

WO 01/05969

PCT/US00/19036

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, e.g., Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

10 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 ETRN 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.

15 *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.

20 In another embodiment of the invention, ETRN, 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 ETRN and the agent being tested may be measured.

25 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 ETRN, or fragments thereof.

WO 01/05969

PCT/US00/19036

and washed. Bound ETRN is then detected by methods well known in the art. Purified ETRN 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.

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

10 In additional embodiments, the nucleotide sequences which encode ETRN 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 preferred specific 15 embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

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 preferred specific 20 embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/143,816, are hereby expressly incorporated by reference.

EXAMPLES

25 I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. 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 30 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 35 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

WO 01/05969

PCT/US00/19036

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
5 vector system (Stratagene) or SUPERSCRIP^T 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-
10 1000 bp) using S₁PHACRYL 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., PBLUESCRIP^T plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Genomics, Palo Alto CA). Recombinant
15 plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

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
20 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
30 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 (PE Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ
35 Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB

WO 01/05969

PCT/US00/19036

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 (PE Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (PE Biosystems) in conjunction with standard ABI protocols and base calling software, or other sequence analysis systems known in the art. Reading frames within the cDNA 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 VI.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 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, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the 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.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences 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 PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and 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 amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS,

WO 01/05969

PCT/US00/19036

DOMO, PRODOM, Prosite, 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, e.g., Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:6-10. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Analysis of Polynucleotide Expression

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

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

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

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: (the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences)). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding ETRN occurred. Analysis involved the categorization of cDNA libraries by

WO 01/05969

PCT/A/00/19036

organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Chromosomal Mapping of ETRN Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:6-10 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:6-10 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 5). 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.

The genetic map locations of SEQ ID NO:6 and SEQ ID NO:9 are described in The Invention as ranges, or intervals, of human chromosomes. More than one map location is reported for SEQ ID NO:9, indicating that previously mapped sequences having similarity, but not complete identity, to SEQ ID NO:9 were assembled into their respective clusters. 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. Diseases associated with the public and Incyte sequences located within the indicated intervals are also reported in the Invention where applicable.

VI. Extension of ETRN Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:6-10 were 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, 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

WO 01/05969

PCT/US00/19036

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.

5 High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg²⁺, (NH₄)₂SO₄, and β-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 5K+ 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.

15 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 mini-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 25 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 35 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;

WO 01/05969

PCT/US00/19036

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

In like manner, the polynucleotide sequences of SEQ ID NO:6-10 are used to obtain 5' regulatory sequences using the procedure above, along with oligonucleotides designed for such extension, and an appropriate genomic library.

10 VII. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:6-10 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of 15 [γ -³²P] 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 10⁷ counts per minute of the labeled probe is used in a typical membrane-based 20 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 25 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.

VIII. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing 30 photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schemm (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link 35 elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding

WO 01/05969

PCT/US00/19036

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.)

- 5 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.
- 10 After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

- Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMi.V 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 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.
- 20
- 25
- 30

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
- 35

WO 01/05969

PCT/US00/19036

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).

5 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.

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

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene).
15 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

20 Hybridization reactions contain 9 µl of sample mixture consisting of 0.2 µg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65 °C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the
25 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

30 Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas (0 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
35

WO 01/05969

PCT/US00/19036

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
5 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
10 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
15 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
20 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
25 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).

IX. Complementary Polynucleotides

Sequences complementary to the ETRN-encoding sequences, or any parts thereof, are used to
30 detect, decrease, or inhibit expression of naturally occurring ETRN. 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 ETRN. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence
35 and used to prevent promoter binding to the coding sequence. To inhibit translation, a

WO 01/05969

PCT/US00/19036

complementary oligonucleotide is designed to prevent ribosomal binding to the ETRN-encoding transcript.

X. Expression of ETRN

Expression and purification of ETRN is achieved using bacterial or virus-based expression systems. For expression of ETRN 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 ETRN upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of ETRN 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 ETRN 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, F.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, ETRN 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 ETRN at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunofluorescence 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 ETRN obtained by these methods can be used directly in the assays shown in Examples XI and XV.

XI. Demonstration of ETRN Activity

ETRN activity is demonstrated by oxidation or reduction of NADP. Substrates such as Asn-βGal, biocytidine, or ubiquinone-10 may be used. The reaction mixture contains 1-2 mg/ml HORP,

WO 01/05969

PCT/US00/19036

1.5 mM substrate, and 2.4 mM NAD(P)⁺ in 0.1 M phosphate buffer, pH 7.1 (oxidation reaction), or 2.0 mM NAD(P)H, in 0.1 M Na₂HPO₄ buffer, pH 7.4 (reduction reaction); in a total volume of 0.1 ml. FAD may be included with NAD, according to methods well known in the art. Changes in absorbance are measured using a recording spectrophotometer. The amount of NAD(P)H is stoichiometrically equivalent to the amount of substrate initially present, and the change in A₃₄₀ is a direct measure of the amount of NAD(P)H produced; AA₃₄₀ = 6620[NADH]. ETRN activity is proportional to the amount of NAD(P)H present in the assay. ETRN activity is proportional to the increase in extinction coefficient of NAD(P)H coenzyme at 340 nm, which is a measure of oxidation activity, or the decrease in extinction coefficient of NAD(P)H coenzyme at 340 nm, which is a measure of reduction activity (Dalziel, K. (1963) *J. Biol. Chem.* 238:2850-2858).

XII. Functional Assays

ETRN function is assessed by expressing the sequences encoding ETRN 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 plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), 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 ETRN on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding ETRN 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

WO 01/05969

PCT/US00/19036

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 ETRN and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIII. Production of ETRN Specific Antibodies

ETRN 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 ETRN 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 (PE Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-ETRN activity by, for example, binding the peptide or ETRN to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-labeled goat anti-rabbit IgG.

XIV. Purification of Naturally Occurring ETRN Using Specific Antibodies

Naturally occurring or recombinant ETRN is substantially purified by immunoaffinity chromatography using antibodies specific for ETRN. An immunoaffinity column is constructed by covalently coupling anti-ETRN 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 ETRN are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of ETRN (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/ETRN 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 ETRN is collected.

XV. Identification of Molecules Which Interact with ETRN

ETRN, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules

WO 01/05969

PCT/US00/19036

previously arrayed in the wells of a multi-well plate are incubated with the labeled ETRN, washed, and any wells with labeled ETRN complex are assayed. Data obtained using different concentrations of ETRN are used to calculate values for the number, affinity, and association of ETRN with the candidate molecules.

5 Alternatively, molecules interacting with ETRN are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, *Nature* 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

ETRN 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
10 between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

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
15 invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

WO 01/05969

PCT/US00/19036

Table 1

Peptide SEQ ID NO.	Nucleotide SEQ ID NO.	Clone ID	Library	Fragment
1	6	1319812	BLAN004	9045332 (COLN001), 1379812F6 (BLAN004), 1319812H1 (BLAN004), 1326227H1 (OCML001), 1549633H1 (SPIN004), 1922786E (BRST001), 1964261R6 (BRST004), 3270977H1 (BRA-NOT30), 3505563H1 (ADREN011)
2	7	2052013	LIVRE02	1717630F6 (UCMG002), 2052013H1 (LIVRE02)
3	6	2693187	LUNG023	1431691F6 (KIDN009), 1540176R1 (SINT001), 1994432E6 (BRST003), 3292442F6 (BOKRE01)
4	9	4315138	BRAN001	263103H1 (RNP2AG03), 1428238F6 (SINT001), 1813156F6 (PROST02), 1994911R6 (BRST003), 2450737F6 (BRAN001), 2450737F6 (ENDAN01), 2457623E6 (ENDAN01), 2892695F6 (LUNG-ET04), 2993723H1 (KIDNE02), 3697155H1 (SININ005), 3718966H1 (SHCC003), 4315138H1 (BRAN001), 5758279H1 (BRAN001)
5	10	5135362	OWARDIT04	878769T1 (CHYR002), 1421357F1 (KIDN009), 1456573F6 (COLN001), 3126634F6 (LUNG-UT12), 5135362H1 (OWARDIT04)

WO 01/05969

PCT/US00/19036

Table 2

Peptide SEQ ID NO.	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequence(s)	Homologous Sequence	Analytical Methods and Databases
1	262	S1 S182 T186 S222 S232 T255	N38 M185 N201	Signal peptide: V.- C28 Transmembrane motifs: L11-L10, M43-Y59, L12C-F13, F-59-V216 Cytochrome b561 motif: G17-Q219	G1651201 cytochrome b561	NCBI SPSCAN MEME BLAST_PRODB BLAST_DDBJ BLAST_GenBank
2	141	S3 S18 T137		ABC transporter motif: M1870 G186 G186 M1-R141	94455876 alpha D-globin	NCBI HMER_FFAN EMBL BLAST_GenBank
3	376	S7 T78 S30 S82 T167 S122 T357 T359 S68 S111 T219 T232 S127 T332 T335 T361 T366	N76 N203	ATP/GTP binding site: S49 T56 ATP-binding cobalamin synthetase motif: V73-S222	G152705 Cobalamin biosynthase-like placental protein	NCBI BLAST_PRODB BLAST_DDBJ BLAST_GenBank
4	162	T92 T81		FAD synthetase motif: R9-Q152	9843671 FAD synthetase	NCBI BLAST_PRODB BLAST_DDBJ BLAST_GenBank
5	265	S108 S117 T81 S244		Ubiquinone biosynthesis protein motif: Y38-T225	92240138 Coq4p ubiquinone biosynthesis protein	NCBI BLAST_PRODB BLAST_GenBank

WO 01/05969

PCT/US00/19836

Table 3

Nucleotide SEQ. ID. NO.	Selected Fragment	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
6	489-533	Reproductive (0.244) Hematopoietic/Immune (0.178) Nervous (0.122)	Cancer (0.489) Inflammation/Trauma (0.289) Cell proliferation (0.156)	pMVCY
7	55-99	Developmental (0.636) Hematopoietic/Immune (0.182) Nervous (0.091) Reproductive (0.087)	Cell proliferation (0.636) Cancer (0.273)	pMVCY
8	589-613	Reproductive (0.587) Nervous (0.171) Gastrointestinal (0.147)	Cancer (0.423) Inflammation/Trauma (0.347) Fever (0.394)	pMVCY
9	325-369	Nervous (0.228) Gastrointestinal (0.215) Cardiovascular (0.165) Reproductive (0.165)	Cancer (0.240) Inflammation/Trauma (0.354) Cell proliferation (0.177)	pMVCY
10	395-439	Reproductive (0.315) Nervous (0.146) Gastrointestinal (0.135)	Cancer (0.506) Inflammation/Trauma (0.337) Cell proliferation (0.157)	pMVCY

WO 01/05969

PCT/1/00/1936

Table 4

Nucleotide SEQ ID NO.	Library	Library Comment
6	BLADN0104	Library was constructed using RNA isolated from bladder tissue of a 28-year-old Caucasian male, who died from a self-inflicted gunshot wound.
7	LIVRE002	Library was constructed using RNA isolated from liver tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation.
8	LUNG0073	Library was constructed using RNA isolated from left lobe lung tissue removed from a 58-year-old Caucasian male. Pathology for the associated tumor tissue indicated metastatic grade 3 (of 4) osteosarcoma. Patient history included soft tissue cancer, secondary cancer of the lung, prostate cancer, and an acute duodenal ulcer with hemorrhage. Family history included prostate cancer, and acute leukemia.
9	BRAZ0070	Library was constructed using RNA isolated from amygdala (and adjacent area) tissue removed from the brain of a 38-year-old Caucasian male.
10	OVARI0104	This library was constructed using RNA isolated from diseased left ovary tissue removed from a 22-year-old Caucasian female during a left ovarian cystectomy. Pathology indicated a mature cystic teratoma (dermoid cyst) of the left ovary which consisted of aggregate pieces of fibrous tissue and hair. Patient history included capsular disruption of spleen, mononucleosis, and chlamydia (treated at age 21).

WO 01/05969

PCT/US00/19036

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequence and masks ambiguous bases in nucleic acid sequences.	PE Biosystems, Foster City, CA.	
ABI/PARACHE EDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	PE Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	PE 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: blast, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) <i>J. Mol. Biol.</i> 215:403-410; Altschul, S.F. et al. (1997) <i>Nucleic Acids Res.</i> 25:3389-902.	ESY: Probability value=1.0E-8 or less Full Length sequence: Probability value=1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, ifasta, fastx, ifastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci. USA</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183:63-98; and Smith, T.F. and M.S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489.	ESY: fast E value=1.0E-6 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequence: fastx score=1.00 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCks, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) <i>Nucleic Acids Res.</i> 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) <i>Methods Enzymol.</i> 266:38-105; and Atwood, T.K. et al. (1997) <i>J. Chem. Inf. Comput. Sci.</i> 37:417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and, if applicable, 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) <i>J. Mol. Biol.</i> 235:1501-1511; Sonnhammer, E.L.L. et al. (1988) <i>Nucleic Acids Res.</i> 26:320-322.	Score=10-50 bits for PFAM hits, depending on individual protein families

WO 01/05969

PCT/US00/19036

Table 5 (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; Birsoy, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores>CCG-specified "HG31" value for that particular Prosite motif. Generally, score=14.2, 1.
Phred	A base-calling algorithm that examines untrimmed sequence traces with high sensitivity and probability.	Evling, B. et al. (1998) Genome Res. 8:175-185; Tsing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phrap Revised Assembly Program including SWAY 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. 2:462-489; 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
Crossed	A graphical tool for viewing and editing Phrap assemblies.	Godton, D. et al. (1998) Genome Res. 8:195-202.	
SFScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:14; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Birsoy, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 01/05969

PCT/US00/19036

What is claimed is:

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

WO 01/05969

PCT/US00/19036

10. An isolated antibody which specifically binds to a polypeptide of claim 1.
11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:
- 5 a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10,
b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:6-10,
c) a polynucleotide sequence complementary to a),
d) a polynucleotide sequence complementary to b), and
10 e) an RNA equivalent of a)-d).
12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.
- 15 13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization
20 complex is formed between said probe and said target polynucleotide or fragments thereof, and
b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.
- 25 15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
30 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.
16. A pharmaceutical composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.
- 35

WO 01/05969

PCT/US00/19036

17. A pharmaceutical composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-5.

18. A method for treating a disease or condition associated with decreased expression of functional ETRN, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:
10 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
b) detecting agonist activity in the sample.

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

21. A method for treating a disease or condition associated with decreased expression of functional ETRN, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:
20 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
b) detecting antagonist activity in the sample.

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

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

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:
35 a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

WO 01/05969

PCT/US00/19036

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

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

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1.

b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound.

and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

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

a) exposing a sample comprising the target polynucleotide to a compound, and

b) detecting altered expression of the target polynucleotide.

20

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

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

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

c) quantifying the amount of hybridization complex; and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

WO 01/05969

PCT/US00/9036

SEQUENCE LISTING

<110> INCYTE GENOMICS, INC.
LAL, Preeti
TANG, Y. Tom
YUE, Henry
BAUGHN, Mariah R.
LI, Dyoung Aina M.

<120> ELECTRON TRANSFER PROTEINS

<130> PF-0723 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/143,816

<151> 1999-07-14

<160> 10

<170> PERL Program

<210> 1

<221> 263

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1119612CD1

<400> 1

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

WO 01/05969

PCT/US00/19036

Gly Arg Gln Pro Tyr Lys Ser Trp
 260

<210> 2
 <211> 141
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2952013CD1

<400> 2
 Met Leu Ser Ala Gln Glu Arg Ala Gln Ile Ala Gln Val Trp Asp
 1 5 10 15
 Leu Ile Ala Gly His Glu Ala Gln Phe Gly Ala Glu Leu Leu Leu
 20 25 30
 Arg Leu Phe Thr Val Tyr Pro Ser Thr Lys Val Tyr Phe Pro His
 35 40 45
 Leu Ser Ala Cys Gln Asp Ala Thr Gln Leu Leu Ser His Gly Gln
 50 55 60
 Arg Met Leu Ala Ala Val Gly Ala Ala Val Gln His Val Asp Asn
 65 70 75
 Leu Arg Ala Ala Leu Ser Pro Leu Ala Asp Leu His Ala Leu Val
 80 85 90
 Leu Arg Val Asp Pro Ala Asn Phe Pro Leu Leu Ile Gln Cys Phe
 95 100 105
 His Val Val Leu Ala Ser His Leu Gln Asp Glu Phe Thr Val Gln
 110 115 120
 Met Gln Ala Ala Trp Asp Lys Phe Leu Thr Gly Val Ala Val Val
 125 130 135
 Leu Thr Glu Lys Tyr Arg
 140

<210> 3
 <211> 376
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2693167CD1

<400> 3
 Met Leu Pro Ala Val Gly Ser Val Asp Glu Glu Glu Asp Pro Ala
 1 5 10 15
 Glu Glu Asp Cys Pro Glu Leu Val Pro Ile Glu Thr Thr Gln Ser
 20 25 30
 Glu Glu Glu Glu Lys Ser Gly Leu Gly Ala Lys Ile Pro Val Thr
 35 40 45
 Ile Ile Thr Gly Tyr Leu Gly Ala Gly Lys Thr Thr Leu Leu Asn
 50 55 60
 Tyr Ile Leu Thr Glu Gln His Ser Lys Arg Val Ala Val Ile Leu
 65 70 75
 Asn Glu Ser Gly Glu Gly Ser Ala Leu Glu Lys Ser Leu Ala Val
 80 85 90
 Ser Gln Gly Gly Glu Leu Tyr Glu Glu Tyr Leu Glu Leu Arg Asn
 95 100 105
 Gly Cys Leu Cys Cys Ser Val Lys Asp Asn Gly Leu Arg Ala Ile
 110 115 120
 Glu Asp Leu Met Gln Lys Lys Gly Lys Phe Asp Asp Ile Leu Leu
 125 130 135
 Glu Thr Thr Gly Leu Ala Asp Pro Gly Ala Val Thr Ser Met Phe
 140 145 150
 Trp Val Asp Ala Glu Leu Gly Ser Asp Ile Tyr Leu Asp Gly Ile
 155 160 165
 Ile Thr Ile Val Asp Ser Lys Tyr Gly Leu Lys His Leu Thr Glu
 170 175 180
 Glu Lys Pro Asp Gly Leu Ile Asn Glu Ala Thr Arg Gln Val Ala

WO 01/05969

PCT/US00/19036

```

185      190      195
Leu Ala Asp Ala Ile Leu Ile Asn Lys Thr Asp Leu Val Pro Glu
200      205      210
Glu Asp Val Lys Lys Leu Arg Thr Thr Ile Arg Ser Ile Asn Gly
215      220      225
Leu Gly Gln Ile Leu Gln Thr Gln Arg Ser Ser Leu Gln Lys Lys
230      235      240
Leu Gln His Val Pro Gly Thr Gln Pro His Leu Asp Gln Ser Ile
245      250      255
Val Thr Ile Thr Phe Glu Val Pro Gly Asn Ala Lys Gln Glu His
260      265      270
Leu Asn Met Phe Ile Gln Asn Leu Leu Trp Glu Lys Asn Val Arg
275      280      285
Asn Lys Asp Asn His Cys Met Glu Val Ile Arg Leu Lys Gly Leu
290      295      300
Val Ser Ile Lys Asp Lys Ser Gln Gln Val Ile Val Gln Gly Val
305      310      315
His Glu Leu Tyr Asp Leu Glu Gln Thr Pro Val Ser Trp Lys Asp
320      325      330
Asp Thr Glu Arg Thr Asn Arg Leu Val Leu Leu Gly Arg Asn Leu
335      340      345
Asp Lys Asp Ile Leu Lys Gln Leu Phe Ile Ala Thr Val Thr Glu
350      355      360
Thr Glu Lys Gln Trp Thr Thr Arg Phe Gln Glu Asp Gln Val Cys
365      370      375
Thr

```

```

<210> 4
<211> 162
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 4315138CD1

```

```

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

```

```

<210> 5
<211> 265
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 5135362CD1

```

WO 01/05969

PCT/US00/19036

```

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

```

```

<210> 6
<211> 2414
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1319812CB1

```

```

<400> 6
gctgagggcc ggctcgtagt gggaggcttc gcgttcaact gaacatttcc gccctcctctg 60
gatgggggat gagagggggg gctgatgcat cgcagcagtc ctggaccttg acttcggact 120
tggagaagaa agcccyagga ggaacgggag ctaggggggt agcagcttgg cgcgcgggga 180
gcgggcaccc cctggctcac ggaactcaat tccaggagag cctcgggcgc gactggaagt 240
gctggagacc tatzagaagg cggggcgcga ccgcgcgctc tgcctctggg eggcattgcc 300
agccggctgt agccattcag gccagctctt cctgcctctc ctagggaacc cgggagcgcc 360
agctccggcg cctggtagcg agaggcgggt tccggagatc ccggcctcac tctctccacc 420
tgtggttagg ggtgagtcct gcaaatgtta agtcatctgc tcaaggtgcc catttcgacg 480
gaattggagc ccagcccaat tctctgagcc tatcattagg gctaaaggag tgnctgatca 540
gaatgggttc tggacggctc tactttcctt gccctgctgt ggggtccctg gccctctatgt 600
gcatcctctt cctatctcac tggatgcagt actggtgttg tggctttgcc tggaaatggca 660
gcatctaat gtcaactgg caccagctgc ttatggttgc tggcctggcg gtattctatg 720
gaggctgctc actggtgtac cgcctgcctc agtctggggc ggggcccacg ctgctctgga 780
aactcttcca tgcagcctcg cactctatgg cctctctctt cactlllgtg gggctgggtg 840
ctgtctttac gtctcacaac catggagagg ctgcceacct ctactccctt cacagctggc 900
tggcctctac cactgtcttc cctctcgcct gccagtggtt cctggctttt gctgtcttcc 960
tcttgcctcg gccctccatc tggctgcgca gccctccaaa acctatccac gcttttttgg 1020
gagcggccat cctctctctg tccatgcgat ccgtcatttc gggcattaat gagaagcttt 1080
tctctagttt gaaaaacac accagggccat acccaagcct gcccaatgag ggggtcttgg 1140
ccaaacagcag cgggatgctg gtggtagcct ttgggctgct ggtctctcac acctctctgg 1200

```

WO 01/05969 PCT/US00/19036

cttcaatcttg gaagcccccga gggccggggg tccctaacga cagacagcty ctgclacagg 1260
fgggccctgg atccgggcel tcccccstga ctacgctctc tctcaacggc aggcagccct 1320
acaactccclg gtgacctgct ctcccagaa cagagcccty cccagatgt cccagtagcg 1380
atgagtaaca gaqgtqctg lggacttccr cnaactctcc ctgctggatc agggcctccc 1440
tgcctcccgz tggccaggic tggccctgct ctcttggcag gcccaccacc cctctgacca 1500
ctctgcaact caaccagcag ctgatgccc aattgtgggt tccagtgtg agcagccctg 1560
ggagcccttg caccctccag aggggttccr tggctgagac canattgctt cccctggccc 1620
caccctggct gcttgcctgg cccaacctg cgtctctgtg catctagag cltagactgt 1680
tgcctlcttl cagggcaaga aatagggtgg agagcgggaa ggtcttgcct cccaaagctt 1740
gctgctgtgg cttttttgce ttctccaaag accgactgcc aggtcccagd ctttagactg 1800
ctgctcttag taagcaagtg agagccctgg ggtttggagc ccacctacc tctggcagca 1860
tcagcatccl actcctggca acatcaggcc aagttccacc ccagctcac attgccagat 1920
gtttggagaa ggttaaatat tgaccgtctt gactggctgg agcctcaaaa gccactggga 1980
tgtctctcag gcaactgggt cccctgacca gctccccctc tccatagggg taagcattcc 2040
actggtttat gaagccctgag tttccatlaa tctgttaaga atcaagctg tctttgttca 2100
ggctgtctat ccaaaaatca aatagcctgg gtggcttaaa caacaagtgc tctctctcca 2160
cagttctgga gctctagac atccagctg aagggctgag caaalgtgat gttggcag 2220
ggccagcttc ctcaagacc abcttctccc caaacctca catggtggaa agggcctgag 2280
agctctctgg ggtccaccct tcaaacccc atcaactcct aaagcccca cgtctcaaca 2340
ctaccatagt ggggtgtag cttcaaccr aggaatttgg ggggacaaa acattcagtc 2400
tctaaataca gcca 2414

<210> 7
<211> 519
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2052013CB1

<400> 7
cggctagagg cggccgccat gctcagccc caggagcgc cccaaatcgn gaggctctgg 60
gacctgattg cgggcacaaga ggcgaattc gggcggagc tgcctctcag gctcttcccg 120
gtgtacccca gcaccaaagyl ctacttcccg caectggagc cctgcacaga cggagaccag 180
ctgctgagcc cgggcagcgy calgtggcgy gctgtggcgy cggcggctga ccaactggac 240
aacctgcacc cggcctctag cctgctggcy gactctgacc cgtctgctgt ggcctggacc 300
ccagccaact tccgctgct aatccagtgt ttcaccctg tgcctggccc ccaactgcag 360
gagaggttca ccgtccaaat gcaagcggcy tgggacaagt tccctgactgg tctggcctgg 420
gtgctgccc aaaaataccn cttagccctg tgcctgcccag gcttggctgt gtgctgttca 480
ataaacagag gcccgaaacc tctgaaaaaa aaaaaaaaa 519

<210> 8
<211> 1632
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2693187CB1

<400> 8
ggcagttctc gtccagagcc caggtaatcc gggcgggacc agctagcgtc gcgatgtgat 60
gagctcaggg cccggccagg cccgggctgg cgtcctggyc gtycccggct gcgtaagcc 120
gttttggctc cagcggctca gctgggctag ggaactgcty tagycccga aittaccggc 180
tgtttggatct gtgaatgagc aaaaagatcc tgggagagag gatgtctcgy aattgtctcc 240
cattagagcy accgaaagcy agggagagya aaagtctggc ctggcgcaca agatccagct 300
caaatatcac accgggtatt taggtgctgy gaagacaaca ctctcgaact atatttttag 360
agagcaaat agtaaaagag tagcggctcat tttaaatgaa cctggggagag gaagtgcgct 420
gggaaaatcc ttactgttca gccaaagctg agagctctat gaagatggc tgaacttag 480
aaacggtgc ctctgctgtt cagtgaagga caatggcctt agagctatg agaatttga 540
gcaaaaagag gggaaaattcy atgacatact gtttagagacc actggtattag cagaccctgg 600
tgcattgact tctatgtctt ggtttgactc tpaatctaggy agtgatattt acctgatgg 660
tatcataact atctgtgact caaatctgct atcaaacct ttaacagag agaacctga 720
tggccttacc aatgaagcta ctaggcagct tgccttagca gatgccatcc tcaattaaa 780
aacagaccct gttccagaag aagctgtaaa gaaatlaaga acgacaatta gatccataa 840
tggactagya caaatcttag aaacacaaag atcaagcttg cagaaaaaac ttcagctatg 900

WO 01/05969

PCT/US00/19036

```

gcaggaacaa caacctcacc ttgatccagag tattgttara atcacatttg aagtaaccagg 960
aaatgcacaag caaadaacacc tlaactatggt taltccagat ctccctgtagg aaaaagaatgt 1020
gagaacaacag gacacatcact gcatgagatg caaaagcag aagaqaltgg tgcacatcaa 1080
aacaaaataca caaccagtya ttgtccaggg tgcaccagag ctctatgaltc tcaagagcaac 1140
tcaagtacagc tgggaagatly acactgcagag aacaaatcga ttgtctctcc ttggcagaaa 1200
ttlagataaag gacatcctra aacagctglt ttagcttact tgcacagaaa caaaaaagca 1260
gtggacaaca cgtttccaag aagatcaagt ttgtacataa caatagagggc altcttacc 1320
Aaaaggattg gataaataaaa ataagttctt actgggtata tccaagcat ttatttacta 1380
ctctagttaa gaattccact aacttttaaa atggtatttg tttacagca tacataaagt 1440
gtagcaactc acctctgac aacattcaaa atccatcacg tctatataa taccctcttt 1500
tttaagaat ggtatttccac aaaaatact ttgaattg gctttggag gttacaata 1560
tactgacca ttgaagttt AATAataatt ggtatgacc acattccac catttggoot 1620
tttttccct ag 1632

```

```

<210> 9
<211> 2615
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature
<222> Incyte ID No: 4315138C81

```

```

<400> 9
acgggcaactl atttgggctg tctccgctgc cctccagcgg aghagcagct cctcaggcgc 60
cdggggatg ttgttgggtc ggagggctca ggaalccga gtggccagt gtccggggccg 120
tccgctcact gacgctcact cggctcaggg agagcctagt ctccgtctcc cggctcaggt 180
cggggtctgg gcccgcctgc gacgycgggt ccgaacagag ccgagaccgg gccgcagccc 240
atgcccagag aggarctcat tatgagccac ctgccttact tctcccggg tcaagtgggt 300
cggggcttcc gcccggctcc caagcagctg ggcctcccca cagctaatit lccgagcaa 360
gtgttagata atcttccagc tgalatatcc actggtattt actatggttg ggcagttgt 420
ggaatcgagc atgtccataa gatggtgggt agcatagcat ggaaccata tcaacaagaat 480
acgaagaat ctatgaaaac acatatactg cataccctca aagggactt ctatggggaa 540
acctcaatg tggcactgt tggctaccg agaccagaaa aacacttga tctcttagag 600
tcaattattc cagcaatbca agttagattt gaagaagctc agaacgact agagttaca 660
gaacatttga aaatcaaaa agacaatttc tccagtttt ctaaaagca aataacgaa 720
ggocactgat gaaaaattgt attatttatt catccactgt tttctagtt ttctgtgttt 780
attactgttc agcctactct tggttatata ctgaaaatca aacctttac lacagttgt 840
aattagttta aacctacaa tttgtttatc atataatgct tcaactatta aattaagccc 900
atcatatcat agacttaaaa agattgagtt aaaaaaaga ttaaaaaat atattattgat 960
taaaatgac cactgaaaag ggtttacatg tctataagta tgggttaaaa ggcaggtcac 1020
tagtttagga tgaaaactaa aattttctag ttaaaaagcc agcatattg taacttagag 1080
gtgtgtactt ctaaaacagag aaggtctata aaagttaacg cattaagccc aagttagtt 1140
gtccacacgg gtaatcccag cactttggga ggccagaggt ggaggattgc ttagctcat 1200
gagttcaaga ccagtttggg caacatagcg agactccatc tctataaata aaaaaaagg 1260
aaatgacata aatagaaggt tcatagttta tcaataactc cgggggaaa aactcagact 1320
ttctatttt gatattttgt gcaataatac attaaaaa tgaagattgt tgtaaagtt 1380
tgttttagtc attttctttg cagtcacatc attgtaccac ctccatata: gtatagagct 1440
gttcaatic ttagctctct ggactccccc tgcctatctg ttgttgttg ttgttgaaga 1500
cagtaaaaata agtgtttbca aatctatag ttctcttttt gttgagact gtttctctt 1560
ttaaatgta aatattcttg ttcttcaaca tttttctttg gttcttttt ctttttttag 1620
gaaaaaacia acaacagact tcaactctag gttttctcaa gatttaagt aacacatta 1680
cacatataca ttctctaaag aacacaaaat gtttctctcc tagcaaaact atttaagagc 1740
caattaactc atgaaatagt tataaaatta ttttaatat ttaaggcctt gaaatata 1800
tccaaaggcaa atcaaatgg gtytauttgt atatgcacga tgaatttaa tgcataaagag 1860
tacttacagc ttgagggaaa agcctaaagt tttagttaaa tagcttcat actgtggatt 1920
tttctctca ctactaata tctggcaat taagttaca gatgcttgg gcccctca 1980
acagttactt gctacagttc actgtgcaa cagatttgat ggtgaggg ggaataggt 2040
agaagcagat agaaggtat tttaaagtt atttattatg aatcatttc cctcaactg 2100
cataaatgaa tagaactga ttttaataaa aatattctgt gtaaaaagta ctcaaaaa 2160
carctaaagc ttttggagaa gagggaacta tataaalagt tggattgoyt ttgtctttt 2220
atttcttat aaaaatagt caaaaacagt ttttaatttt aqaatctta tactgaaag 2280
aaactcaat taccgtggc tttatgaagt aaagttaaaa acaaacatag gaaatgtgaa 2340
aacatttgtg ttgtttttgt ctgactaac tgcctgtaaa tgcctatctt aggtgttga 2400
aagttcttg atatgtact gtcctttcac ctctgttaa atgcccaly cttttgaaa 2460
tcaatagtg tttgttttag ggttaattg tgtcaatc tttcaaac ttttataca 2520
tgtattataa gaagttacag tagtaccact ttgattagtc tgaatgtttl caggctctga 2580

```

WO 01/05969

PCT/US00/19036

```

aatuaatttt gactaaq1ga aaaacaaaaa aaaa          2615
<210> 10
<211> 1645
<212> DNA
<213> Homo sapiens
<220>
<221> misc feature
<223> InseyTe ID No: 51553620B1
<400> 10
accgcctgga cctggacgga agtacccgca ggcgccgccc ccaaatgtgg tcccttccac 60
ggcccccggc atgttccaca gccggaagag gttcgcattt tatagtcttc ggggaaaaac 120
ggctgtggag aaggaaatag gcccgggggc cgaagcagcc tggctccgtg tcccttccag 180
aacctctctg gggcaggttg acatggcggag agctcaggat cggctggacgt agacggcaga 240
cagttccact ggttttccct cgcctacttg gcttaactgc ctctcggctg tgaagcgatg 300
ttccacctcg aaagcccgta gylacgcctg cagaa1cgtt lttlcag1ga gttttgacc 360
ctccagcctt cggcgcctcg acagaactgc ggccttcttc gtcaccgccc atccctccgg 420
gaccctccct gccatggcga ctctcgtcgc cctcgtcttc cgtcggcttt ggggctccc 480
ggycctacag cggcctcgcg cagaatcgc cctccggcct agagcgcagc ggcgccgccc 540
gctatactcg caccaccttc ccaactcccc gctgcagaaa gctgtgttgg ccgcgggctc 600
cgggggagtg ggcctctata accctaccg ccacgacatg gctcagcttc taggggagac 660
cbagggacac cgcacctgga aggtctctag ggcacagatg agggggatc cagaggg1gc 720
ccagatcccg caggagcctc cccggatttc gacatccacc ctcgacctgg gcaagctcca 780
gagcctcctg gaagcctccc tcggtcgcga gtatctccct tctctggalg tgaacaggtt 840
ctcccagac accgagcacc ccaacctgct cgtggatgat gaggagctag cgtatgcat 900
tcagcgtac cgggaggtgc acgacatgct tcacacctg cggggatgc ccaaccaact 960
cctgggggag atcgttg1ga aatggtttga gctctccag actggcctgc ccatglcal 1020
cctgggtgca tcttttggac cgatcngact tggcctcag agcctcaag tctgtgtctc 1080
ggagtgtac ccatgggctg tcagaacgg gcgcagacc ccatgtctcc tcaacctgta 1140
ctatggcgg cgtcgggac agtccctgag ggtctcgggg gaggacctgg gcaattacag 1200
accctccatz cagctccagg gcttggcctg agtctctgag cccggggggc ctggcctacc 1260
tcccacatc cctgcttccc ttggagcgag agggctccct tgcctacct tctctctctt 1320
ccttgaacac tgcctctgg acaacatta tcataattg tcataccac tctcagttgg 1380
ccttgaagac gaacctcgca gggagaagc agtaccgtg cartccaggg gggaccagca 1440
gctaccacag gaggaccrat catgaacagt atcaatcgtc tggctctatg ctgggatctc 1500
gcagtgtctc tcttgaacct cctcccagcc agccaggttt cctgggggccc aggtctggctc 1560
tctcaccagg agcagggct acaccaatt ccaaaagcct gagaagagag aagctggagg 1620
ggggcctggt gttgaaata agcgtccca 1645

```

【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
25 January 2001 (25.01.2001)

PCT

(10) International Publication Number
WO 01/05969 A3

- (51) International Patent Classification: C12N 15/12, S10, C07K 1447, 16/18, C12Q 1/68, G01N 33/50, A61K 38/17, A01K 67/027 [US/US]: 14244 Santiago Road, San Leandro, CA 94577 (US), LLI, Dnyong, Aims, M. [US/US]: 55 Park Belmont Place, San Jose, CA 95136 (US).
- (21) International Application Number: PCT/US00/19036 (74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).
- (22) International Filing Date: 12 July 2000 (12.07.2000) (81) Designated States (national): AE, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GR, HU, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TH, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (25) Filing Language: English
- (26) Publication Language: English
- (36) Priority Data: 60/143,816 14 July 1999 (14.07.1999) US
- (63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application: US 60/143,816 (CIP) Filed on 14 July 1999 (14.07.1999) (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), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).
- (71) Applicant (for all designated States except US): INCYTE GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US). Published: -- with international search report
- (72) Inventor; and (75) Inventors/Applicants (for US only): LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). TANG, V, Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). BAUGHN, Mariah, R. (88) Date of publication of the international search report: 19 July 2001
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



WO 01/05969 A3

(54) Title: ELECTRON TRANSFER PROTEINS

(57) Abstract: The invention provides human electron transfer proteins (ETRN) and polynucleotides which identify and encode ETRN. 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 expression of ETRN.

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 00/19036
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C12N5/10 C07K14/47 C07K16/18 C12Q1/68 G01N33/50 A61K36/17 A01K67/027		
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 G01N A61K A01K		
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)		
EPD-Internal, EMBL, BIOSIS		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 54963 A (FERRIE ANN H; HUMAN GENOME SCIENCES INC ; GREENE JOHN M (US); YOUNG) 10 December 1998 (1998-12-10) abstract seq. 116 (pages 369-370): compare nt 1069-1965 with nt 1245-2141 of seq. 10 6. claims 1-23	1-19, 22, 25-28
A	SRIVASTAVA M.: "Genomic structure and expression of the human gene encoding cytochrome b561, and integral protein of the chromaffin granule membrane." J. BIOL. CHEM., vol. 270, no. 39, 1995, pages 22714-22720, XP062150821 the whole document	1-19, 22, 25-28
<input type="checkbox"/> Further documents are filed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "C" 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 treatment (see, specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but after the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the applicant but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family		
Date of the actual completion of the international search		Date of the date of the international search report
23 October 2000		17. 01. 01
Name and mailing address of the ISA European Patent Office, P.O. Box 1818, Lankastrian 2 NL-2200 LZ Rijswijk Tel: (+31-70) 340 2940, Tx: 3-651 cuc nl, Fax: (+31-70) 340 5216		Authorized officer Galli, I

Form PCT/ISA 210 (2000) (English) July 1992

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 00/19036
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.	<input checked="" type="checkbox"/> Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely: Although claim 18 is 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.	<input checked="" type="checkbox"/> Claims Nos.: 20, 21, 23, 24 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: see FURTHER INFORMATION sheet PCT/ISA/210	
3.	<input type="checkbox"/> 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.	<input type="checkbox"/> As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.	
2.	<input type="checkbox"/> As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.	
3.	<input type="checkbox"/> 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.	<input checked="" type="checkbox"/> 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.: 1-28 partly	
Remark on Protest		
		<input type="checkbox"/> The additional search fees were accompanied by the applicant's protest.
		<input type="checkbox"/> No protest accompanied the payment of additional search fees.

International Application No. PCT/US 00/19036

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

1. Claims: 1-28, partly

An isolated polypeptide comprising the amino acid sequence of seq. ID 1 or a fragment thereof.

A corresponding nucleic acid (seq. ID 6), vectors, recombinant host cells, antibodies, transgenic animals, methods of screening for ligands and modulators, pharmaceutical compositions and therapeutical applications.

2. Claims: 1-28, partly

Idem as subject-matter 1, but limited to seq. IDs 2 and 7.

3. Claims: 1-28, partly

Idem as subject-matter 1, but limited to seq. IDs 3 and 8.

4. Claims: 1-28, partly

Idem as subject-matter 1, but limited to seq. IDs 4 and 9.

5. Claims: 1-28, partly

Idem as subject-matter 1, but limited to seq. IDs 5 and 10.

International Application No. PCT/US 00/19036

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 20,21,23,24

Claims 20,21,23,24 refer to agonists and antagonists of the electron transfer protein claimed, without however providing a true technical characterisation. Moreover, no such compounds are defined in the application. In consequence, said claims are ambiguous and vague, and their subject matter is not sufficiently disclosed and supported pursuant to Art. 5 and 6 PCT. No search can be carried out for such purely speculative claims, the wording of which is, in fact, a mere recitation of the results to be achieved.

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

INTERNATIONAL SEARCH REPORT
Information on patent family members

Int. Application No.
PCT/US 00/19036

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WD 9854963 A	10-12-1998	AU 7812098 A	21-12-1998
		EP 1838801 A	04-10-2000
		AU 7811998 A	21-12-1998
		AU 6552198 A	29-09-1998
		EP 0973882 A	26-01-2000
		WO 9840483 A	17-09-1998

フロントページの続き

(51) Int.Cl. ⁷	F I	テーマコード(参考)
A 6 1 P 3/10	A 6 1 P 5/38	4 C 0 8 4
A 6 1 P 5/38	A 6 1 P 7/00	4 H 0 4 5
A 6 1 P 7/00	A 6 1 P 7/06	
A 6 1 P 7/06	A 6 1 P 9/04	
A 6 1 P 9/04	A 6 1 P 9/10	
A 6 1 P 9/10	A 6 1 P 11/06	
A 6 1 P 11/06	A 6 1 P 13/12	
A 6 1 P 13/12	A 6 1 P 15/00	
A 6 1 P 15/00	A 6 1 P 15/08	
A 6 1 P 15/08	A 6 1 P 15/10	
A 6 1 P 15/10	A 6 1 P 19/02	
A 6 1 P 19/02	A 6 1 P 19/10	
A 6 1 P 19/10	A 6 1 P 21/04	
A 6 1 P 21/04	A 6 1 P 25/00	
A 6 1 P 25/00	A 6 1 P 25/04	
A 6 1 P 25/04	A 6 1 P 29/00	1 0 1
A 6 1 P 29/00	A 6 1 P 31/10	
A 6 1 P 31/10	A 6 1 P 31/12	
A 6 1 P 31/12	A 6 1 P 31/18	
A 6 1 P 31/18	A 6 1 P 33/00	
A 6 1 P 33/00	A 6 1 P 35/00	
A 6 1 P 35/00	A 6 1 P 35/02	
A 6 1 P 35/02	A 6 1 P 37/08	
A 6 1 P 37/08	A 6 1 P 43/00	
A 6 1 P 43/00	C 0 7 K 14/47	
C 0 7 K 14/47	C 0 7 K 16/18	
C 0 7 K 16/18	C 1 2 N 1/15	
C 1 2 N 1/15	C 1 2 N 1/19	
C 1 2 N 1/19	C 1 2 N 1/21	
C 1 2 N 1/21	C 1 2 P 21/02	C
C 1 2 N 5/10	C 1 2 Q 1/68	A
C 1 2 P 21/02	G 0 1 N 33/15	Z
C 1 2 Q 1/68	G 0 1 N 33/50	Z
G 0 1 N 33/15	G 0 1 N 33/53	M
G 0 1 N 33/50	G 0 1 N 33/566	
G 0 1 N 33/53	G 0 1 N 37/00	1 0 2
G 0 1 N 33/566	C 1 2 N 5/00	A
G 0 1 N 37/00	A 6 1 K 37/02	

(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), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG), AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, 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, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, U A, UG, US, UZ, VN, YU, ZA, ZW

(72) 発明者 タング、トム・ワイ

アメリカ合衆国カリフォルニア州 9 5 1 1 8 ・ サンノゼ ・ ランウィックコート 4 2 3 0

- (72)発明者 ユエ、ヘンリー
 アメリカ合衆国カリフォルニア州 9 4 0 8 7 ・ サニーベイル・ルイスアベニュー 8 2 6
- (72)発明者 ボーゲン、マライア・アール
 アメリカ合衆国カリフォルニア州 9 4 5 7 7 ・ サンレアンドロ・サンティアゴロード 1 4 2 4 4
- (72)発明者 リュ、デュング・アイナ・エム
 アメリカ合衆国カリフォルニア州 9 5 1 3 6 ・ サンノゼ・パークベルモントブレイス 5 5

F ターム(参考) 2G045 AA25 AA40 CB01 DA12 DA13 DA14 DA36 DA77 FB02 FB03
 FB04 FB06 FB07 HA16
 4B024 AA01 AA11 BA43 BA80 CA04 CA07 CA09 CA11 DA01 DA02
 DA05 DA11 EA01 EA02 EA03 EA04 FA02 GA11 HA01 HA12
 4B063 QA01 QA08 QA18 QA19 QQ01 QQ42 QQ52 QR08 QR33 QR42
 QR55 QR59 QR62 QR82 QS25 QS34 QS36 QX02
 4B064 AG01 CA01 CA19 CC24 DA01 DA13
 4B065 AA01X AA57X AA87X AA90Y AB01 BA01 CA24 CA25 CA44 CA46
 4C084 AA17 ZA02 ZA36 ZA45 ZA51 ZA59 ZA68 ZA75 ZA81 ZA89
 ZA94 ZA97 ZB02 ZB13 ZB15 ZB26 ZB27 ZB31 ZB33 ZB35
 ZB38 ZC35 ZC42
 4H045 AA10 AA11 AA20 AA30 BA10 BA41 CA40 DA00 DA76 EA20
 EA50 FA72 FA74

专利名称(译)	电子传递蛋白		
公开(公告)号	JP2004500032A	公开(公告)日	2004-01-08
申请号	JP2001511182	申请日	2000-07-12
[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	ラルプリーティ タングトムワイ ユエヘンリー ボーグンマライアール リュデュングアイナエム		
发明人	ラル、プリーティ タング、トム・ワイ ユエ、ヘンリー ボーグン、マライア・アール リュ、デュング・アイナ・エム		
IPC分类号	G01N33/50 A61K38/00 A61K45/00 A61P1/04 A61P1/16 A61P3/10 A61P5/38 A61P7/00 A61P7/06 A61P9/04 A61P9/10 A61P11/06 A61P13/12 A61P15/00 A61P15/08 A61P15/10 A61P19/02 A61P19/10 A61P21/04 A61P25/00 A61P25/04 A61P29/00 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37/08 A61P43/00 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/09 C12N15/12 C12P21/02 C12Q1/68 G01N33/15 G01N33/53 G01N33/566 G01N37/00		
CPC分类号	A01K2217/05 A61K38/00 A61P1/04 A61P1/16 A61P11/06 A61P13/12 A61P15/00 A61P15/08 A61P15/10 A61P19/02 A61P19/10 A61P21/04 A61P25/00 A61P25/04 A61P29/00 C07K14/47		
FI分类号	C12N15/00.ZNA.A A61K45/00 A61P1/04 A61P1/16 A61P3/10 A61P5/38 A61P7/00 A61P7/06 A61P9/04 A61P9/10 A61P11/06 A61P13/12 A61P15/00 A61P15/08 A61P15/10 A61P19/02 A61P19/10 A61P21/04 A61P25/00 A61P25/04 A61P29/00.101 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P35/00 A61P35/02 A61P37/08 A61P43/00 C07K14/47 C07K16/18 C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.M G01N33/566 G01N37/00.102 C12N5/00.A A61K37/02		
F-TERM分类号	2G045/AA25 2G045/AA40 2G045/CB01 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/DA77 2G045/FB02 2G045/FB03 2G045/FB04 2G045/FB06 2G045/FB07 2G045/HA16 4B024/AA01 4B024/AA11 4B024/BA43 4B024/BA80 4B024/CA04 4B024/CA07 4B024/CA09 4B024/CA11 4B024/DA01 4B024/DA02 4B024/DA05 4B024/DA11 4B024/EA01 4B024/EA02 4B024/EA03 4B024/EA04 4B024/FA02 4B024/GA11 4B024/HA01 4B024/HA12 4B063/QA01 4B063/QA08 4B063/QA18 4B063/QA19 4B063/QQ01 4B063/QQ42 4B063/QQ52 4B063/QR08 4B063/QR33 4B063/QR42 4B063/QR55 4B063/QR59 4B063/QR62 4B063/QR82 4B063/QS25 4B063/QS34 4B063/QS36 4B063/QX02 4B064/AG01 4B064/CA01 4B064/CA19 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA01X 4B065/AA57X 4B065/AA87X 4B065/AA90Y 4B065/AB01 4B065/BA01 4B065/CA24 4B065/CA25 4B065/CA44 4B065/CA46 4C084/AA17 4C084/ZA02 4C084/ZA36 4C084/ZA45 4C084/ZA51 4C084/ZA59 4C084/ZA68 4C084/ZA75 4C084/ZA81 4C084/ZA89 4C084/ZA94 4C084/ZA97 4C084/ZB02 4C084/ZB13 4C084/ZB15 4C084/ZB26 4C084/ZB27 4C084/ZB31 4C084/ZB33 4C084/ZB35 4C084/ZB38 4C084/ZC35 4C084/ZC42 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/BA41 4H045/CA40 4H045/DA00 4H045/DA76 4H045/EA20 4H045/EA50 4H045/FA72 4H045/FA74		
优先权	60/143816 1999-07-14 US		
外部链接	Espacenet		

摘要(译)

本发明提供了鉴定和编码ETRN的人电子转移蛋白 (ETRN) 和多核苷酸。 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ETRN表达有关的疾病的方法。

核苷酸 位置	核苷酸 位置	核苷酸 位置	核苷酸 位置	核苷酸 位置
1	11942	12074	12106	12138
2	26201	26333	26465	26597
3	26818	26950	27082	27214
4	41519	41651	41783	41915
5	51530	51662	51794	51926