

(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2005-500818

(P2005-500818A)

(43) 公表日 平成17年1月13日(2005.1.13)

(51) Int. Cl. ⁷		F I		テーマコード (参考)	
C 1 2 N	15/09	C 1 2 N	15/00	Z N A A	2 G O 4 5
A 6 1 K	38/00	A 6 1 K	45/00		4 B O 2 4
A 6 1 K	45/00	A 6 1 P	1/04		4 B O 2 9
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	1/18		4 B O 6 4
		審査請求	未請求	予備審査請求	有 (全 271 頁) 最終頁に続く

(21) 出願番号	特願2002-565105 (P2002-565105)	(71) 出願人	301005050
(86) (22) 出願日	平成14年1月4日 (2002.1.4)		インサイト・ゲノミックス・インコーポレイテッド
(85) 翻訳文提出日	平成15年6月30日 (2003.6.30)		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/US2002/000254		
(87) 国際公開番号	W02002/064792	(74) 代理人	100089266
(87) 国際公開日	平成14年8月22日 (2002.8.22)		弁理士 大島 陽一
(31) 優先権主張番号	60/260, 168	(72) 発明者	ボーグン、マライア・アール
(32) 優先日	平成13年1月5日 (2001.1.5)		アメリカ合衆国カリフォルニア州94577・サンレアンドロ・サンティアゴロード 14244
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/262, 857		
(32) 優先日	平成13年1月19日 (2001.1.19)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/262, 736		
(32) 優先日	平成13年1月19日 (2001.1.19)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 疾患検出および治療用分子

(57) 【要約】

本発明は、ヒトの疾患検出および治療用分子 (M D D T) および、M D D Tを同定しコードするポリヌクレオチドを提供する。本発明はまた、発現ベクター、宿主細胞、抗体、アゴニストおよびアンタゴニストをも提供する。本発明はまた、M D D Tの異常発現に関連する障害を診断、治療、または予防する方法をも提供する。

【特許請求の範囲】

【請求項 1】

以下の (a) 乃至 (d) からなる群から選択した単離されたポリペプチド。

(a) SEQ ID NO: 1-26 (配列番号 1 乃至 26) からなる群から選択した或るアミノ酸配列を持つポリペプチド

(b) SEQ ID NO: 1-26 からなる群から選択したアミノ酸配列に対して少なくとも 90% が同一であるような天然アミノ酸配列を持つポリペプチド

(c) SEQ ID NO: 1-26 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片

(d) SEQ ID NO: 1-26 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片 10

【請求項 2】

SEQ ID NO: 1-26 からなる群から選択したアミノ酸配列を有する、請求項 1 に記載の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする、単離されたポリヌクレオチド。

【請求項 4】

請求項 2 のポリペプチドをコードする、単離されたポリヌクレオチド。

【請求項 5】

SEQ ID NO: 27-52 からなる群から選択したポリヌクレオチド配列を有する、請求項 4 に記載の単離されたポリヌクレオチド。 20

【請求項 6】

請求項 3 に記載のポリヌクレオチドに機能的に連結したプロモーター配列を持つ組換えポリヌクレオチド。

【請求項 7】

請求項 6 に記載の組換えポリヌクレオチドを用いて形質転換した細胞。

【請求項 8】

請求項 6 に記載の組換えポリヌクレオチドを持つ遺伝形質転換体。

【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、 30

(a) 前記ポリペプチドの発現に好適な条件下で、請求項 1 のポリペプチドをコードする或るポリヌクレオチドに機能的に連結されたプロモーター配列を持つ組換えポリヌクレオチドで形質転換される細胞を培養する過程と、

(b) そのように発現した前記ポリペプチドを回収する過程、とからなることを特徴とする方法。

【請求項 10】

前記ポリペプチドが、SEQ ID NO: 1-26 からなる群から選択したアミノ酸配列を持つことを特徴とする、請求項 9 に記載の方法。

【請求項 11】

請求項 1 に記載のポリペプチドと特異結合するような単離された抗体。 40

【請求項 12】

以下の (a) 乃至 (e) からなる群から選択した単離されたポリヌクレオチド。

(a) SEQ ID NO: 27-52 からなる群から選択した或るポリヌクレオチド配列を有するポリヌクレオチド

(b) SEQ ID NO: 27-52 からなる群から選択したポリヌクレオチド配列と少なくとも 90% が同一であるような天然ポリヌクレオチド配列を有するポリヌクレオチド

(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド

(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド

(e) (a) ~ (d) の RNA 等価物

【請求項 13】

請求項 1 2 に記載のポリヌクレオチドの少なくとも 6 0 の連続したヌクレオチドを持つ単離されたポリヌクレオチド。

【請求項 1 4】

請求項 1 2 のポリヌクレオチドの配列を持つ標的ポリヌクレオチドをサンプル中から検出する方法であって、

(a) 前記サンプル中の前記標的ポリヌクレオチドに相補的な或る配列を有する少なくとも 2 0 の連続したヌクレオチド群を持つ或るプローブと前記サンプルとをハイブリダイズし、該プローブが、或るハイブリダイゼーション複合体が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群との間に形成される条件下で、前記標的ポリヌクレオチドへ特異的にハイブリダイズする過程と、

10

(b) 前記ハイブリダイゼーション複合体の有無を検出し、該複合体が存在する場合にはオプションでその量を検出する過程、とを含む方法。

【請求項 1 5】

前記プローブが少なくとも 6 0 の連続したヌクレオチドを含むことを特徴とする請求項 1 4 に記載の方法。

【請求項 1 6】

請求項 1 2 のポリヌクレオチドの配列を持つ標的ポリヌクレオチドをサンプル中から検出する方法であって、

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

20

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

【請求項 1 7】

請求項 1 のポリペプチドと、薬剤として許容できる賦形剤とを有することを特徴とする組成物。

【請求項 1 8】

前記ポリペプチドが、SEQ ID NO: 1-26 からなる群から選択されたアミノ酸配列を持つことを特徴とする、請求項 1 7 の組成物。

【請求項 1 9】

30

機能的な MDDT の発現の低下に関連する疾患や病態の治療方法であって、そのような治療を要する患者への請求項 1 7 の組成物の投与を含む治療方法。

【請求項 2 0】

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

(a) 請求項 1 のポリペプチドを有する或るサンプルを或る化合物に曝す過程と、

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

【請求項 2 1】

請求項 2 0 の方法で同定したアゴニスト化合物を有し、薬剤として許容できる賦形剤をも有する組成物。

40

【請求項 2 2】

機能的な MDDT の発現の低下に関連する疾患や病態の治療方法であって、そのような治療を要する患者への請求項 2 1 の組成物の投与を含む治療方法。

【請求項 2 3】

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

(a) 請求項 1 のポリペプチドを有する或るサンプルを或る化合物に曝す過程と、

(b) 前記サンプルにおいてアンタゴニスト活性を検出する過程、とを含むことを特徴とするスクリーニング方法。

50

【請求項 24】

請求項 23 の方法で同定したアンタゴニスト化合物と、薬剤として許容できる賦形剤とを有する組成物。

【請求項 25】

機能的 M D D T の過剰発現に関連する疾患や病態の治療方法であって、そのような治療を要する患者への請求項 24 の組成物の投与を含む治療方法。

【請求項 26】

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

(a) 請求項 1 に記載のポリペプチドを適切な諸条件下で少なくとも 1 つの試験化合物と混合する過程と、

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

【請求項 27】

請求項 1 のポリペプチドの活性をモジュレート（調節）する化合物をスクリーニングする方法であって、

(a) 請求項 1 に記載のポリペプチドの活性が許容された諸条件下で、請求項 1 に記載のポリペプチドを少なくとも 1 つの試験化合物と混合する過程と、

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

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

【請求項 28】

請求項 5 の配列を持つ標的ポリヌクレオチドの発現を改変するのに効果的な化合物をスクリーニングする方法であって、

(a) 前記標的ポリヌクレオチドの発現に好適な諸条件下で、前記標的ポリヌクレオチドを有する或るサンプルを或る化合物に曝露する過程と、

(b) 前記標的ポリヌクレオチドの、改変された発現を検出する過程と、

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

【請求項 29】

試験化合物の毒性を算定する方法であって、

(a) 核酸群を有する或る生体サンプルを前記試験化合物で処理する過程と、

(b) 処理した前記生体サンプルの核酸群と、請求項 12 のポリヌクレオチドの少なくとも 20 の連続するヌクレオチド群を有する或るプローブをハイブリダイズさせる過程であって、このハイブリダイゼーションが、前記プローブと前記生体サンプルの或る標的ポリヌクレオチドとの間で或る特異的なハイブリダイゼーション複合体が形成される諸条件下で行われ、前記標的ポリヌクレオチドが、請求項 12 のポリヌクレオチドのポリヌクレオチド配列またはその断片を有するポリヌクレオチドである、前記過程と、

(c) ハイブリダイゼーション複合体の収量を定量する過程と、

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

【請求項 30】

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

(a) 前記抗体が前記ポリペプチドに結合し、抗体とポリペプチドとの複合体が形成されるのに適した諸条件下で、前記生体サンプルを請求項 11 に記載の抗体と混合する過程と

10

20

30

40

50

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

【請求項 3 1】

請求項 1 1 の抗体であって、

- (a) キメラ抗体
- (b) 単鎖抗体
- (c) F a b 断片
- (d) F (a b ')₂ 断片
- (e) ヒト化抗体、のいずれかであることを特徴とする抗体。

10

【請求項 3 2】

請求項 1 1 の抗体と、許容できる賦形剤とを有する組成物。

【請求項 3 3】

被検者における M D D T の発現に関連する症状又は疾患の診断方法であって、請求項 3 2 に記載の組成物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。

【請求項 3 4】

前記抗体が標識されることを特徴とする請求項 3 2 の組成物。

【請求項 3 5】

被検者における M D D T の発現に関連する症状又は疾患の診断方法であって、請求項 3 4 に記載の組成物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。

20

【請求項 3 6】

請求項 1 1 の抗体の特異性を持つポリクローナル抗体を調製する方法であって、

(a) 抗体応答を誘発する諸条件下で、S E Q I D N O : 1 - 2 6 からなる群から選択した或るアミノ酸配列またはその免疫原性断片を有する或るポリペプチドを用いて或る動物を免疫化する過程と、

(b) 前記動物から抗体を単離する過程と、

(c) 前記単離された抗体を該ポリペプチドでスクリーニングし、それによって、S E Q I D N O : 1 - 2 6 からなる群から選択した或るアミノ酸配列を有する或るポリペプチドに特異結合するようなポリクローナル抗体を同定する過程、とを含むことを特徴とする方法。

30

【請求項 3 7】

請求項 3 6 の方法で産生したポリクローナル抗体。

【請求項 3 8】

請求項 3 7 のポリクローナル抗体と好適なキャリアとを含む組成物。

【請求項 3 9】

請求項 1 1 に記載の抗体の特異性を有するモノクローナル抗体を作製する方法であって、

(a) 抗体応答を誘発する諸条件下で、S E Q I D N O : 1 - 2 6 からなる群から選択した或るアミノ酸配列またはその免疫原性断片を有する或るポリペプチドを用いて或る動物を免疫化する過程と、

(b) 前記動物から、抗体を産出する細胞を単離する過程と、

(c) 不死化した細胞と前記抗体産出細胞とを融合し、モノクローナル抗体を産出するハイブリドーマ細胞を形成する過程と、

(d) 前記ハイブリドーマ細胞を培養する過程と、

(e) S E Q I D N O : 1 - 2 6 からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するようなモノクローナル抗体を前記培養物から単離する過程、とを含むことを特徴とする方法。

40

【請求項 4 0】

請求項 3 9 に記載の方法で産出したモノクローナル抗体。

【請求項 4 1】

請求項 4 0 に記載のモノクローナル抗体と適切なキャリアとを含む組成物。

50

【請求項 4 2】

F a b 発現ライブラリをスクリーニングすることにより産出されることを特徴とする請求項 1 1 に記載の抗体。

【請求項 4 3】

組換え免疫グロブリンライブラリをスクリーニングすることにより産出されることを特徴とする、請求項 1 1 に記載の抗体。

【請求項 4 4】

S E Q I D N O : 1 - 2 6 からなる群から選択したアミノ酸配列を有するポリペプチドをサンプル中に検出する方法であって、

(a) 前記抗体と前記ポリペプチドとの特異結合を許容する条件下で、或るサンプルと共に請求項 1 1 に記載の抗体をインキュベートする過程と、 10

(b) 特異結合を検出する過程とを含み、該特異結合が、S E Q I D N O : 1 - 2 6 からなる群から選択したアミノ酸配列を有するポリペプチドがサンプル中に存在することを標示することを特徴とする方法。

【請求項 4 5】

或るサンプルからの S E Q I D N O : 1 - 2 6 からなる群から選択したアミノ酸配列を有するポリペプチドを精製する方法であって、

(a) 前記抗体と前記ポリペプチドとの特異結合を許容する条件下で、或るサンプルと共に請求項 1 1 に記載の抗体をインキュベートする過程と、

(b) 前記サンプルから前記抗体を分離し、S E Q I D N O : 1 - 2 6 からなる群から選択した或るアミノ酸配列を有する精製ポリペプチドを得る過程、とを含むことを特徴とする方法。 20

【請求項 4 6】

マイクロアレイの少なくとも 1 つのエLEMENT が請求項 1 3 に記載のポリヌクレオチドであることを特徴とするマイクロアレイ。

【請求項 4 7】

ポリヌクレオチド群を有する或るサンプルの発現プロフィールを作製する方法であって、

(a) 該サンプルの該ポリヌクレオチド群を標識する過程と、

(b) ハイブリダイゼーション複合体の形成に適した諸条件下で請求項 4 6 のマイクロアレイのエLEMENT 群をサンプルの標識されたポリヌクレオチド群と接触させる過程と、 30

(c) 該サンプル中の該ポリヌクレオチド群の発現を定量化する過程、とを含むことを特徴とする方法。

【請求項 4 8】

或る固体基板上の固有の物理的位置に付着された種々のヌクレオチド分子を有するアレイであって、少なくとも 1 つの前記ヌクレオチド分子が、或る標的ポリヌクレオチドの少なくとも 3 0 の連続したヌクレオチド群と特異的にハイブリダイズ可能な最初のオリゴヌクレオチドまたはポリヌクレオチド配列を有し、前記の標的ポリヌクレオチドが請求項 1 2 に記載のポリヌクレオチドであることを特徴とするアレイ。

【請求項 4 9】

請求項 4 8 に記載のアレイで、前記のオリゴヌクレオチドまたはポリヌクレオチドの最初の配列が前記の標的ポリヌクレオチドの少なくとも 3 0 の連続したヌクレオチド群に完全に相補的であることを特徴とするアレイ。 40

【請求項 5 0】

請求項 4 8 に記載のアレイで、前記のオリゴヌクレオチドまたはポリヌクレオチドの最初の配列が前記の標的ポリヌクレオチドの少なくとも 6 0 の連続したヌクレオチド群に完全に相補的であることを特徴とするアレイ。

【請求項 5 1】

請求項 4 8 に記載のアレイで、前記のオリゴヌクレオチドまたはポリヌクレオチドの最初の配列が前記の標的ポリヌクレオチドに完全に相補的であることを特徴とするアレイ。

【請求項 5 2】

請求項 48 に記載のアレイで、マイクロアレイであることを特徴とするアレイ。

【請求項 53】

請求項 48 に記載のアレイで、更に前記のオリゴヌクレオチドまたはポリヌクレオチドの最初の配列を有する或るヌクレオチド分子にハイブリダイズした前記の標的ポリヌクレオチドを有することを特徴とするアレイ。

【請求項 54】

請求項 48 に記載のアレイで、或るリンカーが少なくとも 1 つの前記のヌクレオチド分子と前記の固体基板とを連結していることを特徴とするアレイ。

【請求項 55】

請求項 48 に記載のアレイで、該基板上の固有の物理的位置の各々が複数のヌクレオチド分子を含み、任意の単一の固有の物理的位置でのその複数のヌクレオチド分子は同一の配列を有し、該基板上の固有の物理的位置の各々は、該基板上の別の固有の物理的位置でのヌクレオチド分子群の配列とは異なる或る配列を有するヌクレオチド分子群を含むことを特徴とするアレイ。 10

【請求項 56】

SEQ ID NO: 1 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 57】

SEQ ID NO: 2 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 58】

SEQ ID NO: 3 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。 20

【請求項 59】

SEQ ID NO: 4 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 60】

SEQ ID NO: 5 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 61】

SEQ ID NO: 6 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 62】

SEQ ID NO: 7 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 63】

SEQ ID NO: 8 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。 30

【請求項 64】

SEQ ID NO: 9 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 65】

SEQ ID NO: 10 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 66】

SEQ ID NO: 11 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 67】

SEQ ID NO: 12 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 68】

SEQ ID NO: 13 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。 40

【請求項 69】

SEQ ID NO: 14 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 70】

SEQ ID NO: 15 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 71】

SEQ ID NO: 16 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 72】

SEQ ID NO: 17 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 73】

SEQ ID NO: 18 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。 50

【請求項 74】

SEQ ID NO: 19 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 75】

SEQ ID NO: 20 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 76】

SEQ ID NO: 21 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 77】

SEQ ID NO: 22 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 78】

SEQ ID NO: 23 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 79】

SEQ ID NO: 24 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 80】

SEQ ID NO: 25 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 81】

SEQ ID NO: 26 のアミノ酸配列を有する請求項 1 に記載のポリペプチド。

【請求項 82】

SEQ ID NO: 27 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 83】

SEQ ID NO: 28 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 84】

SEQ ID NO: 29 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 85】

SEQ ID NO: 30 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 86】

SEQ ID NO: 31 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 87】

SEQ ID NO: 32 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 88】

SEQ ID NO: 33 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 89】

SEQ ID NO: 34 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 90】

SEQ ID NO: 35 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 91】

SEQ ID NO: 36 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 92】

SEQ ID NO: 37 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 93】

10

20

30

40

50

SEQ ID NO : 38 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 94】

SEQ ID NO : 39 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 95】

SEQ ID NO : 40 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 96】

SEQ ID NO : 41 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。 10

【請求項 97】

SEQ ID NO : 42 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 98】

SEQ ID NO : 43 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 99】

SEQ ID NO : 44 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。 20

【請求項 100】

SEQ ID NO : 45 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 101】

SEQ ID NO : 46 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 102】

SEQ ID NO : 47 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 103】

SEQ ID NO : 48 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。 30

【請求項 104】

SEQ ID NO : 49 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 105】

SEQ ID NO : 50 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【請求項 106】

SEQ ID NO : 51 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。 40

【請求項 107】

SEQ ID NO : 52 のポリヌクレオチド配列を有する請求項 12 に記載のポリヌクレオチド。

【発明の詳細な説明】

【0001】

(技術分野)

本発明は、疾患検出および治療用分子の核酸配列及びアミノ酸配列に関する。本発明はまた、これらの配列を利用した、細胞増殖異常、自己免疫／炎症性の疾患、発達障害、および神経障害の診断・治療・予防に関する。本発明はさらに、疾患検出および治療用分子の 50

核酸配列及びアミノ酸配列の発現における外来化合物の効果についての算定に関する。

【0002】

(発明の背景)

推定によれば、哺乳類DNAの2%のみがタンパク質をコードしており、どの時点でも、タンパク質をコードする遺伝子群の僅かな部分のみが特定細胞内で実際に発現される。多細胞生物における種々のタイプの細胞は構造も機能も劇的に異なっており、特定細胞の個性を授けるものは、その独自のパターンの遺伝子発現である。また、異なる細胞タイプは、オーバーラップするが特有のセットの遺伝子群を、発生期を通じて発現する。生物の発達と生存とに寄与する、細胞成長、細胞増殖、細胞分化、免疫応答、アポトーシス及び他のプロセスは、遺伝子発現の調節によって支配される。適切な遺伝子調節はまた、細胞を効率よく機能させる。これを確実にするため、その機能が要求される遺伝子群のみが所定の時期に発現される。遺伝子発現に影響を与える因子の例には、細胞間伝達を仲介し様々な細胞型の作用を協調させる細胞外シグナルを含む。遺伝子発現は、DNAとRNAとの転写レベル、および、mRNA翻訳のレベルにおいて調節される。

10

【0003】

遺伝子と遺伝子産物とに異常な発現または突然変異があると、種々のヒト疾患（例えば癌その他の細胞増殖異常など）を起こしたり、またはそれらに対する感受性を増加させることがある。これらの遺伝子と遺伝子産物の同定が、疾患の早期検出のためのマーカーと、疾患の予防と治療との為の標的とを見出すために進展している努力の基礎となる。例えば癌は、身体ほぼ全ての組織に影響する細胞増殖異常の或るタイプを言う。癌の発生すなわち発癌は、しばしば、異常発現または突然変異を通じて正常遺伝子が癌を起こす遺伝子すなわち癌原遺伝子へ変換することと相関する。腫瘍性タンパク質（癌原遺伝子の産物）の例には、細胞増殖に影響する種々の分子を含む。このような分子とは例えば成長因子、成長因子受容体、細胞内シグナル伝達因子、核内転写因子、および、細胞周期制御タンパク質である。一方、腫瘍抑制遺伝子は、細胞増殖抑制に関係する。腫瘍抑制遺伝子の機能を低減または抑止する突然変異の結果、異常な細胞増殖と癌とが起こる。このように、多種多様な遺伝子と遺伝子産物とが細胞増殖異常（癌など）に関連することが見出されているが、まだ発見されていない多くの遺伝子と遺伝子産物が存在する可能性がある。

20

【0004】

DNAベースのアレイは、遺伝子発現と遺伝的変異性とを試験する、効率的で高処理の方法を提供できる。例えばSNP（すなわち一塩基多型）は、最も多いタイプのヒト遺伝的変異である。DNAベースのアレイは、数百、数千もの遺伝子群におけるSNPの発見を劇的に加速できる。同様に、このようなアレイを用いて、SNP遺伝子型解析（SNP genotyping）をなす。この解析では、個体群または集団群からのDNAサンプルのアッセイを、選択したSNPの存在について行う。これらのアプローチは、ヒトゲノムにおける全ての遺伝的変異の体系的同定や、特定の遺伝的変異と疾患感受性、薬物治療応答性その他の医学的関連情報との相関の体系的同定を最終的にもたらすだろう（例えばWang, D. G. 他（1998）Science 280:1077-1082等を参照）。

30

【0005】

DNAベースのアレイ技術は、全体的な遺伝子発現パターンの迅速な分析に特に重要である。例えば遺伝子的疾病素質、疾患、または治療処置は、直接または間接に、或る組織における多数の遺伝子の発現に影響することがある。この場合、発現される全ての遺伝子と、それらの遺伝子とその特定組織において発現されるレベルとのプロファイルすなわち転写イメージを作製することは有用である。或る疾患を患う個体または集団、または特定の治療を受けている個体または集団から作製されたプロファイルは、対照の個体または集団から作製されたプロファイルと比較されうる。このような分析には、遺伝子機能の知識を必要としない。理由は、発現プロファイルを数学的に分析でき、この分析では各遺伝子を単にマーカーとして扱うからである。更に遺伝子発現プロファイルは、例えば或る発生段階や特定組織において、または疾患や処置に応答して発現される全ての遺伝子を同定する

40

50

ことによって、生物学的経路の詳細な分析を補助しうる（例えばL a n d e r, E. S. 他（1996）S c i e n c e 274:536-539等を参照）。

【0006】

数種の遺伝子は、疾患と関連することが、染色体上の位置から知られている。その例として、マウスとヒトとの筋緊張性ジストロフィー（DM）領域にある遺伝子群がある。DMの根底にある突然変異はDMキナーゼ蛋白をコードする或る遺伝子に限局されているが、別の活性遺伝子であるDMR-N9は、このDMキナーゼ遺伝子に非常に近接した位置にある（J a n s e n, G. 他（1992）N a t. G e n e t. 1:261-266）。DMR-N9は、或る650個のアミノ酸のタンパク質をコードし、このタンパク質にはWDリピートがある。WDリピートは細胞シグナル伝達タンパクに見られるモチーフである。DMR-N9は全ての神経組織と睪丸とに発現されるので、重篤例のDMにおける精神症状と睪丸症状との出現におけるDMR-N9の役割を示唆している（J a n s e n, G. 他（1995）H u m. M o l. G e n e t. 4:843-852）。

他の複数の遺伝子の同定が、その発現パターンに、または疾病症候群との関連に基づいてなされた。例えば細胞内オルガネラに対する自己抗体は、全身性リウマチ性疾患の患者に見られる。最近同定されたタンパク質であるg o l g i n-67はゴルジ自己抗原の或るファミリーに属しており、 α ヘリックスのコイルドコイルドメインを持つ（E y s t a t h i o y, T. 他（2000）J. A u t o i m m u n. 14:179-187）。S t a c 遺伝子は、脳に特異的で、発達に応じて調節される遺伝子として同定された。S t a c タンパク質はSH3ドメインを持つ。また、ニューロン特異的シグナル伝達に関わると思われる（S u z u k i, H. 他（1996）B i o c h e m. B i o p h y s. R e s. C o m m u n. 229:902-909）。

新規の疾患検出および治療用分子、並びにそれらをコードするポリヌクレオチドの発見により、新規の組成物を提供することで当分野の要望に応えることができる。この新規の組成物は、細胞増殖異常、自己免疫／炎症性の疾患、発達障害、および神経疾患の診断・治療・予防において有用であり、また、疾患検出および治療用分子の核酸配列及びアミノ酸配列の発現における外来性化合物の効果についての評価にも有用である。

【0007】

（発明の概要）

本発明は、総称して「MDDT」、個別にはそれぞれ「MDDT-1」、「MDDT-2」、「MDDT-3」、「MDDT-4」、「MDDT-5」、「MDDT-6」、「MDDT-7」、「MDDT-8」、「MDDT-9」、「MDDT-10」、「MDDT-11」、「MDDT-12」、「MDDT-13」、「MDDT-14」、「MDDT-15」、「MDDT-16」、「MDDT-17」、「MDDT-18」、「MDDT-19」、「MDDT-20」、「MDDT-21」、「MDDT-22」、「MDDT-23」、「MDDT-24」、「MDDT-25」および「MDDT-26」と呼ぶ、精製されたポリペプチドである、疾患検出および治療用分子を提供する。或る実施例において本発明は、（a）SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列を持つポリペプチド、（b）SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を持つ天然アミノ酸配列を有するポリペプチド、（c）SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列を持つポリペプチドの生物学的活性断片、および（d）SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した、単離されたポリペプチドを提供する。一実施例で本発明は、SEQ ID NO: 1-26のアミノ酸配列を持つ、単離されたポリペプチドを提供する。

【0008】

また、本発明は（a）SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列を持つポリペプチド、（b）SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を持つ或る天然アミノ酸配列を有するポリペプチド、（c）SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列

を持つポリペプチドの生物学的活性断片、または (d) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列を持つポリヌクレオチドの免疫原性断片、からなる群から選択した或る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。一実施態様では、該ポリヌクレオチドは、SEQ ID NO: 1-26 からなる群から選択した或るポリペプチドをコードする。別の実施態様では、ポリヌクレオチドは SEQ ID NO: 27-52 からなる群から選択される。

【0009】

本発明は更に、(a) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列を有するポリペプチドの生物学的活性断片、および (d) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した或る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連結したプロモーター配列を持つ組換えポリヌクレオチドを提供する。一実施態様で本発明は、この組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様で本発明は、この組換えポリヌクレオチドを持つ遺伝形質転換体を提供する。

10

【0010】

更に本発明は、(a) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列を持つポリペプチドと、(b) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列との少なくとも90%以上の同一性を有する或る天然アミノ酸配列を持つポリペプチドと、(c) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列を持つポリペプチドの生物学的活性断片と、(d) SEQ ID NO: 1-26 からなる群から選択したアミノ酸配列を持つポリペプチドの免疫原性断片、とからなる群から選択したポリペプチドを製造する一方法を提供する。製造方法は、(a) 或る細胞を該ポリペプチドの発現に適した条件下で培養し、この細胞を組換えポリヌクレオチドを用いて形質転換する過程と、(b) そのように発現したポリペプチドを回収する過程からなる。この組換えポリヌクレオチドは、該ポリペプチドをコードするポリヌクレオチドに対し機能的に連結したプロモーター配列を持つ。

20

【0011】

本発明は更に、(a) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO: 1-26 からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、および (d) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドに特異結合する、単離された抗体を提供する。

30

【0012】

本発明は更に、(a) SEQ ID NO: 27-52 からなる群から選択したポリヌクレオチド配列を持つポリヌクレオチド、(b) SEQ ID NO: 27-52 からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有する或る天然ポリヌクレオチド配列を持つポリヌクレオチド、(c) (a) に相補的なポリヌクレオチド、(d) (b) に相補的なポリヌクレオチド、および (e) (a) ~ (d) のRNA等価物、からなる群から選択した、単離されたポリヌクレオチドを提供する。一実施態様で該ポリヌクレオチドは、少なくとも60の連続したヌクレオチド群からなる。

40

【0013】

本発明は更に、サンプル中の標的ポリヌクレオチドを検出する一方法を提供する。ここで、標的ポリヌクレオチドは、(a) SEQ ID NO: 27-52 からなる群から選択した或るポリヌクレオチド配列を持つポリヌクレオチド、(b) SEQ ID NO: 27-52 からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を

50

有する天然ポリヌクレオチド配列を持つポリヌクレオチド、(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド、(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド、および(e) (a) ~ (d) のRNA等価物、からなる群から選択した、或るポリヌクレオチドの配列を持つ。検出方法は、(a) サンプル中の上記標的ポリヌクレオチドに相補的な或る配列からなる少なくとも20の連続したヌクレオチド群からなる或るプローブを用いて該サンプルをハイブリダイズする過程と、(b) 該ハイブリダイゼーション複合体の有無を検出し、複合体が存在すればオプションでその量を検出する過程からなる。該プローブと該標的ポリヌクレオチドあるいはその断片との間でハイブリダイゼーション複合体が形成されるような条件下で、プローブは、該標的ポリヌクレオチドに対し特異的にハイブリダイズする。一実施態様では、プローブは少なくとも60の連続したヌクレオチド群からなる。

10

【0014】

本発明はまた、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは、(a) SEQ ID NO: 27-52からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(b) SEQ ID NO: 7-52からなる群から選択したポリヌクレオチド配列と少なくとも90%の相同性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド、(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド、および(e) (a) ~ (d) のRNA等価物からなる群から選択された配列のポリヌクレオチドを有する。検出方法は、(a) ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、(b) 増幅した標的ポリヌクレオチドまたはその断片の有無を検出し、該標的ポリヌクレオチドまたはその断片が存在すればオプションでその量を検出する過程からなる。

20

【0015】

本発明は更に、或る有効量のポリペプチドと薬物として許容し得る或る賦形剤とからなる、或る組成物を提供する。有効量のポリペプチドは、(a) SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列を含むポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択される。一実施態様では、SEQ ID NO: 1-26からなる一群から選択されたアミノ酸配列を持つ組成物を提供する。更に、本発明は、この組成物を機能的MDDTの発現の低下に関連した疾患やその症状の治療を必要とする患者に投与する過程を含む方法を提供する。

30

【0016】

本発明はまた、(a) SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO: 1-26からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-26からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択したポリペプチドのアゴニストとしての有効性を確認するために、或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドを有するサンプルを或る化合物に曝す過程と、(b) サンプル中のアゴニスト活性を検出する過程からなる。別法では、本発明は、この方法で同定したアゴニスト化合物と許容される医薬用賦形剤とを有する、或る組成物を提供する。更なる別法では、本発明は、機能的MDDTの発現の低下を伴う疾患や症状の治療を要する患者への、この組成物の投与方法を提供する。

40

【0017】

本発明は更に、(a) SEQ ID NO: 1-26からなる群から選択した或るアミノ酸

50

配列を持つポリペプチド、(b) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO: 1-26 からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-26 からなる群から選択したアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択したポリペプチドのアンタゴニストとしての有効性につき、或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドを含むサンプルを或る化合物に曝す過程と、(b) サンプル中のアンタゴニスト活性を検出する過程からなる。別の一実施態様で本発明は、この方法で同定したアンタゴニスト化合物と薬物として許容し得る賦形剤とを有する組成物を提供する。更なる別法で本発明は、機能的MDDTの過剰な発現に関連した疾患やその症状の治療を必要とする患者への、この組成物の投与を含む方法を提供する。

10

【0018】

本発明は更に、(a) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO: 1-26 からなる群から選択したアミノ酸配列を持つポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1-26 からなる群から選択したアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドに特異結合する或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドを適切な条件下で少なくとも1つの試験化合物に混合させる過程と、(b) この試験化合物と該ポリペプチドとの結合を検出し、それにより該ポリペプチドに特異結合する化合物を同定する過程からなる。

20

【0019】

本発明は更に、(a) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列を持つポリペプチド、(b) SEQ ID NO: 1-26 からなる群から選択したアミノ酸配列との少なくとも90%の同一性を有する或る天然アミノ酸配列を持つポリペプチド、(c) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列を持つポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1-26 からなる群から選択した或るアミノ酸配列を持つポリペプチドの免疫原性断片、からなる群から選択した或るポリペプチドの活性をモジュレートする或る化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) 該ポリペプチドの活性にとり許容し得る条件下で、該ポリペプチドを少なくとも1つの試験化合物と混合する過程と、(b) 該ポリペプチドの活性をこの試験化合物の存在下で算定する過程と、(c) この試験化合物の存在下での該ポリペプチドの活性をこの試験化合物の不存在下での該ポリペプチドの活性と比較する過程からなり、この試験化合物の存在下での該ポリペプチドの活性の変化は、該ポリペプチドの活性をモジュレートする化合物を標示する。

30

【0020】

更に本発明は、SEQ ID NO: 27-52 からなる群から選択した或るポリヌクレオチド配列を持つ標的ポリヌクレオチドの発現を改変する効果につき、或る化合物をスクリーニングする一方法を提供する。この方法は、(a) この標的ポリヌクレオチドを有するサンプルを或る化合物に曝露する過程と、(b) この標的ポリヌクレオチドの発現の改変を検出する過程と、(c) 可変量のこの化合物の存在下でのこの標的ポリヌクレオチドの発現と、この化合物の不存在下での発現とを比較する過程とからなる。

40

【0021】

本発明は更に、試験化合物の毒性の算定方法を提供する。この方法には、以下の過程がある。(a) 核酸群を有する生体サンプルを試験化合物で処理する過程。(b) 処理済み生体サンプルの核酸群をハイブリダイズする過程。この過程には、次のようなプローブを用いる。(i) SEQ ID NO: 27-52 からなる群から選択した或るポリヌクレオチド配列を持つポリヌクレオチド、(ii) SEQ ID NO: 27-52 からなる群から

50

選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有する天然ポリヌクレオチド配列を持つポリヌクレオチド、(iii)(i)に相補的な配列を持つポリヌクレオチド、(iv)(ii)のポリヌクレオチドに相補的なポリヌクレオチド、(v)(i)~(iv)のRNA等価物、からなる群から選択した或るポリヌクレオチドの少なくとも20の連続したヌクレオチド群からなるプローブである。ハイブリダイゼーションは、上記プローブと生体サンプル中の標的ポリヌクレオチドとの間に特異的ハイブリダイゼーション複合体が形成されるような条件下で生じる。上記標的ポリヌクレオチドは、(i)SEQ ID NO: 27-52からなる群から選択した或るポリヌクレオチド配列を持つポリヌクレオチド、(ii)SEQ ID NO: 27-52からなる群から選択した或るポリヌクレオチド配列との少なくとも90%の同一性を有する天然ポリヌクレオチド配列を持つポリヌクレオチド、(iii)(i)のポリヌクレオチドに相補的なポリヌクレオチド、(iv)(ii)のポリヌクレオチドに相補的なポリヌクレオチド、(v)(i)~(iv)のRNA等価物、からなる群から選択する。あるいは、標的ポリヌクレオチドは、上記(i)~(v)からなる群から選択したポリヌクレオチド配列の断片を持つ。毒性の算定方法には更に、以下の過程がある。(c)ハイブリダイゼーション複合体の量を定量する過程と、(d)処理済み生体サンプル中のハイブリダイゼーション複合体の量を、非処理の生体サンプル中のハイブリダイゼーション複合体の量と比較する過程である。処理済み生体サンプル中のハイブリダイゼーション複合体の量の差異が、試験化合物の毒性を示す。

10

20

【0022】

(本発明の記述について)

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

【0023】

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

30

【0024】

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

40

【0025】

(定義)

用語「MDDT」は、天然、合成、半合成或いは組換え体などの、全ての種(特にウシ、ヒツジ、ブタ、ネズミ、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MDDTのアミノ酸配列を指す。

【0026】

用語「アゴニスト」は、MDDTの生物学的活性を強める、或いは模倣する分子を指す。アゴニストの例としては、タンパク質、核酸、糖質、小分子その他の任意の化合物や組成物の内、MDDTと直接に相互作用することによって、或いはMDDTが関与する生物学

50

的経路の種々の構成成分に作用することによってMDDTの活性を調節するものを含むうる。

【0027】

「対立遺伝子変異体／変異配列」は、MDDTをコードする遺伝子の、別の形態である。対立遺伝子変異配列群は、核酸配列における少なくとも1つの突然変異から生じ得る他、変容したmRNA群を生じ得る。また、ポリペプチド群を生じ得るが、それらのポリペプチドの構造または機能は、変容することもしないこともある。或る遺伝子は、その天然型の対立遺伝子変異配列を全く持たない場合もあり、1個以上持つこともある。対立遺伝子変異配列を生じさせる通常の突然変異性変化は一般に、ヌクレオチドの自然な欠失、付加または置換による。これら各種の変化は、単独であるいは他の変化と共に、或る配列内で1回以上、生じ得る。

10

【0028】

MDDTをコードする「変容した／改変された」核酸配列としては、様々なヌクレオチドの欠失、挿入、あるいは置換がありながら、MDDTと同じポリペプチド、あるいはMDDTの機能特性の少なくとも1つを備えるポリペプチドをもたらす配列もある。この定義には、MDDTをコードするポリヌクレオチド配列にとって正常な染色体の遺伝子座ではない位置での、対立遺伝子変異配列との不適当あるいは予期しないハイブリダイゼーションを含み、また、MDDTをコードするポリヌクレオチドの或る特定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あるいは検出困難な多型性を含む。コードされるタンパク質も「変容する／改変される」ことがあり、また、サイレント変化を生じて機能的には等価なMDDTと成るような、アミノ酸残基の欠失、挿入、あるいは置換を持ち得る。意図的なアミノ酸置換は、生物学的或いは免疫学的にMDDTの活性が保持される範囲で、残基の極性、電荷、溶解度、疎水性、親水性、及び／または両親媒性、についての類似性に基づいて成され得る。例えば、負に帯電したアミノ酸にはアスパラギン酸およびグルタミン酸があり、正に帯電したアミノ酸にはリジンおよびアルギニンがある。親水性値が近似した非荷電極性側鎖を持つアミノ酸には、アスパラギンとグルタミン、セリンとトレオニンがある。親水性値が近似した非荷電側鎖を持つアミノ酸には、ロイシンとイソロイシンとバリン、グリシンとアラニン、フェニルアラニンとチロシンがある。

20

【0029】

用語「アミノ酸」および「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、あるいはそれらのいずれかの断片を指し、天然分子または合成分子を指す。「アミノ酸配列」が或る天然タンパク質分子の配列を指す場合、「アミノ酸配列」および類似の用語は、記載したその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。

30

【0030】

用語「増幅」は、或る核酸配列の付加的複製物を作製する行為に関する。増幅には通常、当業者に公知のポリメラーゼ連鎖反応（PCR）技術を用いる。

【0031】

用語「アンタゴニスト」は、MDDTの生物学的活性を阻害あるいは減弱する分子を指す。アンタゴニストとしては、MDDTと直接相互作用すること、あるいはMDDTが関与する生物学的経路の各成分に作用することによってMDDTの活性をモジュレートする、抗体などタンパク質、核酸、糖質、小分子、または任意の他の化合物や組成物があり得る。

40

【0032】

用語「抗体」は、エピトープの決定基と結合できる、無傷の免疫グロブリン分子やそれらの断片、例えばFab、F(ab')₂、およびFv断片を指す。MDDTポリペプチドと結合する抗体は、免疫化抗原として、無傷のポリペプチドを用いて、または、目的の小ペプチドを持つ断片を用いて作製可能である。マウス、ラット、ウサギなどの動物を免疫化するために用いるポリペプチドまたはオリゴペプチドの由来は、RNAの翻訳の場合や、または化学合成があり得る。また、それらは所望により、キャリアタンパク質に抱合

50

し得る。通常用いられるキャリアであってペプチドと化学結合するキャリアの例は、ウシ血清アルブミン、サイログロブリン、および、スカシガイのヘモシアニン（K L H）などがある。結合したこのペプチドは、前記動物を免疫化するために用いる。

【0033】

用語「抗原決定基」は、特定の抗体と接触する、分子の領域（すなわちエピトープ）を指す。タンパク質またはタンパク質断片を用いて宿主動物を免疫化する場合、タンパク質の多数の領域が、抗原決定基（タンパク質の特定の領域または3次元構造）に特異結合する抗体の産生を誘発し得る。或る抗原決定基は、或る抗体への結合について、無傷抗原（すなわ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】

用語「アプタマー（aptamer）」は、核酸またはオリゴヌクレオチド分子であって、特定の分子ターゲットに結合する分子を指す。アプタマーは*in vitro*での進化プロセスに由来する（例えば、SELEX（Systematic Evolution of Ligands by EXponential Enrichmentの略、試験管内選択法）、米国特許第5,270,163号に記述）。これは、大規模な組み合わせライブラリ群から標的特異的アプタマー配列を選択するプロセスである。アプタマーの構成は二本鎖または一本鎖であり、デオキシリボヌクレオチド、リボヌクレオチド、ヌクレオチド誘導体または他のヌクレオチド様分子を含み得る。アプタマーの各ヌクレオチド成分は修飾された糖基を持つことがあり（例えば、リボヌクレオチドの2'-OH基は2'-Fまたは2'-NH₂によって置換され得る）、これは、ヌクレアーゼへの抵抗性または血中でのより長い寿命など、所望の特性を向上し得る。アプタマーを他の分子（例えば高分子量のキャリア）と抱合させることにより、循環系からのこのアプタマーのクリアランスを遅らせ得る。アプタマー類は、例えば或る架橋剤の光活性化によって、それらの同種リガンド群と特異的に架橋させ得る（例えば、Brody, E. N. および L. Gold (2000) J. Biotechnol. 74:5-13を参照）。

【0035】

用語「イントラマー（intramer）」は、*in vivo*で発現されるアプタマーを指す。例えば、ワクシニアウイルスに基づく或るRNA発現系を用いて、白血球の細胞質内で特定のRNAアプタマー類が高レベルに発現されている（Blind, M. 他 (1999) Proc. Natl Acad. Sci. USA 96:3606-3610）。

【0036】

用語「スピーゲルマー（spiegelmer）」は、アプタマーの内、L-DNA、L-RNAなどの左旋性ヌクレオチド誘導体または左旋性ヌクレオチド様分子などを指す。左旋性のヌクレオチド群を持つアプタマー類は、右旋性ヌクレオチド群を持つ基質に通常は作用する、天然酵素類による分解に対して耐性がある。

【0037】

用語「アンチセンス」は、或る特定の核酸配列の「センス」（コーディング）鎖との塩基対を形成し得る任意の組成物を指す。アンチセンス組成物としては、DNAや、RNAや、ペプチド核酸（PNA）や、修飾されたバックボーン連結たとえばホスホロチオ酸、メチルホスホン酸またはベンジルホスホン酸などを有するオリゴヌクレオチドや、修飾された糖基たとえば2'-メトキシエチル糖または2'-メトキシエトキシ糖などを有するオリゴヌクレオチドや、あるいは修飾された塩基たとえば5-メチルシトシン、2-デオキシウラシルまたは7-デアザ-2'-デオキシグアノシンなどを有するオリゴヌクレオチドがあり得る。アンチセンス分子は、化学合成または転写など、任意の方法で製造することができる。相補的アンチセンス分子は、細胞に導入されると、細胞が産生した天然核酸配列との塩基対を形成し、二重鎖を形成して転写または翻訳を妨害する。「負」または「マイナス」という表現は、或る参考DNA分子のアンチセンス鎖を指すことがあり、「正」または「プラス」という表現は、或る参考DNA分子のセンス鎖を意味し得る。

【0038】

10

20

30

40

50

用語「生物学的に活性」は、或る天然分子の構造的、調節的、ある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、天然あるいは組換え体のMDDT、合成のMDDT、またはそれらの任意のオリゴペプチドが、適当な動物の或いは細胞の特異的免疫応答を誘発して特異抗体群と結合する能力を指す。

【0039】

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、「5'-AGT-3'」は、その相補配列「3'-TCA-5'」との対を形成する。

【0040】

「～のポリヌクレオチド配列を含む（持つ）組成物」または「～のアミノ酸配列を含む（持つ）組成物」は、広い意味で、指定のポリヌクレオチド配列若しくはアミノ酸配列を持つ、任意の組成物を指す。この組成物には、乾燥製剤または水溶液が含まれ得る。MDDTをコード、若しくはその断片をコードするポリヌクレオチド配列を持つ組成物は、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も可能である。ハイブリダイゼーションにおいては、塩（例えばNaCl）、界面活性剤（例えばドデシル硫酸ナトリウム；SDS）およびその他の成分（例えばデンハート液、粉乳、サケの精子のDNAなど）を含む水溶液中にプローブを分散させることができる。

【0041】

「コンセンサス配列」は、不要な塩基を除くためにDNA配列解析を繰り返し行い、XL-PCRキット（Applied Biosystems, Foster City CA）を用いて5'および/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）などの断片構築用コンピュータプログラムを用いて1つあるいはそれ以上のオーバーラップする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。伸長および構築の両方を行ってコンセンサス配列を産生する配列もある。

【0042】

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と予測される置換を指す。すなわち、そのタンパク質の構造、特にその機能が保存され、置換によって大きくは変わらない。下表は、タンパク質内で元のアミノ酸と置換され得るアミノ酸であり、保存的なアミノ酸置換と認められるアミノ酸を示す。

元の残基

保存的な置換

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

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

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

【0043】

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内の変化、あるいは1個以上のヌクレオチドが欠如するヌクレオチド配列内の変化を指す。

【0044】

用語「誘導体」は、化学修飾されたポリヌクレオチドまたはポリペプチドを指す。例えば、アルキル基、アシル基、ヒドロキシル基またはアミノ基による水素の置換は、ポリヌクレオチドの化学修飾に含まれ得る。ポリヌクレオチド誘導体は、天然分子の生物学的または免疫学的機能を少なくとも1つは保持しているポリペプチドをコードする。ポリペプチド誘導体は、次のようなプロセスによって修飾されたポリペプチドである。すなわち、グリコシル化、ポリエチレングリコール化(pegylation)、あるいは任意の同様なプロセスであって、誘導元のポリペプチドの少なくとも1つの生物学的若しくは免疫学的機能を保持するプロセスである。

【0045】

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

【0046】

「示差発現」は、少なくとも2例のサンプルを比較して判定する、増加(上方調節)、あるいは減少(下方調節)、または欠損した、遺伝子またはタンパク質の発現を指す。このような比較は例えば、処理済サンプルと不処理サンプル、または病態サンプルと健常サンプルとの間で行われ得る。

【0047】

「エキソンシャッフリング」は、異なるコード領域(エキソン)群の組換えを意味する。或るエキソンは、コードするタンパク質の1つの構造的または機能的ドメインを代表し得るため、安定したサブストラクチャー群の新たな組み合わせによって新規のタンパク質群を構築することが可能であり、新たなタンパク質機能の進化を促進できる。

【0048】

用語「断片」は、MDDTの、またはそれ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)と配列は同一であるが親配列より長さが短いものを指す。或る断片は、定義された配列の全長から1ヌクレオチド/アミノ酸残基を差し引いた長さよりも短い長さを有し得る。例えば或る断片は、5~1000の連続したヌクレオチドまたはアミノ酸残基を有し得る。プローブ、プライマー、抗原、治療用分子として、あるいはその他の目的のために用いられる或る断片は、少なくとも長さ5、10、15、20、25、30、40、50、60、75、100、150、250若しくは500の、連続したヌクレオチドあるいはアミノ酸残基であり得る。断片は、或る分子の特定領域から優先的に選択し得る。例えば、或るポリペプチド断片は、定義された或る配列内に見られるような或るポリペプチドの最初の250または500アミノ酸(または最初の25%または50%)から選択した、或る長さの連続したアミノ酸を持ち得る。これらの長さは明らかに例として挙げているものであり、本発明の実施態様では、配列表、表および図面を含む本明細書が支持する任意の長さであり得る。

【0049】

SEQ ID NO: 27-52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: 27-52を特異的に同定する固有のポリヌクレオチド配列の領域を持つ。SEQ ID NO: 27-52の或る断片は、例えば、ハイブリダイゼーションや増幅技術、またはSEQ ID NO: 27-52に関連ポリヌクレオチド配列

10

20

30

40

50

から区別する類似の方法に有用である。SEQ ID NO: 27-52の或る断片の正確な長さは、また、その断片が対応するSEQ ID NO: 27-52の領域は、その断片に意図した目的に基づき当業者が慣例的に決定し得る。

【0050】

SEQ ID NO: 1-26の或る断片は、SEQ ID NO: 27-52の或る断片によってコードされる。SEQ ID NO: 1-26の或る断片は、SEQ ID NO: 1-26を特異的に同定する固有のアミノ酸配列の領域を持つ。例えば、SEQ ID NO: 1-26の或る断片は、SEQ ID NO: 1-26を特異的に認識する抗体の開発における免疫原性ペプチドとして有用である。SEQ ID NO: 1-26の或る断片の正確な長さは、また、この断片に対応するSEQ ID NO: 1-26の領域は、その断片に意図する目的に基づき、当業者が慣例的に判定し得る。

10

【0051】

「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）と、それに続く1オープンリーディングフレームおよび翻訳終止コドンを有する配列である。或る「完全長」ポリヌクレオチド配列は、或る「完全長」ポリペプチド配列をコードする。

【0052】

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

【0053】

ポリヌクレオチド配列に適用される「一致率」または「%一致」の語は、標準化されたアルゴリズムを用いてアラインメントされた、2つ以上のポリヌクレオチド配列間で一致する残基の割合を意味する。標準化アルゴリズムは、2配列間のアラインメントを最適化するため、標準化された再現性のある方法で比較対象の2配列内にギャップ群を挿入し得るので、2つの配列をより有意に比較できる。

20

【0054】

ポリヌクレオチド配列間の一致率を判定するには、MEGALIGN version 3.12e配列アラインメントプログラムなどに組込まれているCLUSTAL Vアルゴリズムの、デフォルトのパラメータ群を用い得る。このプログラムは、LASERGENソフトウェアパッケージ（一組の分子生物学的分析プログラム）（DNASTAR, Madison WI）の一部である。このCLUSTAL Vは、Higgins, D. G. および P. M. Sharp (1989) CABIOS 5:151-153、Higgins, D. G. 他 (1992) CABIOS 8:189-191に記載されている。ポリヌクレオチド配列をペアワイズアラインメントする際のデフォルトパラメータは、Ktuple=2、gap penalty=5、window=4、「diagonals saved」=4と設定される。「weighted」残基重み付け表がデフォルトで選択される。CLUSTAL Vによる一致率の報告は、アラインメントされたポリヌクレオチド配列間の「percent similarity（類似性パーセント）」としてなされる。

30

【0055】

あるいは、米国国立バイオテクノロジー情報センター（NCBI）のBasic Local Alignment Search Tool（BLAST）が、一般的に用いられ、且つ、無料で利用可能な配列比較アルゴリズム一式を提供している（Altschul, S. F. 他 (1990) J. Mol. Biol. 215:403-410）。BLASTアルゴリズムは、幾つかの情報源から入手可能であり、メリーランド州ベセスダにあるNCBIおよびインターネット（http://www.ncbi.nlm.nih.gov/BLAST/）からも入手可能である。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」と呼ばれるツールも入手可能であり、これは

40

50

2つのヌクレオチド配列を直接のペアワイズで比較するために用いられる。「BLAST 2 Sequences」は、<http://www.ncbi.nlm.nih.gov/gorf/bl2.html>にアクセスして、対話形式で利用できる。「BLAST 2 Sequences」ツールは、blastn および blastp（以下に記載）の両方に用い得る。BLASTプログラムは、一般的には、ギャップ（gap）などのパラメータをデフォルト設定にセットして用いる。例えば、2つのヌクレオチド配列を比較するには、デフォルトパラメータに設定した「BLAST 2 Sequences」ツール Version 2.0.12（2000年4月21日）を用いてblastnを実行し得る。デフォルトパラメータの設定例を以下に示す。

【0056】

10

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

一致率は、ある定義された配列の全長（例えば特定のSEQ IDナンバーで定義された配列）で測定し得る。あるいは、より短い長さ、例えば、定義された、より大きな配列から得た断片（例えば少なくとも20、30、40、50、70、100または少なくとも200の連続したヌクレオチドの断片）の長さの一致率も測定し得る。ここに挙げた長さは単なる例示的なものに過ぎず、表、図および配列リストを含めた本明細書に記載された配列が支持する任意の断片長を用いて、一致率を測定し得る或る長さを説明し得ることを理解されたい。

20

【0057】

高度の同一性を示さない核酸配列が、それにもかかわらず遺伝子コードの縮重が原因で、類似のアミノ酸配列をコードする場合がある。この縮重を利用して核酸配列内で変化を生じさせて、全ての核酸配列が実質上同一のタンパク質をコードするような多数の核酸配列を生成し得ることを理解されたい。

30

【0058】

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

【0059】

ポリペプチド配列間の一致率は、MEGALIGN version 3.12e配列アラインメントプログラムに組込まれているようなCLUSTAL Vアルゴリズムのデフォルトのパラメータ群を用いて判定できる（参照先と共に上述）。CLUSTAL Vを用いて、ポリペプチド配列をペアワイズアラインメントする際のデフォルトパラメータは、Ktuple=1、gap penalty=3、window=5、「diagonals saved」=5と設定される。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。ポリヌクレオチドの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一致率は、CLUSTAL Vによって「percent similarity（類似性パーセント）」として報告される。

40

【0060】

あるいは、NCBI BLASTソフトウェア一式を用い得る。例えば、2つのポリペプチド配列をペアワイズで比較する場合、「BLAST 2 Sequences」ツールV

50

ersion 2.0.12 (2000年4月21日)のblastpをデフォルトパラメータに設定して用い得る。デフォルトパラメータの設定例を以下に示す。

【0061】

Matrix: BLOSUM62

Open Gap: 11 及び Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

一致率は、ある定義されたポリペプチド配列の全長（例えば特定のSEQ IDナンバーで定義された配列）で測定し得る。あるいは、より短い長さ、例えば、定義された、より大きなポリペプチド配列から得た断片（例えば少なくとも15、20、30、40、50、70、または少なくとも150の連続した残基の断片）の長さの一致率も測定し得る。ここに挙げた長さは単なる例示的なものに過ぎず、表、図および配列リストを含めた本明細書に記載された配列が支持する任意の断片長を用いて、一致率を測定し得る或る長さを説明し得ることを理解されたい。

10

【0062】

「ヒト人工染色体（HAC）」は直鎖状の微小染色体であり、約6 kb ~ 10 MbのサイズのDNA配列を持ち得る。また、染色体の複製、分離および維持に必要な、全てのエレメントを持つ。

20

【0063】

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を改変した抗体分子を指す。

【0064】

「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、或る一本鎖ポリヌクレオチドが或る相補的な一本鎖と塩基対を形成するアニーリングのプロセスである。特異的ハイブリダイゼーションは、2つの核酸配列が高い相補性を共有することの指標である。特異的ハイブリダイゼーション複合体は許容されるアニーリング条件下で形成され、1回以上の「洗浄」ステップ後もハイブリダイズされたままである。洗浄ステップは、ハイブリダイゼーションプロセスのストリンジェンシーを決定する際に特に重要であり、よりストリンジェントな条件では、非特異結合（すなわち完全には一致しない核酸鎖対間の結合）が減少する。核酸配列のアニーリングに関する許容条件は、当業者が慣例的に決定できる。アニール条件はどのハイブリダイゼーション実験でも一定であり得るが、洗浄条件は、所望のストリンジェンシー、したがってハイブリダイゼーション特異性を得るように、実験ごとに変更し得る。許容されるアニーリング条件は、例えば、温度が68℃で、約6×SSC、約1%（w/v）のSDS、並びに約100 μg/mlの、せん断して変性したサケ精子DNAの存在下である。

30

【0065】

一般に、ハイブリダイゼーションのストリンジェンシーは或る程度、洗浄ステップを実行する温度を基準にして表すことができる。このような洗浄温度は通常、所定のイオン強度およびpHにおける特定配列の融点（T_m）より約5~20℃低くなるように選択する。このT_mは、所定のイオン強度およびpHの条件下で、完全一致プローブに標的配列の50%がハイブリダイズする温度である。T_mを計算する式および核酸のハイブリダイゼーション条件はよく知られており、Sambrook, J. 他（1989）Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NYに記載されており、特に2巻の9章を参照されたい。

40

【0066】

本発明のポリヌクレオチド間の高ストリンジェンシー条件のハイブリダイゼーションには、約0.2×SSCおよび約0.1%のSDSの存在下、68℃で1時間の洗浄条件を

50

含む。別法では、温度は約 65℃、60℃、55℃、または 42℃を用い得る。SSC 濃度は、約 0.1% の SDS 存在下で、約 0.1～2×SSC の範囲で変更し得る。通常は、ブロッキング剤を用いて非特異ハイブリダイゼーションを阻止する。このようなブロッキング剤には、例えば、約 100～200 μg/ml の、せん断した変性サケ精子 DNA がある。特定条件下で、例えば RNA と DNA のハイブリダイゼーションでは、有機溶剤、例えば約 35～50% v/v の濃度のホルムアミドを用いることもできる。洗浄条件の有用なバリエーションは、当業者には自明であろう。ハイブリダイゼーションは、特に高ストリンジェント条件下では、ヌクレオチド間の進化的な類似性を示唆し得る。進化的類似性は、ヌクレオチド群、およびヌクレオチドがコードするポリペプチド群について、或る同様の役割を強く示唆する。

10

【0067】

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

【0068】

用語「挿入」あるいは「付加」は、1個以上のアミノ酸残基あるいはヌクレオチドがそれぞれ追加される、アミノ酸配列あるいは核酸配列における変化を指す。

20

【0069】

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

【0070】

用語「免疫原性断片」は、例えば哺乳動物など生物に導入すると免疫応答を引き起こし得る、MDDT のポリペプチド断片またはオリゴペプチド断片を指す。用語「免疫原性断片」にはまた、本明細書で開示するまたは当分野で既知のあらゆる抗体生産方法に有用な、MDDT のあらゆるポリペプチド断片またはオリゴペプチド断片をも含む。

30

【0071】

用語「マイクロアレイ」は、或る基板上的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。

【0072】

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。

【0073】

用語「モジュレート」または「活性を調節」は、MDDT の活性を変化させることを指す。例えば、モジュレートによって、MDDT のタンパク質活性の増減が、或いは結合特性の、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き得る。

40

【0074】

「核酸」および「核酸配列」の語は、ヌクレオチド、オリゴヌクレオチド、ポリヌクレオチドまたはこれらの断片を指す。「核酸」および「核酸配列」の語はまた、ゲノム起源または合成起源の DNA または RNA であって一本鎖または二本鎖であるかあるいはセンス鎖またはアンチセンス鎖を表し得るような DNA または RNA や、ペプチド核酸 (PNA) や、任意の DNA 様または RNA 様物質を指すこともある。

【0075】

「機能的に連結した」は、第1の核酸配列と第2の核酸配列が機能的な関係にある状態を指す。例えば、或るプロモーターが或るコード配列の転写または発現に影響を及ぼす場合には、そのプロモーターはそのコード配列に機能的に連結している。機能的に連結した D

50

N A 配列群は非常に近接するか連続的に隣接することがあり、また、2つのタンパク質コード領域を結合するために必要な場合は同一リーディングフレーム内にあり得る。

【0076】

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

【0077】

MDDTの「翻訳後修飾」としては、脂質化、グリコシル化、リン酸化、アセチル化、ラセミ化、蛋白分解的切断及びその他の当分野で既知の修飾を含み得る。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、MDDTの酵素環境に依存し、細胞の種類によって異なることとなる。

【0078】

「プローブ」とは、核酸配列の内、MDDTやそれらの相補配列、またはそれらの断片をコードし、同一配列や対立遺伝子核酸配列、または関連する核酸配列の検出に用いる配列を指す。プローブは、単離されたオリゴヌクレオチドまたはポリヌクレオチドであって、検出可能な標識またはレポーター分子に結合した配列である。典型的な標識には、放射性アイソトープ、リガンド、化学発光試薬および酵素がある。「プライマー」とは、相補的な塩基対を形成して標的ポリヌクレオチドにアニーリング可能な、通常はDNAオリゴヌクレオチドである短い核酸である。プライマーは次に、DNAポリメラーゼ酵素により、標的DNA鎖に沿って伸長され得る。プライマー対は、例えばポリメラーゼ連鎖反応 (PCR) による、核酸配列の増幅 (および同定) に用い得る。

【0079】

本発明に用いるプローブおよびプライマーは通常、既知の配列の、少なくとも15の連続したヌクレオチド群からなる。特異性を高めるため、長めのプローブおよびプライマー、例えば開示した核酸配列の少なくとも20、25、30、40、50、60、70、80、90、100または少なくとも150の連続したヌクレオチドからなるようなプローブおよびプライマーも用い得る。これよりもかなり長いプローブおよびプライマーもある。表、図面および配列リストを含む本明細書が支持する、任意の長さのヌクレオチドを用い得るものと理解されたい。

【0080】

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

【0081】

プライマーとして用いるオリゴヌクレオチドの選択は、そのような目的のために本技術分野でよく知られているソフトウェアを用いて行う。例えばOLIGO 4.06ソフトウェアは、各100ヌクレオチドまでのPCRプライマー対の選択に有用であり、オリゴヌクレオチドおよび、最大5,000までの大きめのポリヌクレオチドであって32キロベースまでのインプットポリヌクレオチド配列から得た配列を分析するのにも有用である。

類似のプライマー選択プログラムには、拡張能力のための追加機能が組込まれている。例えば、PrimOUプライマー選択プログラム（テキサス州ダラスにあるテキサス大学南西部医療センターのゲノムセンターから一般向けに入手可能）は、メガベース配列から特定のプライマーを選択することが可能であり、したがってゲノム全体の範囲でプライマーを設計するのに有用である。Primer3プライマー選択プログラム（Whitehead Institute/MIT Center for Genome Research（マサチューセッツ州ケンブリッジ）より入手可能）を用いれば、プライマー結合部位として避けたい配列を指定できる「非プライミングライブラリ（mispriming library）」を入力できる。Primer3は特に、マイクロアレイのためのオリゴヌクレオチドの選択に有用である（後二者のプライマー選択プログラムのソースコードは、それぞれの情報源から得てユーザー固有のニーズを満たすように修正し得る）。PrimerGenプログラム（英国ケンブリッジ市の英国ヒトゲノムマッピングプロジェクトリソースセンターから一般向けに入手可能）は、多数の配列アラインメントに基づいてプライマーを設計し、それによって、アラインメントされた核酸配列の最大保存領域または最小保存領域のいずれかとハイブリダイズするようなプライマーの選択を可能にする。したがって、このプログラムは、独自のものであれ保存されたものであれ、オリゴヌクレオチドとポリヌクレオチド断片との同定に有用である。上記選択方法のいずれかによって同定したオリゴヌクレオチドおよびポリヌクレオチド断片は、ハイブリダイゼーション技術において、例えばPCRまたはシーケンシングプライマーとして、マイクロアレイエレメントとして、あるいは核酸のサンプルにおいて完全または部分的相補的ポリヌクレオチドを同定する特異プローブとして有用である。オリゴヌクレオチドの選択方法は、上記の方法に限定されるものではない。

10

20

30

【0082】

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

【0083】

あるいはこのような組換え核酸は、ウイルスベクターの一部と成すことができ、ベクターは例えばワクシニアウイルスに基づくものであり得る。そのようなベクターは哺乳類に接種され、その組換え核酸が発現されて、その哺乳類内で防御免疫応答を誘導するように使用することができる。

【0084】

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

40

【0085】

「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る、化学的または生化学的成分である。レポーター分子には、放射性核種、酵素、蛍光剤、化学発光剤、発色剤、基質、補助因子、阻害因子、磁気粒子およびその他の当分野で既知の成分がある。

【0086】

或るDNA配列に対する「RNA等価物」は、基準となるDNA配列と同じ直鎖のヌクレオチド配列からなるが、生じる全ての窒素性塩基のチミンがウラシルに置換され、糖鎖のバックボーンがデオキシリボースではなくリボースからなる。

【0087】

50

用語「サンプル」は、その最も広い意味で用いられている。M D D T、それをコードする核酸群、またはその断片群を含むと推定されるサンプルとしては、体液と、細胞や細胞から単離した染色体や細胞内小器官（オルガネラ）や膜からの抽出物と、細胞と、溶液中に存在するまたは基板に固定されたゲノム D N A、R N A、c D N A と、組織と、組織プリントなどがあり得る。

【0088】

用語「特異結合」および「特異的に結合する」は、タンパク質若しくはペプチドの、アゴニスト、抗体、アンタゴニスト、小分子、若しくは任意の天然若しくは合成の結合組成物との間の相互作用を指す。この相互作用は、タンパク質の特定の構造（例えば抗原決定基すなわちエピトープ）であって結合分子が認識する構造の有無に依存する。例えば、或る抗体がエピトープ「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」と抗体とを含む或る反応の中に、エピトープ A を持つポリペプチドが、ある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。

10

【0089】

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、単離あるいは分離された核酸配列ある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% 以上除去、最も好ましくは 90% 以上除去されたものを指す。

【0090】

「置換」とは、1 つ以上のアミノ酸残基またはヌクレオチドを、それぞれ別のアミノ酸残基またはヌクレオチドに置き換えることである。

20

【0091】

用語「基板」は、任意の好適な固体あるいは半固体の支持物を指し、膜およびフィルタ、チップ、スライド、ウェハ、ファイバー、磁気または非磁気ビーズ、ゲル、チューブ、プレート、ポリマー、微小粒子、毛細管が含まれる。基板は、ウェル、溝、ピン、チャネル、孔など、様々な表面形態を有することができ、基板表面にはポリヌクレオチドやポリペプチドが結合する。

【0092】

「転写イメージ」または「発現プロフィール」は、所定条件下での所定時間における、或る特定の細胞タイプまたは組織による、遺伝子発現の集合的パターンを指す。

30

【0093】

「形質転換」とは、外来性 D N A が、或る受容細胞に導入されるプロセスを言う。形質転換は、本技術分野で知られている種々の方法に従って自然条件または人工条件下で生じ得るものであり、外来性の核酸配列を原核宿主細胞または真核宿主細胞に挿入する、任意の既知の方法を基にし得る。形質転換の方法は、形質転換する宿主細胞の種類によって選択する。限定するものではないが形質転換方法には、ウイルス感染、電気穿孔法（エレクトロポレーション）、熱ショック、リポフェクションおよび微粒子銃を用いる方法がある。

「形質転換された細胞」には、導入された D N A が、自律的に複製するプラスミドとしてある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。更に、限られた期間に一過的に導入 D N A 若しくは導入 R N A を発現する細胞も含まれる。

40

【0094】

ここで用いる「遺伝形質転換体」とは任意の生物体であり、限定するものではないが動植物を含み、生物体の 1 個以上の細胞が、ヒトの介入によって、例えば本技術分野で公知の形質転換技術によって導入された異種核酸を有するものである。細胞への核酸の導入は、直接または間接的に、細胞の前駆物質に導入することによって行う。これは、計画的な遺伝子操作によって、例えば微量注射法によってあるいは組換えウイルスの導入によって行う。遺伝子操作の語は、古典的な交雑育種あるいは *i n v i t r o* 受精を指すものではなく、組換え D N A 分子の導入を指す。本発明に基づいて予期される遺伝形質転換体には、バクテリア、シアノバクテリア、真菌および動植物がある。本発明の単離された D N A

50

は、本技術分野で知られている方法、例えば感染、形質移入、形質転換またはトランス接合によって宿主に導入することができる。本発明のDNAをこのような生物体に移入する技術は公知であり、前出のSambrookら（1989）等の参考文献に記載がある。

【0095】

特定の核酸配列の「変異体／変異配列」とは、核酸配列1本の或る長さ全体について、該特定核酸配列に対し少なくとも40%の配列同一性を有する核酸配列として決定された配列である。決定には、デフォルトパラメータに設定した「BLAST 2 Sequences」ツールVersion 2.0.9（1999年5月7日）を用いてblastnを実行する。このような核酸対は、所定の長さに対して、例えば少なくとも50%、60%、70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の配列同一性を示し得る。変異配列は、例えば、「対立遺伝子」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」変異配列と記述され得る。スプライス変異配列は参照分子との顕著な同一性を有し得るが、mRNAプロセッシング中のエキソン群の選択的スプライシングによって通常、より多数またはより少数のポリヌクレオチドを有することになる。対応するポリペプチドは、追加機能ドメイン群を有するか、あるいは参照分子には存在するドメイン群が欠落していることがある。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。結果的に生じるポリペプチドは通常、相互に有意のアミノ酸同一性を持つ。多型性変異配列は、或る種の個体間における、或る特定遺伝子のポリヌクレオチド配列内での変異である。多型変異配列にはまた、ポリヌクレオチド配列の1つのヌクレオチド塩基が異なる「1塩基多型性」（SNP）も含み得る。SNPの存在は、例えば特定の個体群、病状または病状性向を示し得る。

【0096】

特定のポリペプチド配列の「変異配列」とは、ポリペプチド配列1本の或る長さ全体について、該特定ポリペプチド配列に対し少なくとも40%の配列同一性を有するポリペプチド配列として決定された配列である。決定には、デフォルトパラメータに設定した「BLAST 2 Sequences」ツールVersion 2.0.9（1999年5月7日）を用いてblastpを実行する。このようなポリペプチド対は、そのポリペプチドの一方の所定の長さに対して、例えば少なくとも50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の配列同一性を示し得る。

【0097】

（発明）

本発明は、新規のヒト疾患検出および治療用分子群（MDDT）および、MDDTをコードするポリヌクレオチド群の発見に基づく。また、これらの組成物を利用した、細胞増殖異常、自己免疫／炎症性障害、発達障害、および神経障害の、診断、治療、および予防の開発に基づく。

【0098】

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

【0099】

表2は、GenBankタンパク質（genpept）データベースに対するBLAST分析で同定した、本発明のポリペプチド群に相同な配列群を示す。列1および列2はそれ

ぞれ、本発明の各ポリペプチドに対するポリペプチド配列識別番号（ポリペプチド SEQ ID NO :）と、それに対応する Incyte ポリペプチド配列番号（Incyte ポリペプチド ID）を示す。列 3 は、最も近い GenBank 相同体の GenBank 識別番号（Genbank ID NO）を示す。列 4 は、各ポリペプチドとその相同体 1 つ以上との間の一致に関する確率スコアを示す。列 5 は、該当箇所には適当な引用を示すとともに GenBank 相同体 1 つ以上の注釈（annotation）を示し、これらはすべて本明細書では参考文献に含まれる。

【0100】

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

【0101】

表 2 および表 3 は共に、本発明の各々のポリペプチドの特性を要約しており、それら特性が、請求の範囲に記載されたポリペプチドが疾患検出および治療用分子であることを確立している。例えば SEQ ID NO : 11 は、大腸菌（*Escherichia coli*）の推定上のヌクレオチド結合タンパク質（GenBank ID g1789285）との 47 % の同一性を有することが Basic Local Alignment Search Tool（BLAST）によって示された（表 2 参照）。BLAST 確率スコアは 1.0×10^{-63} であり、これは観測されたポリペプチド配列を偶然に得る確率を示す。BLIMPS 解析および MOTIFS 解析よりのデータは、SEQ ID NO : 11 がヌクレオチド結合タンパク質であることを更に確証する証拠を提供する。別の例において SEQ ID NO : 15 は、Basic Local Alignment Search Tool（BLAST）による判定では、ウシの生体異物／中鎖脂肪酸 CoA リガーゼ型 XL-III（GenBank ID g5070357）との 57 % の同一性を持つ（表 2 参照）。BLAST 確率スコアは 6.7×10^{-181} であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO : 15 は、また、1 つの AMP 結合酵素ドメインを持つが、これは、隠れマルコフモデル（HMM）を基にした保存されたタンパク質ファミリードメインの PFAM データベースにおいて、統計的に有意な一致を検索して判定された。（表 3 参照）。BLIMPS 解析、MOTIFS 解析、および BLAST 解析よりのデータは、SEQ ID NO : 15 がアシル CoA リガーゼであり AMP 結合ドメインを持つことを示す、更に実証的な証拠を提供する。別の例として、SEQ ID NO : 23 は、残基 1 から 268 まだが、ヒトの NM23-H8（GenBank ID g7580490）との 40 % の同一性を有することが、Basic Local Alignment Search Tool（BLAST）によって示された（表 2 参照）。nm23 遺伝子群は、高度に転移性の細胞株群における、それらの発現の低下に基づいて発見された。様々な種からの Nm23 タンパク質は、ヌクレオシド二リン酸（NDP）キナーゼ活性を示し、組織の発生と分化とに関わる（Freije, J. M. 他（1998）Biochem. Soc. Symp. 63 : 261-271）。BLAST 確率スコアは 1.8×10^{-45} であり、これは観察されたポリペプチド配列アラインメントが偶然に得られる確率を示している。SEQ ID NO : 23 にはまた、1 つのチオレドキシシンファミリー活性部位と、1 つのヌクレオシド二リン酸キナーゼ活性部位とがあることが、PROFILERSCAN 分析で判定された（表 3 参照）。BLIMPS 解析および MOTIFS 解析よりのデータは、SEQ ID NO : 23 が nm23 ファミリー NDP キナーゼであることを更に確証する証拠を提供する。S

10

20

30

40

50

SEQ ID NO: 1-10、SEQ ID NO: 12-14、SEQ ID NO: 16-22、およびSEQ ID NO: 24-26については、同様の方法で分析し、注釈を付けた。SEQ ID NO: 1-26の解析用のアルゴリズム及びパラメータを表7に記載した。

【0102】

表4に示すように、本発明の完全長ポリヌクレオチド配列は、cDNA配列またはゲノムDNA由来のコード（エキソン）配列を用いて、あるいはこれら2種類の配列を任意に組み合わせて構築した。列1は本発明の各ポリヌクレオチドに対するポリヌクレオチド配列識別番号（ポリヌクレオチドSEQ ID NO）および対応するIncycyteポリヌクレオチドコンセンサス配列番号（Incycyte ID）、および塩基対の各ポリヌクレオチド配列の長さを示している。列2は、本発明の完全長ポリヌクレオチド配列を構築するのに用いたcDNA配列および／またはゲノム配列の、また、例えばSEQ ID NO: 27-52を同定するため、或いはSEQ ID NO: 27-52と、関連するポリヌクレオチド配列群とを区別するためのハイブリダイゼーション技術または増幅技術に有用なポリヌクレオチド配列の断片の、開始ヌクレオチド（5'）位置および終了ヌクレオチド（3'）位置を示す。

【0103】

表4の列2に記載したポリヌクレオチド断片は、特に例えば組織特異的cDNAライブラリ群やプールしたcDNAライブラリ群に由来するIncycyte cDNA群を指す場合もある。或いは列2に記載したポリヌクレオチド断片は、完全長ポリヌクレオチド配列の構築（アセンブリ）に寄与したGenBank cDNAまたはESTを指す場合もある。更に、列2に記載したポリヌクレオチド断片は、ENSEMBL（The Sanger Centre、英国ケンブリッジ）データベースに由来する配列を同定する場合もある（すなわち「ENST」の命名を含む配列）。あるいは、列2に記載したポリヌクレオチド断片は、NCBI RefSeq Nucleotide Sequence Records データベースに由来する場合もあり（すなわち「NM」または「NT」の命名を含む配列）、またNCBI RefSeq Protein Sequence Records に由来する場合もある（すなわち「NP」の命名を含む配列）。または列2に記したポリヌクレオチド断片は、「エキソンスティッチング（exon-stitching）」アルゴリズムにより結び合わせたcDNAおよびGenscan予測エキソン群の両方からなる集合を指す場合がある。例えば、ポリヌクレオチド配列の内、FL_XXXXX_X_N1_N2_YYYYY_N3_N4_として同定される配列は「スティッチされた」配列であり、その内、XXXXXXは該アルゴリズムが適用される配列のクラスターの識別番号であり、YYYYYは該アルゴリズムが作成する予測の数であり、N1, 2, 3, ... がある場合には、解析中に手動で編集された特定のエキソン群を表す（実施例5参照）。または、列2のポリヌクレオチド断片は「エキソンストレッチング（exon-stretching）」アルゴリズムにより結び合わせたエキソンの集合を指す場合もある。例えば、ポリヌクレオチド配列の内、FLXXXXXXXX_gAAAAA_gBBBBB_1_Nとして同定される配列は「ストレッチ」配列の識別番号である。ここでXXXXXXはIncycyteプロジェクト識別番号、gAAAAAは「エキソンストレッチング」アルゴリズムを適用したヒトゲノム配列のGenBank識別番号、gBBBBBは一番近いGenBankタンパク質相同体のGenBank識別番号、またはNCBI RefSeq識別番号、Nは特定のエキソンを指す（実施例5を参照）。或るRefSeq配列が「エキソンストレッチング」アルゴリズムのためのタンパク質相同体として使用された場合では、「NM」、「NP」、または「NT」によって表されるRefSeq識別子が、GenBank識別子（すなわち、gBBBBB）の代わりに使用され得る。

【0104】

あるいは、接頭コードは、手動で編集された構成配列、ゲノムDNA配列から予測された構成配列、または組み合わされた配列解析方法に由来する構成配列を同定する。次の表は、構成配列の接頭コードと、接頭コードに対応する配列分析方法の例を列記する（実施例

4 と 5 を参照)。

接頭コード	解析タイプやプログラムの例
GNN	例えば、GENSCAN (Stanford University, CA, USA) または
GFG	FGGENES (Computer Genomics Group, The Sanger Centre,
ENST	Cambridge, UK) を用いた、ゲノム配列群からのエキソン予測
GBI	手動で編集された、ゲノム配列群の解析
FL	スティッチまたはストレッチゲノム配列 (実施例5参照)
INCY	ゲノムへのEST配列群のマッピングからの、完全長転写物とエキソンとの予測。エキソン群と生じる転写物とを予測するために、ゲノム位置とEST構成のデータが組み合わされる。

10

20

30

【0105】

場合によっては、最終コンセンサスポリヌクレオチド配列を確認するために、表4に示すような配列カバレッジと重複する Incyte cDNA カバレッジが得られたが、該当する Incyte cDNA 識別番号は示さなかった。

【0106】

表5は、Incyte cDNA 配列を用いて構築された完全長ポリヌクレオチド配列のための代表的な cDNA ライブラリを示している。代表的な cDNA ライブラリとは Incyte cDNA ライブラリであり、これは、最も頻繁には Incyte cDNA 配列群によって代表されるが、これらは、上記のポリヌクレオチド配列を構築および確認するために用いられた。cDNA ライブラリを作製するために用いた組織およびベクターを表5に示し、表6で説明している。

40

【0107】

本発明には、また、MDDTの変異配列群をも含む。好適なMDDT変異体は、MDDTの機能的或いは構造的特徴を少なくとも1つ有し、かつ、MDDT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変異体である。

【0108】

50

本発明はまた、M D D Tをコードするポリヌクレオチドを提供する。特定の実施態様において、本発明は、M D D Tをコードする、S E Q I D N O : 2 7 - 5 2 からなる一群から選択された1配列を持つポリヌクレオチド配列を提供する。S E Q I D N O : 2 7 - 5 2 のポリヌクレオチド配列は、配列表に示されているように等価R N A配列をも含むが、そこでは窒素塩基チミンの出現はウラシルに置換され、糖のバックボーンはデオキシリボースではなくてリボースから構成されている。

【0109】

本発明はまた、M D D T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、このような変異体ポリヌクレオチド配列は、M D D Tをコードするポリヌクレオチド配列との、少なくとも約70%のポリヌクレオチド配列同一性、あるいは少なくとも約85%のポリヌクレオチド配列同一性、更には少なくとも約95%ものポリヌクレオチド配列同一性を持つこととなる。本発明の或る特定の実施形態は、S E Q I D N O : 2 7 - 5 2 からなる群から選択された1核酸配列との少なくとも約70%のポリヌクレオチド配列同一性、或いは少なくとも約85%のポリヌクレオチド配列同一性、更には少なくとも約95%ものポリヌクレオチド配列同一性を有する、S E Q I D N O : 2 7 - 5 2 からなる群から選択された1配列を持つポリヌクレオチド配列の変異配列を提供する。上記のポリヌクレオチド変異配列は何れも、D M Eの機能的或いは構造的特徴の少なくとも1つを有する或るアミノ酸配列をコードし得る。

【0110】

更に別の例では、本発明の或るポリヌクレオチド変異配列は、M D D Tをコードするポリヌクレオチド配列のスプライス変異配列である。或るスプライス変異配列はM D D Tをコードするポリヌクレオチド配列との顕著な配列同一性を持つ部分複数を有し得るが、m R N Aプロセッシング中のエキソン群の選択的スプライシングによって生ずる、配列の数ブロックの付加または欠失により、通常、より多数またはより少数のポリヌクレオチドを有することになる。或るスプライス変異配列には、約70%未満、または約60%未満、あるいは約50%未満のポリヌクレオチド配列同一性が、M D D Tをコードするポリヌクレオチド配列との間で全長に渡って見られるが、このスプライス変異配列のいくつかの部分には、M D D Tをコードするポリヌクレオチド配列の各部との、少なくとも約70%、あるいは少なくとも約85%、または少なくとも約95%、なおまたは100%の、ポリヌクレオチド配列同一性を有することとなる。例えばS E Q I D N O : 5 2 は、S E Q I D N O : 3 5 の配列からなるポリヌクレオチドのスプライス変異体である。上記したスプライス変異配列は何れも、M D D Tの機能的或いは構造的特徴の少なくとも1つを有する或るアミノ酸配列をコードし得る。

【0111】

遺伝暗号の縮重により、M D D Tをコードする種々のポリヌクレオチド配列が作り出され、中には、既知のいかなる天然遺伝子のポリヌクレオチド配列群とも最小の類似性しか有しない配列もあることは、当業者には理解されよう。したがって本発明には、可能コドン選択に基づく組合せの選択によって産出し得るあらゆる可能なポリヌクレオチド配列のバリエーションを網羅し得る。これらの組合せは、天然M D D Tのポリヌクレオチド配列に適用される標準トリプレット遺伝暗号を基に成されるものであり、このような変異は全て明確に開示されていると考慮されたい。

【0112】

M D D Tとその変異配列とをコードするヌクレオチド配列は一般に、好適に選択されたストリンジェンシー条件下で天然M D D Tのヌクレオチド配列とハイブリダイズ可能であるが、非天然コドン群を含めるなどの実質的に異なるコドン使用を有する、M D D T或いはその誘導体をコードするヌクレオチド配列を作り出すことは、有益であり得る。宿主が特定コドンを利用する頻度に基づいて、特定の真核宿主または原核宿主に発生するペプチドの発現率を高めるようにコドンを選択し得る。コードされるアミノ酸配列を改変せずに、M D D Tおよびその誘導体をコードするヌクレオチド配列を実質的に改変する別の理由には、天然配列から作られる転写物より例えば長い半減期など好ましい特性を備えるR N A

10

20

30

40

50

転写物を作ることもある。

【0113】

本発明にはまた、MDDTとその誘導体とをコードするDNA配列またはそれらの断片の、完全に合成化学による作製も含む。作製後にこの合成配列を、当分野で周知の試薬類を用いて、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できる。更に、合成化学を用いて、MDDTをコード、またはその任意の断片をコードする或る配列の中に突然変異を導入できる。

【0114】

更に本発明には、種々のストリンジェンシー条件下で、請求項に記載されたポリヌクレオチド配列、特に、SEQ ID NO: 27-52及びそれらの断片群にハイブリダイズ可能なポリヌクレオチド配列群が含まれる（例えば、Wahl, G. M. および S. L. Berger (1987) *Methods Enzymol.* 152: 399-407; Kimmel, A. R. (1987) *Methods Enzymol.* 152: 507-511を参照）。アニーリングおよ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」に記載した。

【0115】

DNAシーケンシングの方法は当分野では公知であり、本発明のいずれの実施例も、DNAシーケンシング方法を用い得る。DNAシーケンシング方法には酵素を用い得る。例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amersham, Pharmacia Biotech, Piscataway NJ)を用い得る。あるいは、例えばELONGASE増幅システム (Life Technologies, Gaithersburg MD)において見られるように、ポリメラーゼと校正エキソヌクレアーゼとを併用し得る。好適には、MICROLAB 2200液体転移システム (Hamilton, Reno, NV)、PTC 200サーマルサイクラー (MJ Research, Watertown MA) およびABI CATALYST 800サーマルサイクラー (Applied 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あるいは377 DNAシーケンシングシステム (Applied Biosystems)、MEGABACE 1000 DNAシーケンシングシステム (Molecular Dynamics, Sunnyvale CA)または当分野で公知の他のシステムを用いてシーケンシングを行う。結果として得た配列を、当分野で周知の種々のアルゴリズムを用いて分析する（例えば、Ausubel, F. M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, 7.7ユニット、Meyers, R. A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, 856-853ページを参照）。

【0116】

当分野で周知の、PCR法をベースにした種々の方法と、部分的ヌクレオチド配列とを利用して、MDDT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、上流にある配列を検出し得る。例えば、使用し得る方法の1つである制限部位PCR法は、ユニバーサルプライマーおよびネステッドプライマーを用いてクローニングベクター内のゲノムDNAからの未知の配列を増幅する方法である（例えば、Sarkar, G. (1993) *PCR Methods Applic.* 2: 318-322を参照）。2: 318322.）別の方法に逆PCR法があり、これは多岐の方向に伸長するプライマー群を用いて、環状化した鋳型から未知の配列を増幅する方法である。鋳型は、或る既知のゲノム遺伝子座およびその周辺の配列群からなる制限酵素断片群から得る（例えばTriglia, T. 他 (1988) *Nucleic Acids Res.* 16: 8186を参照）。第3の方法としてキャプチャPCR法があり、これにはヒトおよび酵母人

工染色体DNAの既知の配列群に隣接するDNA断片群をPCR増幅する方法を含む（例えば Lagerstrom, M. 他（1991）PCR Methods Appli c 1: 111-119を参照）。この方法では、PCRを行う前に、複数の制限酵素の消化およびライゲーション反応を用い、未知の配列領域内に組換え二本鎖配列を挿入し得る。未知の配列群を検索するために用い得る他の複数の方法も当分野で公知である（例えば Parker, J. D. 他（1991）Nucleic Acids Res. 19: 3055-3060を参照）。（1991）Nucleic Acids Res. 30553060を参照）。更に、PCR、ネステッドプライマー類、およびPROMOTERFINDERライブラリ類（Clontech, Palo Alto CA）を用いてゲノムDNAをウォーキングし得る。この手順は、ライブラリ類をスクリーニングする必要がなく、イントロン／エキソン接合部の発見に有用である。全てのPCRベースの方法に対して、市販されているソフトウェア、例えばOLIGO 4.06プライマー分析ソフトウェア（National Biosciences, Plymouth MN）あるいは別の好適なプログラムを用いて、長さが約22～30ヌクレオチド、GC含有率が約50%以上、温度約68℃～72℃で鋳型に対してアニーリングするようにプライマーを設計し得る。

【0117】

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

【0118】

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

【0119】

本発明の別の実施例では、MDDT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MDDT、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配列を産生し、これらの配列をMDDTの発現に利用し得る。

【0120】

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

【0121】

本発明のヌクレオチドは、MOLECULAR BREEDING (Maxygen Inc 50

c., 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シャッフリング技術の対象となり得る。これにより、生物活性または酵素活性、あるいは他の分子や化合物と結合する能力など、MDDTの生物学的特性を改変あるいは改良し得る。DNAシャッフリングは、PCRを介する遺伝子断片組換えを用いて遺伝子変異配列のライブラリを作製するプロセスである。ライブラリはその後、所望の特性を持つ遺伝子変異配列群を同定する、選択またはスクリーニングの手順を経る。続いて、これら好適な変異配列をプールし、更に反復してDNAシャッフリングおよび選択／スクリーニングを行い得る。かくして、「人工的な」育種および急速な分子の進化によって、多様な遺伝子が作られる。例えば、ランダムポイント突然変異を持つ単一の遺伝子の断片を組み換えて、スクリーニングし、その後所望の特性が最適化されるまでシャッフリングし得る。あるいは、所定の遺伝子の断片群を同種または異種のいずれかから得た同一遺伝子ファミリーの相同遺伝子の断片と組み換え、それによって天然に存在する複数の遺伝子の遺伝多様性を、方向付けした制御可能な方法で最大化させることができる。

10

【0122】

別の実施態様によれば、MDDT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、全体或いは一部を合成可能である（例えば、Caruthers, M. H. 他 (1980) Nucleic Acids Symp. Ser. 7:215-223; および Horn, T. 他 (1980) Nucleic Acids Symp. Ser. 7:225-232を参照）。別法として、化学的方法を用いてMDDT自体またはその断片を合成することが可能である。例えば、種々の液相または固相技術を用いてペプチド合成を行い得る（例えばCreighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, 55-60ページ; および Roberge, J. Y. 他 (1995) Science 269:202-204を参照）。自動合成はABI 431Aペプチドシンセサイザ (Applied Biosystems) を用いて達成し得る。更に、MDDTのアミノ酸配列、または任意のその一部を、直接合成の際に改変することにより、及び／または他のタンパク質からの配列群または任意のその一部と組み合わせることにより、変異体ポリペプチドを作製、または、或る天然ポリペプチドの1配列を有するポリペプチドを作製し得る。

20

30

【0123】

このペプチドは、分離用高速液体クロマトグラフィを用いて実質的に精製し得る（例えばChiez, R. M. およびF. Z. Regnier (1990) Methods Enzymol. 182:392-421を参照）。この合成ペプチドの組成は、アミノ酸分析またはシーケンシングによって確認できる（例えば前出のCreighton, 28-53ページを参照）。

【0124】

生物学的に活性なMDDTを発現させるために、MDDTをコードするヌクレオチド配列またはそれらの誘導体を好適な発現ベクターに挿入することができる。好適な発現ベクターとは、好適な宿主に挿入されたコーディング配列の転写と翻訳との制御に必要なエレメント群を持つベクターである。必要なエレメントの例には、ベクターに、またMDDTをコードするポリヌクレオチド配列における調節配列、例えばエンハンサー、構成型および発現誘導型のプロモーター、5'および3'の非翻訳領域などが含まれる。必要なエレメント群は、特長および特異性が様々である。特定の開始シグナル類によって、MDDTをコードする配列の、より効果的な翻訳を達成することも可能である。開始シグナルの例には、ATG開始コドンと、コザック配列など近傍の配列とが含まれる。MDDTをコードする配列、およびその開始コドンや、上流の調節配列が好適な発現ベクターに挿入された

40

50

場合は、更なる転写制御シグナルや翻訳制御シグナルは必要ないこともある。しかしながら、コーディング配列あるいはその断片のみが挿入された場合は、インフレイムATG開始コドンなど外来性の翻訳制御シグナルが発現ベクターに含まれるようにすべきである。外来性の翻訳エレメント群および開始コドン群は、様々な天然物および合成物を起源とし得る。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である（例えばScharf, D. 他（1994）Results Probl. Cell Differ. 20:125162を参照）。

【0125】

当業者に周知の方法を用いて、MDDTをコードする配列と、好適な転写および翻訳制御エレメントとを持つ発現ベクターを作製することが可能である。これらの方法としては、in vitro組換えDNA技術、合成技術、およびin vivo遺伝子組換え技術がある（例えばSambrook, J. 他（1989）Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, 4, 8, および16-17章; Ausubel, F. M. 他（1995）Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, 9, 13, および16章を参照）。

10

【0126】

種々の発現ベクター／宿主系を利用して、MDDTをコードする配列の保持および発現が可能である。限定するものではないがこのような発現ベクター／宿主系には、組換えバクテリオファージ、プラスミドまたはコスミドDNA発現ベクターで形質転換させた細菌や、酵母菌発現ベクターで形質転換させた酵母菌など微生物や、ウイルス発現ベクター（例えばバキュロウイルス）に感染した昆虫細胞系や、ウイルス発現ベクター（例えばカリフラワーモザイクウイルス、CaMVまたはタバコモザイクウイルス、TMV）または細菌発現ベクター（例えばTiプラスミドまたはpBR322プラスミド）で形質転換させた植物細胞系、動物細胞系がある（例えば前出Sambrook; 前出Ausubel; Van Heeke, G. および S. M. Schuster（1989）J. Biol. Chem. 264:55035509; Engelhard, E. K. 他（1994）Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. 他（1996）Hum. Gene Ther. 7:1937-1945; Takamatsu, N.（1987）EMBO J. 6:307311; 『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, 191196ページ; Logan, J. および T. Shenk（1984）Proc. Natl. Acad. Sci. USA 81:36553659; および Harrington, J. J. 他（1997）Nat. Genet. 15:345-355を参照）。レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ送達することができる（例えば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; 並びに Verma, I. M. および N. Somia（1997）Nature 389:239-242を参照）。本発明は使用される宿主細胞によって限定されるものではない。

20

30

40

【0127】

細菌系では、MDDT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多数のクローニングベクターおよび発現ベクターを選択し得る。例えば、MDDTをコードするポリ

50

ヌクレオチド配列の慣例的なクローニング、サブクローニング、増殖には、P B L U E S C R I P T (S t r a t a g e n e, L a J o l l a C A) または P S P O R T 1 プラスミド (L i f e T e c h n o l o g i e s) などの多機能の大腸菌ベクターを用いることができる。ベクターのマルチクローニング部位に、M D D T をコードする配列をライゲーションすると l a c Z 遺伝子が破壊され、組換え分子を持つ形質転換された細菌の同定のための比色スクリーニング法が可能となる。更にこれらのベクターは、クローニングされた配列における i n v i t r o 転写、ジデオキシのシークエンシング、ヘルパーファージによる一本鎖のレスキュー、入れ子欠失の生成にも有用であろう (例えば、V a n H e e k e, G. および S. M. S c h u s t e r (1989) J. B i o l. C h e m. 264:55035509 を参照)。例えば抗体類の産生などに多量の M D D T が必要な場合は、M D D T の発現をハイレベルで指示するベクター類が使用できる。例えば、強力な誘導 S P 6 バクテリオファージプロモーターまたは誘導 T 7 バクテリオファージプロモーターを持つベクターが使用できる。

【0128】

酵母の発現系を使用して M D D T を産出することもできる。α 因子、アルコールオキシダーゼ、P G H プロモーターなど、構成型あるいは誘導型のプロモーターを持つ多数のベクターが、出芽酵母菌 (S a c c h a r o m y c e s c e r e v i s i a e) またはピキア酵母 (P i c h i a p a s t o r i s) 内で使用可能である。更に、このようなベクターは、発現したタンパク質の、分泌か細胞内での保持かのどちらかを指示するものであり、安定した増殖のために宿主ゲノムの中に外来配列群を組み込むことを可能にする (例えば前出の A u s u b e l, 1995; B i t t e r, G. A. 他 (1987) M e t h o d s E n z y m o l. 153:516544; および S c o r e r, C. A. 他 (1994) B i o / T e c h n o l o g y 12:181-184 を参照)。

【0129】

植物系も M D D T の発現に使用可能である。M D D T をコードする配列の転写は、ウイルスプロモーター、例えば単独或いは T M V (T a k a m a t s u, N. (1987) E M B O J 6:307-311) 由来のオメガリーダー配列と組み合わせて用いられるような、C a M V 由来の 35S および 19S プロモーターによって駆動し得る。あるいは、R u B i s C O の小サブユニットなどの植物プロモーター、または熱ショックプロモーターを用い得る (例えば C o r u z z i, G. 他 (1984) E M B O J. 3:16711680; B r o g l i e, R. 他 (1984) S c i e n c e 224:838843; および W i n t e r, J. 他 (1991) R e s u l t s P r o b l. C e l l D i f f e r. 17:85105 を参照)。これらの構成物は、直接 DNA 形質転換により、または病原体を媒介とする形質移入により、植物細胞内に導入し得る (例えば『マグローヒル科学技術年鑑』(The M c G r a w H i l l Y e a r b o o k o f S c i e n c e a n d T e c h n o l o g y) (1992) M c G r a w H i l l N e w Y o r k N Y, 191-196 ページを参照)。

【0130】

哺乳類細胞においては、多数のウイルスベースの発現系を利用し得る。アデノウイルスを発現ベクターとして用いる場合、後期プロモーターと 3 連リーダー配列とを持つアデノウイルス転写／翻訳複合体に、M D D T をコードする配列を結合し得る。アデノウイルスゲノムの非必須 E 1 または E 3 領域へ挿入することにより、宿主細胞内で M D D T を発現する感染ウイルスを得ることができる (例えば L o g a n, J. および T. S h e n k (1984) P r o c. N a t l. A c a d. S c i. U S A 81:36553659 を参照)。更に、ラウス肉腫ウイルス (R S V) エンハンサーなどの転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させ得る。S V 40 または E B V をベースにしたベクターを用いてタンパク質を高レベルで発現させることもできる。

【0131】

また、ヒト人工染色体 (H A C) 類を用いて、或るプラスミドに含まれ得る断片やプラスミドから発現し得る断片より大きな DNA の断片群を送達し得る。治療のために約 6 k b

～10MbのHACsが作製され、従来の送達方法（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）で送達されている（例えばHarrington, J. J. 他（1997）Nat. Genet. 15:345-355を参照）。

【0132】

哺乳類系の組換えタンパク質の長期にわたる産生のためには、細胞株におけるMDDTの安定した発現が望ましい。例えば、MDDTをコードする配列を細胞株に形質転換するために、発現ベクター類と、同じ或いは別のベクター上の選択可能マーカー遺伝子とを用い得る。用いる発現ベクターは、複製のウイルス起源、および／または内因性の発現エレメント群を持ち得る。ベクターの導入後、選択培地に移す前に強化培地で約1～2日間、細胞を増殖させ得る。選択可能マーカーの目的は選択剤への抵抗性を与えることであり、選択可能マーカーの存在により、導入した配列をうまく発現するような細胞の成長および回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に適した組織培養技術を用いて増殖可能である。

10

【0133】

任意の数の選択系を用いて、形質転換細胞株を回収できる。限定するものではないがこのような選択系には、tk⁻細胞のために用いられる単純ヘルペスウイルスのチミジンキナーゼ遺伝子と、ap^r⁻細胞のために用いられる単純ヘルペスウイルスのアデニンホスホリボシルトランスフェラーゼ遺伝子がある（例えばWigler, M. 他（1977）Cell 11:223-232; Lowy, I. 他（1980）Cell 22:817-823を参照）。また、選択の基礎として代謝拮抗物質、抗生物質あるいは除草剤への耐性を用いることができる。例えばdhfrはメトトレキセートに対する耐性を与え、neoはアミノグリコシッドネオマイシンおよびG-418に対する耐性を与え、alsはクロルスルフロンに対する耐性を、patはホスフィノトリシンアセチルトランスフェラーゼに対する耐性を各々与える（例えばWigler, M. 他（1980）Proc. Natl. Acad. Sci. USA 77:3567-3570; ColbereGarapin, F. 他（1981）J. Mol. Biol. 150:114を参照）。この他の選択可能な遺伝子、例えば、代謝のための、細胞の必要条件を改変するtrpBおよびhisDも、文献に記載がある（例えばHartman, S. C. およびR. C. Mulligan（1988）Proc. Natl. Acad. Sci. USA 85:8047-8051を参照）。可視マーカー類、例えばアントシアニンや、緑色蛍光タンパク質（GFP; Clontech）、βグルクロニダーゼおよびその基質であるβグルクロニド、またはルシフェラーゼおよびその基質であるルシフェリンを用い得る。これらのマーカーを用いて、トランスフォーマントを同定するだけでなく、特定のベクター系に起因する一過性あるいは安定したタンパク質発現を定量し得る（例えばRhodes, C. A.（1995）Methods Mol. Biol. 55:121-131を参照）。

20

30

【0134】

マーカー遺伝子発現の有無によって目的の遺伝子の存在が示唆されても、その遺伝子の存在および発現の確認が必要な場合もある。例えば、MDDTをコードする配列が或るマーカー遺伝子配列の中に挿入された場合、MDDTをコードする配列群を有する形質転換された細胞群は、マーカー遺伝子機能の欠落により特定できる。または、1つのプロモーターの制御下で、或るマーカー遺伝子が、MDDTをコードする配列とタンデムに配置されることも可能である。誘導または選択に応答したマーカー遺伝子の発現は通常、タンデム遺伝子の発現も示す。

40

【0135】

一般に、MDDTをコードする核酸配列を含み且つMDDTを発現する宿主細胞は、当業者によく知られている種々の方法を用いて同定することが可能である。限定するものではないが当業者によく知られている方法には、DNA-DNA或いはDNA-RNAハイブリダイゼーションと、PCR増幅とがあり、また、核酸配列或いはタンパク質配列の検出、定量、或いはその両方を行うための、膜系、溶液ベース或いはチップベースの技術を含

50

む、タンパク質の生物学的検定法または免疫学的検定法もある。

【0136】

特異的ポリクローナル抗体または特異的モノクローナル抗体を用いてMDDTの発現の検出と計測とを行うための免疫学的方法は、当分野で既知である。このような技術の例としては、酵素結合イムノソルベントアッセイ（ELISA）、ラジオイムノアッセイ（RIA）、蛍光活性化細胞選別（FACS）などが挙げられる。MDDT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、競合結合試験を用いることもできる。これらのアッセイおよび他のアッセイは、当分野で周知である（例えばHampton, R. 他（1990）Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J. E. 他（1997）Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; および Pound, J. D. （1998）Immunochemical Protocols, Humana Press, Totowa NJを参照）。

【0137】

多岐にわたる標識方法および抱合方法が当業者に知られており、様々な核酸アッセイおよびアミノ酸アッセイにこれらの方法を用い得る。MDDT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、エンドラベリング（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、MDDTをコードする配列、またはその任意の断片を、mRNAプローブを生成するためのベクターにクローニングすることも可能である。このようなベクターは、当分野において知られており、市販もされており、T7、T3またはSP6などの好適なRNAポリメラーゼおよび標識されたヌクレオチドを加えて、in vitroでRNAプローブの合成に用いることができる。このような方法は、例えばAmersham Pharmacia Biotech、Promega（Madison WI）、U. S. Biochemicalなどから市販されている種々のキットを用いて実行することができる。検出を容易にするために用い得る好適なレポーター分子あるいは標識には、基質、補助因子、インヒビター、磁気粒子のほか、放射性核種、酵素、蛍光剤、化学発光剤、発色剤などがある。

【0138】

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

【0139】

更に、宿主細胞株の選択は、挿入した配列の発現をモジュレートする能力、または発現したタンパク質を所望の形に処理する能力によって行い得る。限定するものではないがこのようなポリペプチドの修飾には、アセチル化、カルボキシル化、グリコシル化、リン酸化、脂質化およびアシル化がある。タンパク質の「プレプロ」または「プロ」形を切断する翻訳後のプロセッシングを利用して、タンパク質の標的への誘導、折りたたみ、および／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、特定の細胞装置および特徴的な機序を持つ、種々の宿主細胞（例えばCHO、HeLa、MDCK、MEK293、WI38など）は、American Type Culture Collection（ATCC, Manassas, VA）から入手可能であり、外来タンパク質の正し

い修飾および処理を確実にするように選択し得る。

【0140】

本発明の別の実施例では、MDDTをコードする、天然の核酸配列、修飾された核酸配列、または組換えの核酸配列を、或る異種配列に結合させることができ、この結果、上記した任意の宿主系の、或る融合タンパク質の翻訳を生じる。例えば、市販されている抗体を用いて認識可能な異種部分を含むキメラMDDTタンパク質は、MDDT活性阻害剤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および異種ペプチド部分も、市販されている親和性マトリックスを用いて融合タンパク質の精製を促進し得る。限定されるものではないがこのような部分には、グルタチオンSトランスフェラーゼ（GST）、マルトース結合タンパク質（MBP）、チオレドキシン（Trx）、カルモジュリン結合ペプチド（CBP）、6-His、FLAG、c-myc、赤血球凝集素（HA）がある。GSTは固定化グルタチオン上で、MBPはマルトース上で、Trxはフェニルアルシンオキシド上で、CBPはカルモジュリン上で、そして6-Hisは金属キレート樹脂上で、同族の融合タンパク質の精製を可能にする。FLAG、c-mycおよび赤血球凝集素（HA）は、これらのエピトープ標識を特異的に認識する市販のモノクローナル抗体およ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精製を可能にする。また、MDDTをコードする配列と異種タンパク質配列との間にタンパク質分解切断部位を融合タンパク質が持つように遺伝子操作すると、MDDT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および精製方法は、前出のAusubel（1995）10章に記載されている。市販されている種々のキットを用いて融合タンパク質の発現および精製を促進することもできる。

【0141】

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

【0142】

本発明のMDDTまたはその断片を用いて、MDDTに特異結合する化合物をスクリーニングすることができる。少なくとも1つまたは複数の試験化合物を用いて、MDDT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。試験化合物としては、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。

【0143】

或る実施態様では、このように同定された化合物は、MDDTの天然リガンドに密接に関連し、例えばリガンドやその断片であり、または天然基質や、構造的または機能的な擬態物質（mimetic）、あるいは自然結合パートナーである（例えばColigan, J. E. 他（1991）Current Protocols in Immunology 1（2）：5章を参照）。同様に、この化合物は、MDDTが結合する天然受容体に、あるいは例えばリガンド結合部位など少なくともこの受容体の或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MDDT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、哺乳類、酵母、ショウジョウバエ、または大腸菌からの細胞が含まれる。MDDTを発現する細胞またはMDDTを含有する細胞膜分画を試験化合物と接触させて、MDDTまたはこの化合物のどちらかの結合、刺激、または活性の阻害を分析する。

【0144】

或るアッセイは、単に試験化合物をポリペプチドに実験的に結合させ、蛍光色素、放射性同位体、酵素抱合体またはその他の検出可能な標識によりその結合を検出することができる。例えば、このアッセイは、少なくとも1つの試験化合物を、溶液中の或いは固体支持

物に固定されたMDDTと混合するステップと、MDDT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および測定を行うことができる。更にこのアッセイは、無細胞再構成標本、化学ライブラリ、または、天然産物の混合物を用いて実施でき、試験化合物（群）は、溶液中で遊離させるか固体支持体に固定する。

【0145】

本発明のMDDTまたはその断片を用いて、MDDTの活性をモジュレートする化合物をスクリーニングすることが可能である。このような化合物としては、アゴニスト、アンタゴニスト、部分的アゴニストまたは逆アゴニストなどが含まれ得る。一実施例では、MDDTが少なくとも1つの試験化合物と結合する、MDDT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MDDTの活性を、試験化合物不在下でのMDDTの活性と比較する。試験化合物の存在下でのMDDTの活性の変化は、MDDTの活性をモジュレートする化合物の存在を示唆する。別法では、試験化合物を、MDDTの活性に適した条件下で、DMEを有する*in vitro*系すなわち無細胞系と混合してアッセイを実施する。これらアッセイの何れにおいても、MDDTの活性をモジュレートする試験化合物は間接的にモジュレートする場合があります、その際は試験化合物と直接接触する必要がない。少なくとも1つ、または複数の試験化合物をスクリーニングすることができる。

10

【0146】

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、MDDTまたはその哺乳動物相同体をコードするポリヌクレオチドを「ノックアウト」する。このような技術は当技術分野において周知であり、ヒト疾患動物モデルの作製に有用である（米国特許第5,175,383号および第5,767,337号などを参照）。例えば129/SvJ細胞系などのマウス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）などのマーカー遺伝子で破壊した、目的の遺伝子を持つベクターで形質転換する。このベクターは、相同組換えにより、宿主ゲノムの対応する領域に組み込まれる。別法では、Cre-loxP系を用いて相同組換えを行い、組織特異的または発生段階特異的に目的遺伝子をノックアウトする（Marth, J. D. (1996) Clin. Invest. 97:1999-2002; Wagner, K. U. 他(1997) Nucleic Acids Res. 25:4323-4330）。形質転換したES細胞を同定し、例えばC57BL/6マウス株などから採取したマウス細胞胚盤胞に微量注入する。これらの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特定し、これら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、潜在的治療薬や毒性薬物で検査することができる。

20

30

【0147】

MDDT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も可能である。ヒトES細胞は、内胚葉、中胚葉および外胚葉の細胞タイプを含む、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、造血系統および心筋細胞に分化する（Thomson, J. A. 他(1998) Science 282:1145-1147）。

40

【0148】

MDDT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）を作製することも可能である。ノックイン技術を用いて、MDDT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試験し、潜在的医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法で

50

は、例えばL I P A Mを乳汁内に分泌するなど、M D D Tを過剰に発現する哺乳動物近交系は、便利なタンパク質源となり得る（J a n n e , J . 他（1998）B i o t e c h n o l . A n n u . R e v . 4 : 5 5 - 7 4）。

【0149】

（治療）

M D D Tの各領域と疾患検出および治療用分子類との間には、例えば配列およびモチーフの文脈における、化学的および構造的類似性が存在する。また、M D D Tを発現する組織の数例は、表6を見られたい。このように、M D D Tは、細胞増殖異常、自己免疫／炎症の疾患、発生または発達障害、および神経障害において、或る役割を果たすと考えられる。M D D Tの発現または活性の増大に関連する疾患の治療においては、M D D Tの発現または活性を低下させることが望ましい。また、M D D T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、M D D Tの発現または活性を亢進させることが望ましい。

【0150】

したがって、一実施態様において、M D D T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、被験者にM D D Tまたはその断片や誘導体を投与し得る。限定するものではないが、このような疾患には細胞増殖異常、自己免疫／炎症疾患、発達障害、および神経障害が含まれ、細胞増殖異常の中には日光角化症、動脈硬化、アテローム硬化、滑液包炎、硬変、肝炎、混合性結合組織病（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、白血病、リンパ腫、黒色腫、骨髄腫、肉腫、および奇形癌など癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、および子宮の癌が含まれる。自己免疫／炎症疾患には、後天性免疫不全症候群（A I D S）、アジソン病（慢性原発性副腎機能不全）、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、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症外胚葉ジストロフィー（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、リンパ球傷害因子性偶発性リンパ球減少症、新生児溶血性疾患（胎児赤芽球症）、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、過敏性腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、膵炎、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、全身性強皮症、血小板減少性紫斑病、潰瘍性大腸炎、ぶどう膜炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれる。発達障害には尿細管性アシドーシス、貧血、クッシング症候群、軟骨形成不全性小人症、デュシェンヌ／ベッカー型筋ジストロフィー、癲癇、性腺形成異常、W A G R症候群（ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神遅滞）、スミス－マジェニス症候群（S m i t h－M a g e n i s s y n d r o m e）、骨髄異形成症候群、遺伝性粘膜上皮異形成、遺伝性角皮症や、シャルコーマリーツース病および神経線維腫症などの遺伝性神経病、甲状腺機能低下症、水頭症や、S y n d e n h a m 舞蹈病（S y n d e n h a m ’ s c h o r e a）および脳性小児麻痺などの発作障害、二分脊椎、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感音難聴が含まれる。神経障害には、癲癇、虚血性脳血管障害、脳卒中、脳腫瘍、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病およびその他の錐体外路障害、筋萎縮性側索硬化およびその他の運動ニューロン障害、進行性神経性筋萎縮症、網膜色素変性症（色素性網膜炎）、遺伝性運動失調、多発性硬化症および他の脱髄疾患、細菌性およびウイルス性髄膜炎、脳膿瘍、硬膜下膿瘍、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎および神経根炎、ウイルス性中枢神経系疾患と、プリオン病（クールー、クロイツフェルト－ヤコブ病、およびG e r s t m a n n－S t r a u s s l e r－S c h e i n k e r 症候群を含む）、致死性家族性不眠症、神経系性栄養病および代謝病、神経線維腫症、結節硬

化症、小脳網膜血管腫症 (cerebelloretinal hemangioblastomatosis)、脳3叉神経血管症候群、ダウン症を含む中枢神経系性精神遅滞および他の発達障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄疾患、筋ジストロフィー他の神経筋障害、末梢神経疾患、皮膚筋炎および多発性筋炎と、遺伝性、代謝性、内分泌性、および中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺、精神障害（気分性、不安性の障害、統合失調症／分裂病）、季節性感情障害（SAD）、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、遅発性ジスキネジア、ジストニー、パラノイド精神病、帯状疱疹後神経痛、トゥーレット病、進行性核上麻痺、大脳皮質基底核変性 (corticobasal degeneration)、および家族性前頭側頭型痴呆とが含まれる。

10

【0151】

別の実施態様では、限定するものではないが上記した疾患を含む、MDDT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、MDDTまたはその断片や誘導体を発現し得るベクターを被験者に投与しうる。

【0152】

更に別の実施態様では、限定するものではないが上に列記した疾患を含む、MDDT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MDDTを有する組成物を、好適な医薬用キャリアと共に被験者に投与し得る。

【0153】

更に別の実施態様では、限定するものではないが上に列記した疾患を含む、MDDT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、MDDTの活性を調節するアゴニストを被験者に投与し得る。

20

【0154】

更なる実施例では、MDDT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、被験者にMDDTのアンタゴニストを投与し得る。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、自己免疫／炎症疾患、発達障害、および神経障害が含まれる。一実施態様では、MDDT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MDDTを発現する細胞または組織に薬剤を運ぶターゲッティング機構或いは送達機構として間接的に用いられ得る。

【0155】

別の実施例では、MDDT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被験者に投与して、限定するものではないが上記した疾患を含む、MDDT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を成し得る。

30

【0156】

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

40

【0157】

MDDT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できる。詳しくは、精製されたMDDTを用いて抗体を作ったり、治療薬のライブラリ群をスクリーニングしてMDDTと特異結合する薬物を同定できる。MDDTに対する抗体も、本技術分野で公知の方法を用いて製造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、一本鎖抗体、Fab断片、およびFab発現ライブラリによって作られた断片が含まれ得る。中和抗体（すなわち二量体の形成を阻害する抗体）は通常、治療用に好適である。

【0158】

抗体の産生のためには、ヤギ、ウサギ、ラット、マウス、ヒトその他の種々の宿主が、M

50

DDTまたは任意の断片の注入、または免疫原性の特性を備えるそのオリゴペプチドの注入によって免疫化され得る。宿主の種に応じて、種々のアジュバントを用いて免疫応答を高めることもできる。限定するものではないがこのようなアジュバントには、フロイントアジュバントと、水酸化アルミニウムなどのミネラルゲルアジュバントと、リゾレシチン、ブルロニックポリオール、ポリアニオン、ペプチド、油性乳剤、スカシガイのヘモシアニン (KLH)、ジニトロフェノールなどの界面活性剤とがある。ヒトに用いられるアジュバントの中では、BCG (カルメット - ゲラン杆菌) およびコリネバクテリウム - パルバム (*Corynebacterium parvum*) が特に好ましい。

【0159】

MDDTに対する抗体を誘導するために用いるオリゴペプチド、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るアミノ酸配列を持つものが好ましく、一般的には約10個以上のアミノ酸からなるものとなる。これらのオリゴペプチド、ペプチドまたは断片は、天然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。MDDTアミノ酸類の短いストレッチ群は、KLHなど他のタンパク質の配列と融合され、このキメラ分子に対する抗体群が産生され得る。

【0160】

MDDTに対するモノクローナル抗体は、培地内の連続した細胞株によって、抗体分子を産生する任意の技術を用いて作製することが可能である。限定するものではないがこのような技術には、ハイブリドーマ技術、ヒトB細胞ハイブリドーマ技術、およびEBV-ハイブリドーマ技術がある (例えばKohler, G. 他 (1975) *Nature* 256:495497; Kozbor, D. 他 (1985) *J. Immunol. Methods* 81:31-42; Cote, R. J. 他 (1983) *Proc. Natl. Acad. Sci. USA* 80:20262030; および Cole, S. P. 他 (1984) *Mol. Cell Biol.* 62:109-120を参照)。

【0161】

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、好適な抗原特異性および生物学的活性を備える分子を得るために用いられる (例えばMorrison, S. L. 他 (1984) *Proc. Natl. Acad. Sci. USA* 81:68516855; Neuberger, M. S. 他 (1984) *Nature* 312:604608; および Takeda, S. 他 (1985) *Nature* 314:452454を参照)。別法では、当分野で既知の方法を用いて、一本鎖抗体の産生のための記載された技術を適用して、MDDT特異的一本鎖抗体を生成し得る。関連した特異性を有するがイディオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリ類からチェーンシャッフリングによって産生することもできる (例えばBurton D. R. (1991) *Proc. Natl. Acad. Sci. USA* 88:10134-10137を参照)。

【0162】

抗体類の産生は、リンパ球集団における*in vivo*産生の誘導によって、或いは文献に開示されているように非常に特異的な結合試薬の免疫グロブリンのライブラリ類またはパネル類のスクリーニングによっても行い得る (例えばOrlandi, R. 他 (1989) *Proc. Natl. Acad. Sci. USA* 86:38333837; Winter, G. 他 (1991) *Nature* 349:293-299を参照)。

【0163】

MDDTに対する特異結合部位を持つ抗体断片をも産生し得る。例えば、限定するものではないが、このような断片には、抗体分子のペプシン消化によって作製されるF(ab')₂断片と、F(ab')₂断片のジスルフィド架橋を還元することによって作製されるFab断片とがある。あるいは、Fab発現ライブラリを作製することによって、所望

の特異性を持つモノクローナル F a b 断片を迅速且つ容易に同定することが可能となる（例えば H u s e , W . D . 他（1989） S c i e n c e 246 : 1275 - 1281 を参照）。

【0164】

種々の免疫学的検定（イムノアッセイ）を用いてスクリーニングし、所望の特異性を有する抗体を同定することができる。確立された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる免疫放射定量測定法または競合結合試験に対する数々のプロトコルは、本技術分野において公知である。このようなイムノアッセイは通常、M D D T とその特異抗体との間の複合体形成の計測を含む。2つの非干渉性 M D D T エピトープに対して反応性を持つモノクローナル抗体群を用いる、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、競合結合試験も利用できる（P o u n d、前出）。

【0165】

ラジオイムノアッセイ技術と共に S c a t c h a r d 分析などの様々な方法を用いて、M D D T に対する抗体の親和性を評価し得る。親和性を結合定数 K_a で表すが、この K_a は、平衡状態下での M D D T - 抗体複合体のモル濃度を遊離抗体と遊離抗原とのモル濃度で除して得られる値である。ポリクローナル抗体類は多様な M D D T エピトープに対する親和性が不均一であり、或るポリクローナル抗体試薬に関して判定した K_a は、M D D T 抗体群の平均親和性または結合活性を表す。特定の M D D T エピトープに単一特異的なモノクローナル抗体試薬の K_a は、親和性の真の測定値を表す。 K_a 値が $10^9 \sim 10^{12} \text{ L} / \text{mol}$ の高親和性抗体試薬は、M D D T - 抗体複合体が激しい操作に耐えなければならないイムノアッセイに用いるのが好ましい。 K_a 値が $10^6 \sim 10^7 \text{ liter} / \text{mol}$ の低親和性抗体試薬は、M D D T 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（i m m u n o p u r i f i c a t i o n）および類似の処理に用いるのが好ましい（C a t t y , D . （1988） A n t i b o d i e s , V o l u m e I : A P r a c t i c a l A p p r o a c h , I R L P r e s s , W a s h i n g t o n D C ; L i d d e l l , J . E . および A . C r y e r （1991） A P r a c t i c a l G u i d e t o M o n o c l o n a l A n t i b o d i e s , J o h n W i l e y & S o n s , N e w Y o r k N Y）。

【0166】

ポリクローナル抗体試薬の抗体価および結合活性を更に評価して、後に使う或る適用例に対するこのような試薬の品質および適性を決定することができる。例えば、少なくとも $1 \sim 2 \text{ mg} / \text{ml}$ の特異的な抗体、好ましくは $5 \sim 10 \text{ mg} / \text{ml}$ 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、M D D T - 抗体複合体を沈殿させなければならない処理に用いられる。抗体の特異性、抗体価、結合活性、様々な適用例における抗体の品質や使用に対する指針については、一般に入手可能である（例えば前出の C a t t y、同 C o l i g a n 他を参照）。

【0167】

本発明の別の実施例では、M D D T 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を、治療目的で使用できる。ある実施態様では、M D D T をコードする遺伝子のコーディング領域や調節領域に相補的な配列やアンチセンス分子（D N A、R N A、P N A、または修飾したオリゴヌクレオチド）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、アンチセンスオリゴヌクレオチドまたはより大きな断片が、M D D T 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（例えば A g r a w a l , S . , 編集（1996） A n t i s e n s e T h e r a p e u t i c s , H u m a n a P r e s s I n c . , T o t a w a N J を参照）。

【0168】

治療に用いる場合、アンチセンス配列を適切な標的細胞に導入するのに好適な、任意の遺伝子送達系を用いることができる。アンチセンス配列は、転写時に標的タンパク質をコー

ドする細胞配列の少なくとも一部に相補的な配列を作製する発現プラスミドの形で、細胞内に送達し得る（例えば Slater, J. E. 他 (1998) J. Allergy Clin. Immunol. 102 (3) : 469-475 ; および Scanlon, K. J. 他 (1995) 9 (13) : 1288-1296 を参照）。アンチセンス配列はまた、例えばレトロウイルスやアデノ随伴ウイルスベクターなど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、前出の Ausubel, Uckert, W. および W. Walther (1994) Pharmacol. Ther. 63 (3) : 323-347 を参照）。その他の遺伝子送達機構には、リポソーム系、人工的なウイルスエンベロープおよび当分野で公知のその他のシステムが含まれる (Rossi, J. J. (1995) Br. Med. Bull. 51 (1) : 217-225 ; Boado, R. J. 他 (1998) J. Pharm. Sci. 87 (11) : 1308-1315、Morris, M. C. 他 (1997) Nucleic Acids Res. 25 (14) : 2730-2736 など を参照)。

【0169】

本発明の別の実施態様では、MDDT 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。遺伝子治療により、(i) 遺伝子欠損症を治療し（例えば X 染色体連鎖遺伝 (Cavazzana-Calvo, M. 他 (2000) Science 288 : 669-672) により特徴付けられる重症複合型免疫不全 (SCID) - X1 病の場合）、先天性アデノシンデアミナーゼ (ADA) 欠損症に関連する重症複合型免疫不全症候群 (Blaese, R. M. 他 (1995) Science 270 : 475-480 ; Bordignon, C. 他 (1995) Science 270 : 470-475)、嚢胞性繊維症 (Zabner, J. 他 (1993) Cell 75 : 207-216 ; Crystal, R. G. 他 (1995) Hum. Gene Therapy 6 : 643-666 ; Crystal, R. G. 他 (1995) Hum. Gene Therapy 6 : 667-703)、サラセミア (thalassaemia)、家族性高コレステロール血症や、第 VIII 因子若しくは第 IX 因子欠損に起因する血友病 (Crystal, R. G. (1995) Science 270 : 404-410、Verma, I. M. および N. Somia (1997) Nature 389 : 239-242)、(ii) 条件的致死性遺伝子産物を発現させ（例えば制御不能な細胞増殖に起因する癌の場合）、(iii) 細胞内の寄生生物、例えばヒト免疫不全ウイルス (HIV) (Baltimore, D. (1988) Nature 335 : 395-396、Poeschla, E. 他 (1996) Proc. Natl. Acad. Sci. USA. 93 : 11395-11399) などヒトレトロウイルス、B 型若しくは C 型肝炎ウイルス (HBV、HCV)、Candida albicans および Paracoccidioides brasiliensis 等の寄生真菌、並びに熱帯熱マラリア原虫およびクルーズトリパノソーマ等の寄生原虫に対する防御機能を有するタンパク質を発現させることができる。MDDT 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 MDDT 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。

【0170】

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

Boulay, J-L. および H. Recipon (1998) Curr. Opin. Biotechnol. 9:445-450)。

【0171】

MDDT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、PCR2-TOPOTA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。MDDTを発現させるために、(i)構成的に活性なプロモーター (例えば、サイトメガロウイルス (CMV)、ラウス肉腫ウイルス (RSV)、SV40ウイルス、チミジンキナーゼ (TK)、若しくは β -アクチンの遺伝子など)、(ii)誘導性プロモーター (例えば、市販されているT-REXプラスミド (Invitrogen)に含まれている、テトラサイクリン調節性プロモーター (Gossen, M. および H. Bujard (1992) Proc. Natl. Acad. Sci. U. S. A. 89:5547-5551; Gossen, M. 他 (1995) Science 268:1766-1769; Rossi, F. M. V. および H. M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456))、エクジソン誘導性プロモーター (市販されているプラスミドPVGRXRおよびPINDに含まれている: Invitrogen)、FK506/ラパマイシン誘導性プロモーター、またはRU486/ミフェプリストーン誘導性プロモーター (Rossi, F. M. V. および H. M. Blau, 前出)、または(iii)正常な個体に由来する、MDDTをコードする内因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。

【0172】

市販のリボソーム形質転換キット (例えばInvitrogen社のPerfect Lipid Transfection Kit)を用いれば、当業者は実験の各パラメータを最適化する努力をさほど要さず、ポリヌクレオチド群を、培養中の標的細胞群に送達し得る。別法では、リン酸カルシウム法 (Graham, F. L. および A. J. Eb (1973) Virology 52:456-467) 若しくは電気穿孔法 (Neumann, B. 他 (1982) EMBO J. 1:841-845)を用いて形質転換を行う。初代培養細胞にDNAを導入するためには、標準化された哺乳類の形質移入プロトコルの修正が必要である。

【0173】

本発明の別の実施例では、MDDTの発現に関連する遺伝子欠損によって起こる疾患や障害を、(i)レトロウイルス長末端反復配列 (LTR) プロモーター若しくは或る独立プロモーターのコントロール下でMDDT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群と、(iii)追加のレトロウイルス・シス作用性RNA配列と、効率的なベクター増殖に必要なコーディング配列とを伴う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およびPFBNEO)はStratagene社から市販されており、刊行データ (Riviere, I. 他 (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737)に基づく。上記データを引用することを以て本明細書の一部とする。このベクターは、好適なベクター産生細胞株 (VPC L)において増殖される。VPC Lは、各標的細胞上の受容体への親和性を持つエンベロープ遺伝子を、またはVSVGなど汎親和性エンベロープタンパク質を発現する (Armentano, D. 他 (1987) J. Virol. 61:1647-1650; Bender, M. A. 他 (1987) J. Virol. 61:1639-1646; Adam, M. A. および A. D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. 他 (1998)

J. Virol. 72:8463-8471; Zufferey, R. 他 (1998) J. Virol. 72:9873-9880)。Riggに付与された米国特許第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株群を得る方法が開示されており、引用を以て本明細書の一部とする。レトロウイルスベクター類の増殖や、細胞集団(例えばCD4⁺T細胞群)の形質導入、および形質導入した細胞群の患者への戻しは、遺伝子治療の分野では当業者に周知の手法であり、多数の文献に記載がある(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. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290)。

【0174】

別法では、アデノウイルス系遺伝子治療の或る送達系を用いて、MDDTの発現に関連する1つ以上の遺伝子異常を有する細胞群に、MDDTをコードするポリヌクレオチド群を送達する。アデノウイルス系ベクター類の作製およびパッケージングについては、当業者に周知である。複製欠損型アデノウイルスベクター類は、種々の免疫調節タンパク質をコードする遺伝子群を、無損傷の臍島内にインポートする目的で多様に利用し得ることが証明された(Csete, M. E. 他 (1995) Transplantation 27:263268)。潜在的に有用なアデノウイルスベクター類はArmentanoに付与された米国特許第5,707,618号(「Adenovirus vectors for gene therapy」)に記載されており、引用することを以て本明細書の一部とする。アデノウイルスベクター類については、Antinuzzi, P. A. 他 (1999) Annu. Rev. Nutr. 19:511-544 並びに、Verma, I. M. および N. Somia (1997) Nature 18:389:239-242も参照されたい。両文献は、引用を以て本明細書の一部とする。

【0175】

別法では、ヘルペス系遺伝子治療の或る送達系を用いて、MDDTの発現に関連する1つ以上の遺伝子異常を持つ標的細胞群に、MDDTをコードするポリヌクレオチド群を送達する。単純疱疹ウイルス(HSV)系のベクターの利用は、HSVが向性を持つ中枢神経細胞にMDDTを導入する際に特に有用たり得る。ヘルペス系ベクター類の作製およびパッケージングは、当業者に公知である。或る複製適格性単純ヘルペスウイルス(HSV) I型系のベクターが、或るレポーター遺伝子の、霊長類の眼への送達に用いられている(Liu, X. 他 (1999) Exp. Eye Res. 169:385395)。HSV-1ウイルスベクターの作製についても、DeLucaに付与された米国特許第5,804,413号(「Herpes simplex virus strains for gene transfer」)に詳細に開示されており、該特許の引用を以て本明細書の一部とする。米国特許第5,804,413号は、ヒト遺伝子治療などの目的で、好適なプロモーターの制御下において細胞に導入される少なくとも1つの外来遺伝子を有するゲノムを持つ、組換えHSV d92の利用について記載する。上記特許はまた、ICP4、ICP27、およびICP22を欠失した組換えHSV系統の、作製および使用について開示している。HSVベクター類については、Goins, W. F. 他 (1999) J. Virol. 73:519-532 および Xu, H. 他 (1994) Dev. Biol. 163:152-161も参照されたい。両文献は、引用を以て本明細書の一部とする。クローン化したヘルペスウイルス配列群の操作、大ヘルペスウイルスゲノム(large herpesvirus genomes)の様々なセグメントを持つ複数のプラスミドを形質移入した後の組換えウイルスの産生、ヘルペスウイルスの成長および増殖、並びに細胞群へのヘルペスウイルスの感染は、当業者に公知の技術である。

【0176】

別法では、或るαウイルス（正の一本鎖RNAウイルス）ベクターを用いて、MDDTをコードするポリヌクレオチド群を標的細胞群に送達する。プロトタイプのαウイルスであるセムリキ森林熱ウイルス（SFV）の生物学的研究が広範に行われており、遺伝子導入ベクター類は、SFVゲノムに基づく（Garoff, H. および K. -J. Li (1998) Curr. Opin. Biotechnol. 9: 464-469）。αウイルスRNAの複製中に、ウイルスのカプシドタンパク質群を通常はコードする、サブゲノムRNAが産生される。このサブゲノムRNAは、完全長ゲノムRNAよりも高レベルに複製するので、酵素活性（例えばプロテアーゼおよびポリメラーゼ）を持つウイルスタンパク質に比べてカプシドタンパク質群が過剰産生される。同様に、MDDTをコードする配列をαウイルスゲノムのカプシドをコードする領域に導入することによって、ベクター導入細胞において多数のMDDTをコードするRNAが産生され、高いレベルでMDDTが合成される。通常、αウイルスの感染は、数日以内の細胞溶解を伴う。一方、シンドビスウイルス（SIN）の或る変異体を有するハムスター正常腎臓細胞（BHK-21）群が持続的な感染を確立する能力は、αウイルス類の溶解複製を、遺伝子治療に応用し得るように好適に改変可能であることを示唆する（Dryga, S. A. 他 (1997) Virology 228: 74-83）。様々な宿主にαウイルスを導入できるので、様々なタイプの細胞にMDDTを導入することができる。或る集団における或るサブセットの細胞群の特異的形質導入は、形質導入前に細胞の選別を必要とし得る。αウイルスの感染性cDNAクローンの処置方法、αウイルスのcDNAおよびRNAの形質移入方法およびαウイルスの感染方法は、当業者に公知である。

【0177】

転写開始部位由来のオリゴヌクレオチドを用いて遺伝子発現を阻害することも可能である。この位置は、例えば開始部位から数えて約-10と約+10の間である。同様に、三重らせん塩基対の形成方法を用いて阻害が可能となる。三重らせん塩基対形成は、ポリメラーゼ、転写因子または調節分子の結合のために十分に開く、二重らせんの能力を阻害するので有用である。三重らせんDNAを用いる最近の治療の進歩については文献に記載がある（例えばGee, J. E. 他 (1994) in Huber, B. E. および B. I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, 163-177ページを参照）。また、相補配列またはアンチセンス分子を設計し、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の翻訳をブロックし得る。

【0178】

リボザイム類は、酵素性RNA分子である。RNAの特異的切断を触媒するためにリボザイムを用い得る。リボザイム作用のメカニズムは、相補的標的RNAへのリボザイム分子の配列特異的ハイブリダイゼーションと、その後に起こる内ヌクレオチド鎖切断に関与している。例えば、人工的に作製されたハンマーヘッド型リボザイム分子が、MDDTをコードする配列の、内ヌクレオチド鎖切断を特異的且つ効果的に触媒できる可能性がある。

【0179】

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

【0180】

本発明の相補リボ核酸分子およびリボザイムは、核酸分子合成のために当分野でよく知られている任意の方法を用いて作製し得る。作製方法には、固相ホスホラミダイト化学合成など、オリゴヌクレオチドを化学的に合成する方法がある。或いは、MDDTをコードす

るDNA配列の*in vitro*及び*in vivo*転写によってRNA分子を産出し得る。このようなDNA配列は、T7やSP6などの好適なRNAポリメラーゼプロモーターを用いて多様なベクター内に取り込むことが可能である。あるいは、相補的RNAを構成的あるいは誘導的に合成するこれらのcDNA構成物を、細胞株、細胞または組織内に導入することができる。

【0181】

細胞内の安定性を高め、半減期を長くするためにRNA分子を修飾し得る。限定するものではないが可能な修飾としては、分子の5'末端、3'末端、あるいはその両方において隣接配列群を追加することや、分子の主鎖内においてホスホジエステラーゼ結合ではなくホスホロチオエートまたは2'-O-メチルを使用することが含まれる。この概念は、本来はPNA群の産出におけるものであるが、これら全ての分子に拡大することができる。そのためには、内因性エンドヌクレアーゼによって容易には認識されない、アデニン、シチジン、グアニン、チミン、およびウリジンにアセチル、メチル、チオールおよび同様の修飾をしたものや、非従来型塩基、例えばイノシン、クエオシン (*queosine*)、ワイブトシン (*wybutosine*) を加える。

10

【0182】

本発明の更なる実施例には、MDDTをコードするポリヌクレオチドの発現の改変に有効な化合物をスクリーニングする方法を含む。限定するものではないが特異ポリヌクレオチドの発現改変を起こすのに有効であり得る化合物には、オリゴヌクレオチド、アンチセンスオリゴヌクレオチド、三重らせん形成オリゴヌクレオチドや、転写因子などのポリペプチド転写制御因子、および、特異ポリヌクレオチド配列と相互作用し得る非高分子化学的実体がある。有効な化合物は、ポリヌクレオチド発現のインヒビターまたはエンハンサーのいずれかとして作用することによりポリヌクレオチド発現を改変し得る。したがって、MDDT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、MDDT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、MDDT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、MDDT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。

20

【0183】

或る特定ポリヌクレオチドの発現を改変する際の有効性に対して、少なくとも1個または複数個の試験化合物をスクリーニングし得る。試験化合物を得るには、当分野で公知の任意の方法を用い得る。取得方法としては、以下の場合に有効な既知化合物の化学修飾がある。ポリヌクレオチドの発現を改変する場合と、既存の、商用または専用の、天然または非天然の化合物のライブラリから選択する場合と、標的ポリヌクレオチドの化学的および/または構造的特性に基づく化合物を合理的にデザインする場合と、組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合とである。MDDTをコードするポリヌクレオチドを持つサンプルを、このようにして得た試験化合物の少なくとも1つに曝露する。サンプルは例えば、無傷細胞、透過化処理した細胞、*in vitro*無細胞系すなわち再構成生化学系があり得る。MDDTをコードするポリヌクレオチドの発現における改変は、当分野で周知の任意の方法でアッセイする。通常は、MDDTをコードするポリヌクレオチドの配列に相補的なヌクレオチド配列を有するプローブを用いたハイブリダイゼーションにより、特定のヌクレオチドの発現を検出する。ハイブリダイゼーション量を定量することにより、1つ以上の試験化合物に曝露される、または曝露されないポリヌクレオチドの発現の比較のための基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現における変化の検出は、ポリヌクレオチドの発現を改変する際に試験化合物が有効であることを示す。或る特定ポリヌクレオチドの改変発現に有効な化合物のためのスクリーニングを実行でき、例えば分裂酵母 (*Schizosaccharomyces pombe*) 遺伝子発現系 (Atkins, D. 他 (1999) 米国特許第5, 932, 435号、Arndt, G. M. 他 (2000) *Nucleic Acids Res.* 28:E15) またはHeLa細胞等のヒト細胞株 (Clarke, M. L. 他 (2000) *Biochem. Biophys. Res. Commun.* 268

30

40

50

： 8-13）を用いる。本発明の或る特定の実施例は、或る特定ポリヌクレオチド配列に対するアンチセンス活性について、オリゴヌクレオチド（デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾したオリゴヌクレオチド）の組み合わせライブラリをスクリーニングすることに関する（Bruce, T. W. 他（1997）米国特許第5,686,242号、Bruce, T. W. 他（2000）米国特許第6,022,691号）。

【0184】

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

10

【0185】

上記の治療方法はいずれも、例えば、ヒト、イヌ、ネコ、ウシ、ウマ、ウサギ、サルなどの哺乳類を含めて治療が必要な全ての被験体に適用できる。

【0186】

本発明の或る更なる実施態様は、薬物として許容できる或る賦形剤と共に製剤される或る活性成分を一般に有する、或る組成物の投与に関する。賦形剤には例えば、糖、でんぷん、セルロース、ゴムおよびタンパク質がある。様々な剤型が広く知られており、詳細はRemington's Pharmaceutical Sciences (Mack Publishing, Easton PA)の最新版に記載されている。このような組成物は、MDDT、その抗体、擬態物質、アゴニスト、アンタゴニスト、またはインヒビターなどからなる。

20

【0187】

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

【0188】

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

30

【0189】

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

40

【0190】

MDDTまたはその断片を含む高分子を直接細胞内に輸送するべく、特殊な種々の形状の組成物が調製され得る。例えば、細胞不透過性高分子を含むリポソーム製剤は、その高分子の細胞融合と細胞内送達とを促進し得る。別法では、MDDTまたはその断片を、HIV Tat-1タンパク質の陽イオンN末端部に結合することもできる。このようにして生成された融合タンパク質類は、或るマウスモデル系の、脳を含む全ての組織の細胞群に形質導入することがわかっている（Schwarze, S. R. 他（1999）Science 285:1569-1572）。

【0191】

50

任意の化合物に対して、先ず細胞培養アッセイ、例えば新生物性細胞の細胞培養アッセイにおいて、あるいは、動物モデル、例えばマウス、ラット、ウサギ、イヌ、サルまたはブタなどにおいて、治療有効量を推定することができる。動物モデルはまた、好適な濃度範囲および投与経路を決定するためにも用い得る。このような情報を用いて、次にヒトに対する有益な投与量および投与経路を決定することができる。

【0192】

治療有効量とは、症状や容態を回復させる、たとえばMDDTまたはその断片、MDDTの抗体、アゴニストまたはアンタゴニスト、インヒビターなど活性処方成分の量を指す。治療有効度および毒性は、細胞培養または動物実験における標準的な薬学手法によって、例えばED₅₀（集団の50%の治療有効量）またはLD₅₀（集団の50%の致死量）統計を計算するなどして判定できる。毒性効果の、治療効果に対する投与量比が治療指数であり、LD₅₀/ED₅₀比として表すことができる。高い治療指数を示すような組成物が望ましい。細胞培養アッセイと動物実験とから得られたデータは、ヒトに用いる投与量の範囲での調剤に用いられる。このような組成物に含まれる投与量は、毒性を殆どあるいは全く持たず、ED₅₀を含むような血中濃度の範囲にあることが好ましい。用いられる投与形態、患者の感受性および投与の経路によって、投与量はこの範囲内で様々に変わる。

10

【0193】

正確な投与量は、治療が必要な被験者に関する要素を考慮して、現場の医者が決定することになる。十分なレベルの活性成分を与え、あるいは所望の効果を維持すべく、用法および用量を調整する。被験者に関する要素としては、疾患の重症度、患者の一般的な健康状態、患者の年齢、体重および性別、投与の時間および頻度、薬物の配合、反応感受性および治療に対する応答などを考慮し得る。作用期間が長い組成物は、特定の製剤の半減期およびクリアランス率によって3～4日毎に1度、1週間に1度、あるいは2週間に1度の間隔で投与し得る。

20

【0194】

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

30

【0195】

（診断）

別の実施例では、MDDTに特異的に結合する抗体を、MDDTの発現によって特徴付けられる障害の診断に、または、MDDTやそ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得る。診断目的に有用な抗体は、上記の治療の箇所で記載した方法と同じ方法で調合される。MDDTの診断アッセイには、ヒトの体液から、あるいは細胞や組織の抽出物から、抗体および標識を用いてMDDTを検出する方法が含まれる。この抗体は修飾されたものもされていないものも可能であり、レポーター分子との共有結合または非共有結合で標識化できる。レポーター分子としては広くさまざまな種類が本分野で知られており、また使用可能であるが、そのうちのいくつかは上記で説明されている。

40

【0196】

MDDTを測定するための、ELISA、RIA、及びFACSなど、種々のプロトコルは、当分野では周知であり、変容した或いは異常なレベルのMDDT発現を診断するための基盤を提供する。正常或いは標準的なMDDTの発現の値は、正常な哺乳動物、例えばヒトなどの被験体から採取した体液または細胞とMDDTに対する抗体とを複合体の形成に適した条件の下で結合させることによって決定する。標準複合体形成量は、種々の方法、例えば測光法で定量できる。被験体、対照および生検組織からの疾患サンプルで発現す

50

るMDDTの量を標準値と比較する。標準値と被験体との偏差が、疾患を診断するパラメータを確定する。

【0197】

本発明の別の実施態様によれば、MDDT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。用い得るポリヌクレオチドには、オリゴヌクレオチド配列、相補的RNAおよびDNA分子、そしてPNAが含まれる。これらのポリヌクレオチドを用いて、MDDTの発現が疾患と相関し得る生検組織における遺伝子発現を検出し定量し得る。この診断アッセイを用いて、MDDTの存在の有無、更には過剰な発現を判定し、治療時のMDDTレベルの調節を監視する。

【0198】

一実施形態では、MDDTをコードまたは近縁の分子をコードするポリヌクレオチド配列（例えばゲノム配列）を検出可能なPCRプローブ類とのハイブリダイゼーションによって、MDDTをコードする核酸配列を同定できる。プローブが高度に特異的な領域（例えば5'調節領域）から作られている、或いはやや特異性の低い領域（例えば保存されたモチーフ）から作られているかにかかわらず、そ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によって、プローブがMDDTをコードする天然配列のみを同定するかどうか、或いは対立遺伝子や関連配列を同定するかどうかが決まることとなる。

【0199】

プローブはまた、関連する配列の検出に利用でき、また、MDDTをコードする任意の配列との少なくとも50%の配列同一性を有し得る。目的の本発明のハイブリダイゼーションプローブには、DNAあるいはRNAが可能であり、SEQ ID NO: 27-52の配列、あるいはMDDT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。

【0200】

MDDTをコードするDNA群に対して特異的なハイブリダイゼーションプローブ群の作製方法としては、MDDTをコードまたはその誘導体群をコードするポリヌクレオチド配列群を、mRNAプローブ群の作製のためのベクター類にクローニングする方法を含む。mRNAプローブ作製用ベクター類は、当業者に知られており、市販されており、好適なRNAポリメラーゼと好適な標識されたヌクレオチドとを加えることによって、in vitroでRNAプローブを合成するために用い得る。ハイブリダイゼーションプローブ類は、種々のレポーターの集団によって標識され得る。レポーター集団の例としては、³²Pまたは³⁵Sなどの放射性核種が、あるいはアビジン/ビオチン結合系を介してプローブに結合されたアルカリホスファターゼなど酵素標識が挙げられる。

【0201】

MDDTをコードするポリヌクレオチド配列は、MDDTの発現に関係する疾患の診断の為に用い得る。限定するものではないが、このような疾患には細胞増殖異常、自己免疫/炎症疾患、発達障害、および神経障害が含まれ、細胞増殖異常の中には日光角化症、動脈硬化、アテローム硬化、滑液包炎、硬変、肝炎、混合性結合組織病(MCTD)、骨髓線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに、腺癌、白血病、リンパ腫、黒色腫、骨髓腫、肉腫、および奇形癌など癌、具体的には、副腎、膀胱、骨、骨髓、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、および子宮の癌が含まれる。自己免疫/炎症疾患には、後天性免疫不全症候群(AIDS)、アジソン病（慢性原発性副腎機能不全）、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、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症外胚葉ジストロフィー(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球傷害因子性偶発性リンパ球減少症、新生児溶血性疾患（胎児赤芽球症）、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

10

20

30

40

50

ブス病、橋本甲状腺炎、過好酸球増加症、過敏性腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、腓炎、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、全身性強皮症、血小板減少性紫斑病、潰瘍性大腸炎、ぶどう膜炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれる。発達障害には尿細管性アシドーシス、貧血、クッシング症候群、軟骨形成不全性小人症、デュシェンヌ／ベッカー型筋ジストロフィー、癲癇、性腺形成異常、WAGR症候群（ウィルムス腫瘍、無虹彩症、尿生殖器異常、精神遅滞）、スミス－マジェニス症候群（Smith－Magenis syndrome）、骨髄異形成症候群、遺伝性粘膜上皮異形成、遺伝性角皮症や、シャルコーマリーツス病および神経線維腫症などの遺伝性神経病、甲状腺機能低下症、水頭症や、Syndenham 舞蹈病（Syndenham's chorea）および脳性小児麻痺などの発作障害、二分脊椎、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感音難聴が含まれる。神経障害には、癲癇、虚血性脳血管障害、脳卒中、脳腫瘍、アルツハイマー病、ピック病、ハンチントン病、痴呆、パーキンソン病およびその他の錐体外路障害、筋萎縮性側索硬化およびその他の運動ニューロン障害、進行性神経性筋萎縮症、網膜色素変性症（色素性網膜炎）、遺伝性運動失調、多発性硬化症および他の脱髄疾患、細菌性およびウイルス性髄膜炎、脳膿瘍、硬膜下膿瘍、硬膜外膿瘍、化膿性頭蓋内血栓性静脈炎、脊髄炎および神経根炎、ウイルス性中枢神経系疾患と、プリオン病（クールー、クロイツフェルト－ヤコブ病、およびGerstmann－Straussler－Scheinker 症候群を含む）、致死性家族性不眠症、神経系性栄養病および代謝病、神経線維腫症、結節硬化症、小脳網膜血管腫症（cerebellar retinal hemangioblastomatosis）、脳3叉神経血管症候群、ダウン症を含む中枢神経系性精神遅滞および他の発達障害、脳性麻痺、神経骨格異常症、自律神経系障害、脳神経障害、脊髄疾患、筋ジストロフィー他の神経筋障害、末梢神経疾患、皮膚筋炎および多発性筋炎と、遺伝性、代謝性、内分泌性、および中毒性ミオパシーと、重症筋無力症、周期性四肢麻痺、精神障害（気分性、不安性の障害、統合失調症／分裂病）、季節性感情障害（SAD）、静座不能、健忘症、緊張病、糖尿病性ニューロパシー、遅発性ジスキネジア、ジストニー、パラノイド精神病、帯状疱疹後神経痛、トゥーレット病、進行性核上麻痺、大脳皮質基底核変性（corticobasal degeneration）、および家族性前頭側頭型痴呆とが含まれる。MDDTをコードするポリヌクレオチド配列は、変容したMDDT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、サザン法やノーザン法、ドットブロット法、或いはその他の膜系の技術、PCR法や、ディップスティック（dipstick）、ピン（pin）、およびマルチフォーマットのELISA式アッセイ、およびマイクロアレイに使用可能である。このような定性方法または定量方法は、当分野で公知である。

【0202】

或る特定の形態では、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、MDDTをコードするヌクレオチド配列が有用であり得る。MDDT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下で、患者から採取した体液あるいは組織のサンプルに加えることができるであろう。好適なインキュベーション期間が経過したらサンプルを洗浄し、シグナルを定量して標準値と比較する。患者のサンプルのシグナルの量が、対照サンプルと較べて著しく改変された場合は、サンプル内のMDDTをコードするヌクレオチド配列の変異レベルにより、関連する疾患の存在が明らかになる。このようなアッセイは、動物実験、臨床試験における特定の治療効果を推定するため、あるいは個々の患者の治療をモニターするために用いることもできる。

【0203】

MDDTの発現に関連する疾患の診断の基準を提供するために、発現の正常すなわち標準的なプロフィールが確立される。これは、ハイブリダイゼーションあるいは増幅に好適な

条件の下、動物あるいはヒトのいずれかの正常な被験者から抽出された体液あるいは細胞と、MDDTをコードする配列あるいはその断片とを混合させることにより達成され得る。実質的に精製されたポリヌクレオチドを既知量で用いて行った実験から得た値を正常な被験者から得た値と比較することにより、標準ハイブリダイゼーションを定量することができる。このようにして得た標準値は、疾患の徴候を示す患者から得たサンプルから得た値と比較することができる。標準値からの偏差を用いて疾患の存在を確定する。

【0204】

疾患の存在が確定されて治療プロトコルが開始されると、患者の発現レベルが正常な被検者に観察されるレベルに近づき始めたかどうかを測定するため、ハイブリダイゼーションアッセイを定期的に繰り返し得る。連続アッセイから得られた結果を用いて、数日から数ヶ月の期間にわたる治療の効果を示し得る。 10

【0205】

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

【0206】

MDDTをコードする配列から設計されたオリゴヌクレオチドのさらなる診断への利用には、PCRの利用が含まれ得る。これらのオリゴマーは、化学的に合成するか、酵素により生産するか、あるいは*in vitro*で産出し得る。オリゴマーは、好ましくはMDDTをコードするポリヌクレオチドの断片を、或いはMDDTをコードするポリヌクレオチドに対し相補的なポリヌクレオチドの断片を含み、最適な条件下で、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シー条件下で、近縁のDNAあるいはRNA配列の検出、定量、あるいはその両方のため用いることが可能である。 20

【0207】

或る特定の態様において、MDDTをコードするポリヌクレオチド配列群由来のオリゴヌクレオチドプライマー類を用いて、一塩基多型（SNP）を検出し得る。SNPは、多くの場合にヒトの先天性または後天性遺伝病の原因となるような置換、挿入および欠失である。限定するものではないがSNPの検出方法には、SSCP（single-stranded conformation polymorphism）および蛍光SSCP（fSSCP）がある。SSCPでは、MDDTをコードするポリヌクレオチド配列群に由来のオリゴヌクレオチドプライマー類とポリメラーゼ連鎖反応法（PCR）とを用いた、DNAの増幅を行う。このDNAは例えば、病変組織または正常組織、生検サンプル、体液などに由来し得る。DNA内のSNPは、一本鎖形のPCR生成物の2次および3次構造に差異を生じさせる。差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能である。fSSCPでは、オリゴヌクレオチドプライマーを蛍光性に標識する。それによってDNAシーケンシング機などの高処理機器でアンプリマー（amplimer）の検出が可能になる。更に、インシリコSNP（*in silico* SNP, *is* SNP）と呼ばれる配列データベース分析法は、一般的なコンセンサス配列に配列されるような個々のオーバーラップするDNA断片群の配列を比較することにより、多型性を同定し得る。これらのコンピュータベースの方法は、DNAの実験室での調製に、または統計モデルとDNA配列クロマトグラムの自動分析とを用いたシーケンシングのエラーに起因する配列変異を、フィルタリングして除去する。別の態様では、例えば高処理MASSARRAYシステム（Sequenom, Inc., San Diego CA）を用いた質量分析によりSNPを検出し、特徴付ける。 30 40

【0208】

MDDTの発現を定量するために用い得る方法には、ヌクレオチドの放射標識またはビオチン標識、対照核酸の共増幅（coamplification）、および標準曲線から 50

得た結果の補間もある（例えばMelby, P. C. 他（1993）J. Immunol. Methods 159:235244; Duplaa, C. 他（1993）Anal. Biochem. 212:229236を参照）。複数サンプルの定量速度を加速するには、目的のオリゴマーまたはポリヌクレオチドを種々の希釈液中に置き、分光光度法または比色反応によって定量が迅速になるような高処理フォーマットでのアッセイを行い得る。

【0209】

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

10

【0210】

別の実施態様では、MDDT、その断片群または、MDDTに特異的な抗体類を、或るマイクロアレイ上のエレメント群として用い得る。マイクロアレイを用いて、上記のようなタンパク質－タンパク質相互作用、薬物－標的相互作用および遺伝子発現プロファイルをモニターまたは測定することが可能である。

20

【0211】

或る実施態様は、或る組織または細胞タイプの転写イメージを作製する、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タイプによる、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、所与の条件下で所与の時間に発現した遺伝子の数および相対存在量を定量することにより分析し得る（Seilhamer 他米国特許第5,840,484号「Comparative Gene Transcript Analysis」を参照。該特許は特に引用することを以って本明細書の一部となす）。したがって、特定の組織または細胞タイプの転写物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れらの相補体をハイブリダイズすることにより、転写イメージを作製し得る。或る実施態様では、本発明のポリヌクレオチドまたはそれらの相補体が1マイクロアレイ上に複数エレメントの1サブセットを持つような高処理フォーマットでハイブリダイゼーションを発生させる。結果として得られる転写イメージは、遺伝子活性のプロファイルを提供し得る。

30

【0212】

転写イメージは、組織、細胞株、生検またはその生体サンプルから単離した転写物を用いて作製し得る。転写イメージはしたがって、組織または生検サンプルの場合には in vivo、細胞株の場合には in vitro での遺伝子発現を反映する。

40

【0213】

本発明のポリヌクレオチドの発現のプロファイルを作製する転写イメージはまた、in vitro モデル系および薬物の前臨床評価に、あるいは工業的または天然の環境化合物の毒性試験に関連して使用し得る。全ての化合物は、作用および毒性のメカニズムを標示し、しばしば分子フィンガープリントまたは毒性シグネチャと称される、特徴的な遺伝子発現パターンを惹起する（Nuwaysir, E. F. 他（1999）Mol. Carcinog. 24:153-159、Steiner, S. および N. L. Anderson（2000）Toxicol. Lett. 112-113:467-471、該文献は特に引用を以って本明細書の一部となす）。試験化合物が、既知毒性を有する化合物のシグネチャと同一のシグネチャを有する場合には、毒性特性を共有している可

50

能性がある。フィンガープリントまたはシグネチャは、多数の遺伝子および遺伝子ファミリーからの発現情報を含んでいる場合に、最も有用且つ正確である。理想的には、ゲノム全域にわたる発現の測定が、最高品質のシグネチャを提供する。たとえ、発現が任意の試験された化合物によって変容しない遺伝子があったとしても、それらの発現レベルを残りの発現データを標準化することにより使用できるため、それらの遺伝子は重要である。正規化手順は、種々の化合物で処理した後の発現データの比較に有用である。或る毒性シグネチャのエレメント群に遺伝子機能を割り当てることは毒性機序の解釈に役立つが、毒性の予測につながる、シグネチャ群の統計的マッチングには、遺伝子機能の知識は必要でない（例えば2000年2月29日に米国国立環境健康科学研究所（National Institute of Environmental Health Sciences）より発行されたPress Release 00-02を参照されたい。これについては<http://www.niehs.nih.gov/oc/news/toxchip.htm>で入手可能である）。したがって、毒性シグネチャを用いる中毒学的スクリーニングの際に、全ての発現した遺伝子配列を含めることは、重要且つ望ましい。

【0214】

或る実施例では、核酸を有する生体サンプルを試験化合物で処理することにより、この試験化合物の毒性を算定する。処理した生体サンプル中で発現した核酸は、本発明のポリヌクレオチドに特異的な1つ以上のプローブでハイブリダイズし、それによって本発明のポリヌクレオチドに対応する転写レベルを定量し得る。処理した生体サンプル中の転写レベルを、未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は、処理済サンプル中で試験化合物が引き起こす毒性反応を標示する。

【0215】

別の実施態様は、本発明のポリペプチド配列群を用いて或る組織または細胞タイプのプロテオームを分析することに関する。プロテオームの語は、特定の組織または細胞タイプでのタンパク質発現の包括的パターンを指す。プロテオームの各タンパク質成分は、個々に更なる分析の対象とすることができる。プロテオーム発現パターンすなわちプロフィールは、所与の条件下で所与の時間に発現したタンパク質の数および相対存在量を定量することにより分析し得る。したがって、或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タイプのポリペプチドを分離および分析することにより作成し得る。或る実施例では、1次元等電点電気泳動によりサンプルからタンパク質を分離し、2次元ドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に応じて分離するような2次元ゲル電気泳動により、分離が達成される（前出のSteiner および Anderson）。タンパク質は、通常はクーマシーブルー、あるいは銀染色液または蛍光染色液などの物質を用いてゲルを染色することにより、分散した、独自の位置にある点としてゲル中で可視化される。各タンパク質スポットの光学密度は、通常、サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処理または未処理のいずれかの生体サンプルからの、同等に位置したタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を同定する。スポット内のタンパク質は、例えば化学的または酵素的切断とそれに続く質量分析を用いる標準的な方法を用いて部分的にシーケンシングする。或るスポット内のタンパク質の同一性は、その部分配列を、好適には少なくとも5個の連続するアミノ酸残基を、本発明のポリペプチド配列と比較することにより決定し得る。場合によっては、決定的なタンパク質同定のための更なる配列データが得られる。

【0216】

プロテオームのプロフィールは、MDDTに特異的な抗体を用いてMDDT発現レベルを定量することによっても作成可能である。或る実施態様では、マイクロアレイ上のエレメントとしてこれら抗体を用い、マイクロアレイをサンプルに曝して各アレイ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. 他（1999）Anal. Biochem. 270:103-111; Mendoz, L. G. 他（1999）Biotechniques 27: 50

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と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。

【0217】

プロテオームレベルでの毒性シグネチャも中毒学的スクリーニングに有用であり、転写レベルでの毒性シグネチャと並行に分析するべきである。いくつかの組織のいくつかのタンパク質については、転写物の存在量とタンパク質の存在量との相関が乏しいので (Anderson, N. L. および J. Seilhamer (1997) *Electrophoresis* 18: 533-537)、転写イメージにはそれ程影響しないがプロテオームのプロフィールを改変するような化合物の分析において、プロテオーム毒性シグネチャは有用たり得る。更に、体液中の転写物の分析はmRNAの急速な分解のために困難なので、プロテオームのプロフィール作成はこのような場合により信頼し得、情報価値があり得る。

10

【0218】

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

20

【0219】

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

【0220】

マイクロアレイは、本技術分野で既知の方法で調製し、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. USA* 93: 10614-10619、Baldeschweiler 他 (1995) PCT出願第WO95/251116号、Shalon, D. 他 (1995) PCT出願第WO95/35505号、Heller, R. A. 他 (1997) *Proc. Natl. Acad. Sci. USA* 94: 2150-2155、Heller, M. J. 他 (1997) 米国特許第5,605,662号を参照)。様々なタイプのマイクロアレイが公知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, 編集 (1999) Oxford University Press, Londonに記載がある。該文献は、特に引用を以って本明細書の一部となす。

30

40

【0221】

本発明の別の実施態様ではまた、MDDTをコードする核酸配列を用いて、天然のゲノム配列をマッピングするのに有用なハイブリダイゼーションプローブを作製することが可能である。コード配列または非コード配列のいずれかを用いることができ、いくつかの例では、コード配列より非コード配列が好ましい。例えば、多重遺伝子族のメンバー内でのコード配列の保存により、染色体マッピング中に望ましくないクロスハイブリダイゼーションが生じる可能性がある。配列は、或る特定の染色体に、または或る染色体の或る特定領域に、または人為形成の染色体、例えば、ヒト人工染色体 (HAC)、酵母人工染色体 (YAC)、細菌人工染色体 (BAC)、細菌P1産物、あるいは単一染色体cDNAライ

50

ブラリ群に対してマッピングされる（例えばHarrington, J. J. 他（1997）Nat. Genet. 15:345-355; Price, C. M.（1993）Blood Rev. 7:127-134; Trask, B. J.（1991）Trends Genet. 7:149-154を参照）。一度マッピングすると、本発明の核酸配列群を用いて、例えば或る病状の遺伝を特定染色体領域の遺伝とまたは制限酵素断片長多型（RFLP）と関連させるような遺伝子連鎖地図を開発し得る（例えば、Lander, E. S. およびD. Botstein（1986）Proc. Natl. Acad. Sci. USA 83:7353-7357を参照）。

【0222】

蛍光原位置ハイブリッド形成法（FISH）は、他の物理的および遺伝的地図データと相関し得る（例えばHeinz-Ulrich, 他（1995）in Meyers, 前出、965-968ページを参照）。遺伝地図データの例は、種々の科学雑誌あるいはOnline Mendelian Inheritance in Man（OMIM）のウェブサイトに見られる。物理的染色体地図上の、MDDTをコードする遺伝子の位置と、特定の疾患との相関性、あるいは特定の疾患に対する素因との相関性は、この疾患と関連するDNAの領域の決定に役立ち得るため、位置を決定するクローニングの作業を促進し得る。

【0223】

確定した染色体マーカー類を用いた連鎖分析などの物理的マッピング技術、および染色体標本の原位置ハイブリッド形成法を用いて、遺伝地図を拡張することができる。例えばマウスなど別の哺乳類の染色体上に遺伝子を配置することにより、正確な染色体上の遺伝子座が未知でも、関連するマーカー類をしばしば明らかにし得る。この情報は、位置クローニングなどの遺伝子発見技術を用いて疾患遺伝子を探る研究者にとって価値がある。いったん疾患または症候群に関与する遺伝子（群）が、血管拡張性失調症の11q22-23領域など、特定のゲノム領域への遺伝的連鎖によって、大まかに位置決めがなされると、該領域にマップされる任意の配列は、更なる調査のための関連遺伝子あるいは調節遺伝子を提示している可能性がある（例えばGatti, R. A. 他（1988）Nature 336:577-580を参照）。転座、反転などに起因する、健常者、保有者、罹病者の三者間における染色体位置の相違を検出する場合にも、本発明のヌクレオチド配列を用い得る。

【0224】

本発明の別の実施態様では、MDDT、その触媒作用断片群、或いは免疫原断片群、またはそのオリゴペプチド群を、種々の任意の薬物スクリーニング技術における、化合物群のライブラリ類のスクリーニングに用い得る。薬物スクリーニングに用いる断片は、溶液中に遊離しているか、固体支持物に固定されるか、細胞表面上に保持されるか、細胞内に位置することができる。MDDTと試験する薬剤との結合による複合体の形成を測定し得る。

【0225】

別の薬物スクリーニング方法は、目的のタンパク質に対して好適な結合親和性を有する化合物を高い処理能力でスクリーニングするために用いられる（例えばGeysen, 他（1984）PCT出願第WO84/03564号を参照）。この方法においては、多数の様々な低分子の試験用化合物を固体基板上で合成する。試験用化合物は、MDDT、或いはその断片と反応してから洗浄される。結合したMDDTを次に、当分野で周知の種々の方法で検出する。精製したMDDTはまた、上記した薬物スクリーニング技術に用いるプレート上に直接コーティングもできる。別法では、非中和抗体を用いてペプチドを捕捉し、ペプチドを固体支持物に固定することもできる。

【0226】

別の実施例では、MDDTと特異結合可能な中和抗体がMDDTとの結合について試験化合物と競合する、競合的薬物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、抗体を用いて、1つ以上の抗原決定基をMDDTと共有するどのペプチドの存在をも

10

20

30

40

50

検出できる。

【0227】

別の実施態様では、MDDTをコードするヌクレオチド配列群を、将来に開発される分子生物学技術であり、現在知られているヌクレオチド配列群の諸特性（限定はされないが、トリプレット遺伝コード、特異的な塩基対相互作用等を含む）に依存するあらゆる新技術に用い得る。

【0228】

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

10

【0229】

前述した、および以下に記載する全ての特許出願、特許、刊行物の開示を、米国特許出願第60/260,168号、同第60/262,857号、および同第60/262,736号も含めて、言及することをもって本明細書の一部とする。

【0230】

（実施例）

1 cDNAライブラリの作製

Incyte cDNA群の由来は、LIFESEQ GOLDデータベース（Incyte Genomics, Palo Alto CA）に記載されたcDNAライブラリ群である。幾つかの組織はホモジナイズしてグアニジニウムイソチオシアネート溶液に溶解し、他の組織はホモジナイズしてフェノールにまたは変性剤群の好適な混合液に溶解した。混合液の1例であるTRIZOL（Life Technologies）は、フェノールとグアニジンイソチオシアネートとの単相溶液である。結果として得た溶解物は、塩化セシウムのクッション液の上に重層して遠心分離するか、クロロフォルムで抽出した。イソプロパノールか、酢酸ナトリウムとエタノールか、いずれか一方、あるいは別の慣例的方法で、溶解物からRNAを沈殿させた。

20

【0231】

RNAの純度を高めるため、フェノールによるRNAの抽出および沈殿を、必要な回数繰り返した。場合によっては、DNアーゼでRNAを処理した。殆どのライブラリでは、オリゴd(T)連結常磁性粒子（Promega）、OLIGOTEXラテックス粒子（QIAGEN, Chatsworth CA）またはOLIGOTEX mRNA精製キット（QIAGEN）を用いて、ポリ(A)+RNAを単離した。別法では、別のRNA単離キット、例えばPOLY(A) PURE mRNA精製キット（Ambion, Austin TX）を用いて、組織溶解物からRNAを直接単離した。

30

【0232】

場合によってはStratagene社にRNAを提供し、対応するcDNAライブラリ群をStratagene社が作製した。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）を用いて本技術分野で既知の推奨方法または類似の方法でcDNAを合成し、cDNAライブラリを作製した（前出のAusubel, 1997, 5.1-6.6ユニットなどを参照）。逆転写は、オリゴd(T)またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖cDNAに連結反応させ、好適な制限酵素または酵素群でcDNAを消化した。殆どのライブラリに対して、cDNAのサイズ選択（300~1000bp）は、SEPHACRYL S1000、SEPHAROSE CL2BまたはSEPHAROSE CL4Bカラムクロマトグラフィー（Amersham Pharmacia Biotech）、あるいは調製用アガロースゲル電気泳動法を用いて行った。cDNAは好適なプラスミドのポリリンカーの、適合する制限酵素部位にライゲーションされた。好適なプラスミドは、例えばPB LUESCRIPTプラスミド（Stratagene）、PSPORT1プラスミド（Life Technologies）PCDNA2.1プラスミド（Invitro

40

50

en, Carlsbad CA)、PBK-CMVプラスミド(Stratagene)、PCR2-TOPOTAプラスミド(Invitrogen)、PCMV-ICISプラスミド(Stratagene)、pIGEN(Incyte Genomics, Palo Alto CA)、pRARE(Incyte Genomics)、またはpINCY(Incyte Genomics)、またはこれらの誘導体である。組換えプラスミドは、Stratagene社のXL1-Blue、XL1-BlueMRFまたはSOLR、あるいはLife Technologies社のDH5α、DH10BまたはElectroMAX DH10Bなど適格な大腸菌細胞に形質転換した。

【0233】

2 cDNAクローンの単離

実施例1のようにして得たプラスミドの、宿主細胞からの回収は、UNIZAPベクターシステム(Stratagene)を用いた*in vivo*切除によって、あるいは細胞溶解によって行った。プラスミドの精製には、下記の少なくとも1つを用いた。すなわちMagicまたはWIZARD Minipreps DNA精製システム(Promega)、AGTC Miniprep精製キット(Edge Biosystems, Gaithersburg MD)、QIAGEN社のQIAWELL 8 Plasmid、QIAWELL 8 Plus PlasmidおよびQIAWELL 8 Ultra Plasmid 精製システム、R. E. A. L. Prep 96プラスミド精製キットのいずれかである。プラスミドは、沈殿させた後、0.1mlの蒸留水に再懸濁して、凍結乾燥して或いは凍結乾燥しないで4℃で保管した。

【0234】

別法では、高処理フォーマットで直接結合PCR法を用い、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)。宿主細胞の溶解および熱サイクリング過程は、単一反応混合液中で行った。サンプルを加工し、それを384穴プレート内で保管し、増幅したプラスミドDNAの濃度をPICOGREEN色素(Molecular Probes, Eugene OR)およびFLUOROSKANI I 蛍光スキャナ(Labsystems Oy, Helsinki, Finland)を用いて蛍光分析的に定量した。

【0235】

3 シークエンシングおよび分析

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

【0236】

Incyte cDNAに由来するポリヌクレオチド配列は、ベクター、リンカーおよび

10

20

30

40

50

ポリ (A) 配列を除去し、あいまいな塩基をマスクすることによって有効性を確認した。その際、BLASTと、動的プログラミングと、隣接ジヌクレオチド頻度分析とに基づく、アルゴリズムとプログラムとを用いた。次に、Incyte cDNA配列またはそれらの翻訳の問い合わせを、以下のデータベース群に対して行った。すなわち、選抜した公共のデータベース群 (例えばGenBankの霊長類、げっ歯類、哺乳類、脊椎動物、真核生物のデータベースと、BLOCKS、PRINTS、DOMO、PRODOM) と、ヒト、ラット、マウス、線虫 (*Caenorhabditis elegans*)、出芽酵母 (*Saccharomyces cerevisiae*)、分裂酵母 (*Schizosaccharomyces pombe*) および *Candida albicans* からの配列群を持つPROTEOMEデータベース群 (Incyte Genomics, Palo Alto CA)、および、隠れマルコフモデル (HMM) ベースのタンパク質ファミリーデータベース群 (PFAM等) である (HMMは、遺伝子ファミリーのコンセンサス1次構造を分析する確率的アプローチである。例えば、Eddy, S. R. (1996) Curr. Opin. Struct. Biol. 6:361-365を参照のこと)。Eddy, S. R. (1996) Cuff. Opin. Struct. Biol. 6:361-365等を参照) の選択に対するIncyte cDNA配列またはその翻訳を問い合わせた。問合せは、BLAST、FASTA、BLIMPS、およびHMMERに基づくプログラムを用いて行った。Incyte cDNA配列を構築し、完全長のポリヌクレオチド配列を産出した。あるいは、GenBank cDNA群、GenBank EST群、スティッチされた配列群、ストレッチされた配列群、またはGenscan予測コード配列群 (実施例4および5を参照) を用い、Incyte cDNAの集団を完全長まで伸長させた。PhredとPhrapとConsedとに基づくプログラムを用いて構築し、GenMarkとBLASTとFASTAとに基づくプログラムを用いて、cDNAの集団を、オープンリーディングフレームについてスクリーニングした。完全長ポリヌクレオチド配列を翻訳し、対応する完全長ポリペプチド配列を誘導した。あるいは、本発明のポリペプチドは、完全長翻訳ポリペプチドの任意のメチオニン残基で開始し得る。完全長ポリペプチド配列群の続いての分析としての問い合わせを、GenBankタンパク質データベース群 (genpept)、SwissProt、PROTEOMEデータベース群、BLOCKS、PRINTS、DOMO、PRODOMおよびProsite等のデータベースや、PFAM等の隠れマルコフモデル (HMM) ベースのタンパク質ファミリーデータベース群に対し行った。完全長ポリヌクレオチド配列はまた、MACDNASIS PROソフトウェア (日立ソフトウェアエンジニアリング, South San Francisco CA) およびLASERGENEソフトウェア (DNASTAR) を用いて分析した。ポリヌクレオチドとポリペプチドとの配列アラインメントは、MEGALIGNマルチシークエンスアラインメントプログラム (DNASTAR) に組み込まれたCLUSTALアルゴリズムが指定する、デフォルトパラメータを用いて作製する。これは、アラインメントした配列間の一致率も計算する。

【0237】

表7は、Incyte cDNAと完全長配列との分析と構築とに利用したツールとプログラムとアルゴリズムとの概略と、適用可能な説明、参照文献、閾値パラメータを示す。用いたツール、プログラムおよびアルゴリズムを表7の列1に、それらの簡単な説明を列2に示す。列3は好適な参照文献であり、全ての文献は引用を以って全文を本明細書の一部となす。適用可能な場合には、列4は、2つの配列が一致する強さを評価するために用いた、スコア、確率値などのパラメータを示す (スコアが高いほど、または確率値が低いほど、2配列間の同一性が高い)。

【0238】

完全長ポリヌクレオチド配列群とポリペプチド配列群との構築と分析とに用いた上記プログラム群は、SEQ ID NO: 27-52のポリヌクレオチド配列断片群の同定にも利用した。ハイブリダイゼーション技術と増幅技術とに有用な約20~約4000ヌクレオチドの断片群を、表4の列2に示した。

【0239】

4 ゲノムDNAからのコード配列の同定および編集

推定上の疾患検出および治療用分子類の同定には、先ずGenscan遺伝子同定プログラムを、公共のゲノム配列データベース（例えば、gbpriやgbhtg）に対して実行した。Genscanは汎用遺伝子同定プログラムであり、様々な生物からのゲノムDNA配列を分析する（Burge, C. および S. Karlin (1997) J. Mol. Biol. 268:78-94、Burge, C. および S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354 参照）。このプログラムは予測されたエキソン群を連結し、メチオニンから終止コドンに及ぶ、構築されたcDNA配列を形成する。Genscanの出力は、ポリヌクレオチドおよびポリペプチド配列のFASTAデータベースである。Genscanが一度に分析する配列の最大範囲は、30kbに設定した。これらのGenscan予測cDNA配列の内、どの配列が疾患検出および治療用分子をコードするかを決定するために、コードされるポリペプチドを、PFAMモデル群に対し疾患検出および治療用分子について問合せて分析した。潜在的な疾患検出および治療用分子はまた、既に疾患検出および治療用分子として注釈が付けられたIncycyte cDNA配列群に対する相同性により、同定した。これら選択したGenscan予測配列を、次にBLAST分析により、公共データベースgenpeptおよびgbpriと比較した。必要であれば、genpeptからのトップBLASTヒットと比較することによりGenscan予測配列を編集し、余分なまたは省略されたエキソンなど、Genscanが予測した配列におけるエラーを補正した。BLAST分析はまた、Genscan予測配列の、いかなるIncycyte cDNAまたは公共cDNAカバレッジ（coverage）の発見にも用いられ、したがって転写の証拠を提供した。Incycyte cDNAカバレッジが利用できた場合には、この情報を用いてGenscan予測配列を補正または確認した。完全長ポリヌクレオチド配列は、実施例3に記載した構築プロセスを用いて、Incycyte cDNA配列および／または公共cDNA配列でGenscan予測コード配列を構築して得た。あるいは、完全長ポリヌクレオチド配列は、編集後または非編集のGenscan予測コード配列に、完全に由来する。

【0240】

5 ゲノム配列データのcDNA配列データとの統合

スティッチ配列（Stitched Sequence）

部分cDNA配列群を伸長させるため、実施例4に記載のGenscan遺伝子同定プログラムが予測したエキソン群を用いた。実施例3に記載したように構築した部分cDNA群を、ゲノムDNAにマッピングし、また、関連するcDNA群と、1つ以上のゲノム配列からGenscan予測されたエキソン群とを有するクラスタ群に分解した。cDNAとゲノムとの情報を統合すべく、グラフ理論および動的プログラミングに基づく或るアルゴリズムを用いて各クラスタを分析し、潜在的スプライス変異配列群を生成した。配列群を続いて確認、編集または伸長し、完全長配列を創出した。区間の全長が、2つ以上の配列に在るような配列区間群をクラスタ内で同定し、そのように同定された区間群は、推移性により、等しいと考えた。例えば、或る区間が、1つのcDNAと2つのゲノム配列とに在る場合、3つの区間は全て等しいと考えた。このプロセスにより、無関係だが連続したゲノム配列群を、cDNA配列で結び合わせて架橋し得る。このようにして同定した区間を、それらの親配列（parent sequences）に沿って現われる順にスティッチアルゴリズムで「縫い合わせ」、可能な限り最長の配列と変異配列群とを生成した。1種類の親配列に沿って続く区間間の連鎖（cDNA-cDNAまたはゲノム配列-ゲノム配列）は、親の種類を変える連鎖（cDNA-ゲノム配列）より優先した。結果として得たスティッチ配列群を翻訳し、BLAST分析で公共データベースgenpeptおよびgbpriと比較した。Genscanが予測した不正確なエキソン群は、genpeptからのトップBLASTヒットとの比較により補正した。必要な場合、追加cDNA配列群を用いるかゲノムDNAの検査により、配列群を更に伸長させた。

【0241】

ストレッチ配列 (Stretched Sequence)

部分DNA配列群を、BLAST分析に基づく1アルゴリズムにより完全長まで伸長した。まず、BLASTプログラムを用いて、GenBankの霊長類、げっ歯類、哺乳類、脊椎動物および真核生物のデータベースなどの公共データベースに対し、実施例3に記載されたように構築された部分cDNAを問い合わせた。次に、最も近いGenBankタンパク質相同体を、BLAST分析により、Incyte cDNA配列または実施例4に記載のGenScanエキソン予測配列のいずれかと比較した。結果として得られる高スコアリングセグメント対 (HSP) を用いてキメラタンパク質を産出し、翻訳した配列をGenBankタンパク質相同体上にマッピングした。元のGenBankタンパク質相同体に対し、キメラタンパク質内では挿入または欠失が起こり得る。GenBankタンパク質相同体、キメラタンパク質またはその両方をプローブとして用い、公共のヒトゲノムデータベースから相同ゲノム配列を検索した。このようにして、部分的なDNA配列を、相同ゲノム配列の付加によりストレッチすなわち伸長した。結果として得られるストレッチ配列を検査し、完全遺伝子を含んでいるか否かを判定した。

10

【0242】

6 MDDTをコードするポリヌクレオチドの染色体マッピング

SEQ ID NO: 27-52を構築するために用いた配列群を、BLAST他のSmith-Watermanアルゴリズムの実装群を用いて、Incyte LIFESEQデータベースおよび公共ドメインデータベース群の配列と比較した。SEQ ID NO: 27-52と一致するこれらのデータベースの配列を、Phrapなどの構築アルゴリズム (表7) を使用して、連続しオーバーラップする配列のクラスター群に組み入れた。スタンフォード・ヒトゲノムセンター (SHGC)、ホワイトヘッド・ゲノム研究所 (WIGR)、Genethonなどの公的な情報源から入手可能な放射線ハイブリッドおよび遺伝地図データを用いて、いずれかのクラスター化された配列が既にマッピングされているかを判定した。マッピングされた配列が或るクラスターに含まれている場合、そのクラスターの全配列が、個々の配列番号と共に、地図上の位置に割り当てられた。

20

【0243】

地図上の位置は、ヒト染色体の範囲または間隔として表される。センチモルガン間隔の地図上の位置は、染色体のpアームの末端に対して測定する (センチモルガン (cM) は、染色体マーカー間の組換え頻度に基づく計測単位である。平均して、1 cMは、ヒトDNAの1メガベース (Mb) にほぼ等しい。尤も、この値は、組換えのホットスポットおよびコールドスポットに起因して広範囲に変化する)。cM距離は、各クラスター内に配列が含まれる放射線ハイブリッドマーカー類に対して境界を提供するGenethonによってマッピングされた遺伝マーカー群に基づく。NCBI「GeneMap '99」 (<http://www.ncbi.nlm.nih.gov/genemap/>) など一般個人が入手可能なヒトゲノム地図などの資源を用いて、既に同定された疾患遺伝子類が、上記の区間内若しくは近傍にマップされているかを判定できる。

30

【0244】

このようにして、SEQ ID NO: 28は1番染色体の137.6から149.0センチモルガンの区間にマップされ、SEQ ID NO: 30は9番染色体の50.3~75.8センチモルガン、SEQ ID NO: 35は4番染色体の61.0~62.7センチモルガン、SEQ ID NO: 37は4番染色体の145.3~170.9センチモルガン、SEQ ID NO: 38は4番染色体の81.9~115.1センチモルガン、SEQ ID NO: 40は7番染色体の47.9~62.8センチモルガンの区間にマップされた。

40

【0245】

7 ポリヌクレオチド発現の分析

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

50

のハイブリダイゼーションに關与する（例えば前出の S a m b r o o k , 7 章、同 A u s u b e l (1 9 9 5) 4 章および 1 6 章を参照）。

【0246】

B L A S T を応用した類似のコンピュータ技術を用いて、G e n B a n k や L I F E S E Q (I n c y t e G e n o m i c s) などの c D N A データベースにおいて、同一または関連分子を検索した。ノーザン分析は、いくつかの膜系ハイブリダイゼーションよりも非常に速い。更に、任意の特定の一致を厳密なあるいは相同的なものとして分類するか否かを決定するため、コンピュータ検索の感度を修正することができる。検索の基準は積スコアであり、次式で定義される。

【0247】

10

【数1】

(BLAST スコア×配列一致率)

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

【0248】

積スコアは、2つの配列間の類似度と、配列が一致する長さとの両方を考慮している。積スコアは、0～100の正規化された値であり、次のようにして求める。B L A S T スコアにヌクレオチドの配列一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除する。B L A S T スコアを計算するには、或る高スコアリングセグメント対 (H S P) 内の一致する各塩基に+5のスコアを割り当て、各不一致塩基に-4を割り当てる。2つの配列は、2以上のH S Pを共有し得る（ギャップにより隔離される）。2以上のH S Pがある場合には、最高B L A S T スコアのセグメント対を用いて積スコアを計算する。積スコアは、断片的オーバーラップとB L A S T アラインメントの質とのバランスを表す。例えば積スコア100は、比較した2つの配列の短い方の長さ全体にわたって100%一致する場合のみ得られる。積スコア70は、一端が100%一致し、70%オーバーラップしているか、他端が88%一致し、100%オーバーラップしているかのいずれかの場合に得られる。積スコア50は、一端が100%一致し、50%オーバーラップしているか、79%一致し、100%オーバーラップしているかのいずれかの場合に得られる。

20

【0249】

30

あるいは、M D D T をコードするポリヌクレオチド配列を、由来する組織源について分析する。例えば幾つかの完全長配列は、少なくとも一部は、オーバーラップする I n c y t e c D N A 配列群を用いて構築される（実施例3を参照）。各c D N A配列は、ヒト組織から作製されたc D N Aライブラリに由来する。各ヒト組織は、以下の臓器／組織カテゴリーの1つに分類される。すなわち心血管系、結合組織、消化器系、胎芽構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液および免疫系、肝、筋骨格系、神経系、脾臓、呼吸器系、感覚器、皮膚、顎口腔系、非分類性／混合性または尿路である。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。同様に、各ヒト組織は、以下の疾患／条件カテゴリーすなわち癌、細胞株、発達、炎症、神経性、外傷、心血管、プール、その他の1つに分類される。各カテゴリーのライブラリ数を数えて、全カテゴリーの総ライブラリ数で除する。得られるパーセンテージは、M D D T をコードするc D N Aの、組織特異的発現および疾患特異的発現を反映する。c D N A配列およびc D N Aライブラリ／組織の情報は、L I F E S E Q G O L D データベース (I n c y t e G e n o m i c s , P a l o A l t o C A) から得ることができる。

40

【0250】

8 M D D T をコードするポリヌクレオチドの伸長

完全長のポリヌクレオチド配列はまた、完全長分子の適切な断片から設計したオリゴヌクレオチドプライマー群を用いて該断片を伸長させて産生した。或るプライマーは既知の断片の5'伸長を開始するべく合成し、別のプライマーは既知の断片の3'伸長を開始するべく合成した。開始プライマー群の設計にはO L I G O 4 . 0 6 ソフトウェア (N a t

50

ional Biosciences) 或いは別の適切なプログラムを用い、長さが約 22～30ヌクレオチド、GC含有率が約50%以上となり、約68～約72℃の温度で標的配列にアニーリングするようにした。ヘアピン構造およびプライマー-プライマー二量体を生ずるようなヌクレオチド群の伸長は、全て回避した。

【0251】

選択したヒトcDNAライブラリ群を用い、配列を伸長した。2段階以上の伸長が必要または望ましい場合には、付加的プライマーあるいはプライマーのネステッドセットを設計した。

【0252】

高忠実度の増幅が、当業者によく知られている方法を利用したPCR法によって得られた。PCRは、PTC-200サーマルサイクラー(MJ Research, Inc.)を用いて96穴プレート内で行った。反応混合液は、鋳型DNAを有し、また、200nmolの各プライマーを有する。また、 Mg^{2+} と $(NH_4)_2SO_4$ と2-メルカプトエタノールとを含有する反応バッファと、Taq DNAポリメラーゼ(Amersham Pharmacia Biotech)と、ELONGASE酵素(Life Technologies)と、Pfu DNAポリメラーゼ(Stratagene)とを含有する。プライマー対であるPCI AとPCI Bとに対し、以下のパラメータで増幅した。

ステップ1	94℃で3分間	
ステップ2	94℃で15秒間	10
ステップ3	60℃で1分間	
ステップ4	68℃で2分間	
ステップ5	ステップ2、3、及び4を20回繰り返す	
ステップ6	68℃で5分間	
ステップ7	4℃で保存	20

【0253】

別法では、プライマー対であるT7とSK+に対し、以下のパラメータで増幅した。

ステップ1	94℃で3分間	
ステップ2	94℃で15秒間	
ステップ3	57℃で1分間	30
ステップ4	68℃で2分間	
ステップ5	ステップ2、3、及び4を20回繰り返す	
ステップ6	68℃で5分間	
ステップ7	4℃で保存	

【0254】

各ウェルのDNA濃度は、1×TEおよび0.5μlの希釈していないPCR産物に溶解した100μlのPICOGREEN定量試薬(0.25%(v/v) PICOGREEN; Molecular Probes, Eugene OR)を不透明な蛍光光度計プレート(Corning Costar, Acton MA)の各ウェルに分配し、DNAが試薬と結合できるようにして測定した。サンプルの蛍光を計測してDNAの濃度を定量すべく、プレートをFluoroskan II(Labsystems Oy, Helsinki, Finland)でスキャンした。反応混合物のアリコット5～10μlを1%アガロースゲル上で電気泳動法によって解析し、どの反応が配列の伸長に成功したかを判定した。

【0255】

伸長したヌクレオチドは、脱塩および濃縮して384穴プレートに移し、CviJIコラウイルスエンドヌクレアーゼ(Molecular Biology Research, Madison WI)を用いて消化し、音波処理またはせん断し、pUC18ベクター(Amersham Pharmacia Biotech)への再連結を行った。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(0.6～0.50

8%) のアガロースゲル上で分離し、断片を切除し、寒天を Agar ACE (Promega) で消化した。伸長させたクローンを T4 リガーゼ (New England Biolabs, Beverly MA) を用いて pUC 18 ベクター (Amersham Pharmacia Biotech) に再連結し、Pfu DNA ポリメラーゼ (Stratagene) で処理して制限部位のオーバーハングを満たし、適格な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有する培地に移し、それぞれのコロニーを切りとって LB / 2 x カルベニシリン培養液の 384 穴プレートに 37 °C で一晩培養した。

【0256】

細胞を溶解し、Taq DNA ポリメラーゼ (Amersham Pharmacia Biotech) および 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。 10

ステップ 1 94 °C で 3 分間

ステップ 2 94 °C で 15 秒間

ステップ 3 60 °C で 1 分間

ステップ 4 72 °C 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 °C で 5 分間

ステップ 7 4 °C で保存

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

【0257】

同様に、上記手順を用いて完全長ポリヌクレオチド配列を検証した。あるいは、完全長ポリヌクレオチドを用い、上記手順で、そのような伸長のために設計したオリゴヌクレオチド類と、或る適切なゲノムライブラリとを用いて 5' 調節配列を得た。 30

【0258】

9 個々のハイブリダイゼーションプローブの標識及び使用

SEQ ID NO : 27-52 由来のハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノム DNA をスクリーニングする。約 20 塩基対からなるオリゴヌクレオチドの標識について特に記載するが、より大きなヌクレオチド断片に対しても本質的に同一の手順が用いられる。オリゴヌクレオチドは、OLIGO 4.06 ソフトウェア (National Biosciences) 等の最新ソフトウェアを用いて設計し、各オリゴマー 50 pmol と、 $[\gamma\text{-}^{32}\text{P}]$ アデノシン 3 リン酸 (Amersham Pharmacia Biotech) 250 μCi と、T4 ポリヌクレオチドキナーゼ (DuPont NEN, Boston MA) とを混合することにより標識する。標識したオリゴヌクレオチドは、SEPHADEX G-25 超細繊維分子サイズ排除デキストラン ビーズカラム (Amersham Pharmacia Biotech) を用いて実質的に精製する。Ase I、Bgl II、Eco RI、Pst I、Xba I または Pvu II (DuPont NEN) のいずれか 1 つのエンドヌクレアーゼで消化されたヒトゲノム DNA の、典型的な膜ベースのハイブリダイゼーション解析において、毎分 10^7 カウントの標識されたプローブを含むアリコットを用いる。 40

【0259】

各消化物から得た DNA は、0.7 % アガロースゲル上で分画してナイロン膜 (Nytr 50

an Plus, Schleicher & Schuell, Durham NH)に移す。ハイブリダイゼーションは、40℃で16時間行う。非特異的シグナルを除去するため、例えば0.1×クエン酸ナトリウム食塩水および0.5%ドデシル硫酸ナトリウムに一致する条件下で、プロットを室温で順次洗浄する。オートラジオグラフィまたはそれに代わるイメージング手段を用いてハイブリダイゼーションパターンを視覚化し、比較する。

【0260】

10 マイクロアレイ

或るマイクロアレイ上でのアレイエレメント群の結合または合成は、フォトリソグラフィ、圧電印刷（インクジェット印刷、前出のBaldeschweilerなどを参照）、機械的マイクロスポッティング技術およびこれらから派生した技術を用いて達成し得る。上記各技術において基板は、均一な、非多孔性の表面を持つ固体とすべきである（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、ガラスチップおよびシリコンウエハがある。あるいは、ドットプロット法またはスロットプロット法に類似した手順を利用し、熱的、紫外線的、化学的または機械的結合手順を用いて基板の表面にエレメントを配置および結合させてもよい。通常のアレイは、利用可能な、当業者に公知の方法と機械とを用いて作製でき、任意の適正数のエレメントを有し得る（例えばSchena, M. 他（1995）Science 270:467-470; Shalon, D. 他（1996）Genome Res. 6:639-645; Marshall, A. および J. Hodgson（1998）Nat. Biotechnol. 16:27-31を参照）。20

【0261】

完全長cDNA、発現配列タグ（EST）、またはそれらの断片またはオリゴマーが、マイクロアレイのエレメントと成り得る。ハイブリダイゼーションに好適な断片またはオリゴマーを、LASERGENEソフトウェア（DNASTAR）など本技術分野で公知のソフトウェアを用いて選択することが可能である。アレイエレメント群を、生体サンプル中のポリヌクレオチド群とハイブリダイズする。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などの分子タグに抱合させる。ハイブリダイゼーション後、生体サンプルからのハイブリダイズされていないヌクレオチドを除去し、蛍光スキャナを用いて各アレイエレメントでのハイブリダイゼーションを検出する。あるいは、レーザ脱離および質量スペクトロメトリを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上の或るエレメントにハイブリダイズする各ポリヌクレオチドの、相補性の度合と相対存在度とを算定し得る。一実施態様におけるマイクロアレイの調整および使用について、以下に詳述する。30

【0262】

組織または細胞サンプルの調製

全RNAを組織サンプル群から単離するためグアニジニウム チオシアネート法を用い、ポリ（A）⁺ RNAを精製するためオリゴ（dT）セルロース法を用いる。各ポリ（A）⁺ RNAサンプルを、MMLV逆転写酵素、0.05 pg/μlのオリゴ（dT）プライマー（21mer）、1×第一鎖バッファー、0.03 unit/μlのRNアーゼ阻害因子、500 μMのdATP、500 μMのdGTP、500 μMのdTTP、40 μMのdCTP、40 μMのdCTP-Cy3（BDS）またはdCTP-Cy5（Amersham Pharmacia Biotech）を用いて逆転写する。逆転写反応は、GEMBRIGHTキット（Incyte）を用い、200 ngのポリ（A）⁺ RNAを含む体積25 mlで行う。特異的対照ポリ（A）⁺ RNA群は、非コード酵母ゲノムDNAから*in vitro*転写により合成する。37℃で2時間インキュベートした後、各反応サンプル（1つはCy3、もう1つはCy5標識）は、2.5 mlの0.5 M水酸化ナトリウムで処理し、85℃で20分間インキュベートし、反応を停止させてRNAを分解させる。サンプル群の精製には、2つの連続するCHROMA SPIN 30ゲル濾過スピンカラム（CLONTECH Laboratories, Inc.（CLON 40

TECH), Palo Alto CA) を用いる。混合後、2つの反応サンプルのエタノール沈殿を、1 ml のグリコーゲン (1 mg/ml)、60 ml の酢酸ナトリウム、および 300 ml の 100% エタノールで行う。サンプルは次に、Speed VAC (Savant Instruments Inc., Holbrook NY) を用いて完全に乾燥させ、14 μ l の 5 $\times$ SSC/0.2% SDS 中で再懸濁する。

【0263】

マイクロアレイの調製

本発明の配列を用いて、アレイエレメントを作製する。各アレイエレメントは、クローン化した cDNA インサート群により、ベクター含有細菌細胞群から増幅する。PCR 増幅は、cDNA インサートの側面に位置するベクター配列群に相補的なプライマー類を用いる。30 サイクルの PCR によって、1~2 ng の初期量から 5 μ g を超える最終量まで、アレイエレメントを増幅する。増幅したアレイエレメントは、SEPHACRYL-400 (Amersham Pharmacia Biotech) を用いて精製する。 10

【0264】

精製したアレイエレメントは、ポリマーコートされたスライドガラス上に固定する。顕微鏡スライドガラス (Corning) は、0.1% の SDS およびアセトン中で超音波洗浄し、処理中および処理後に多量の蒸留水で洗浄する。スライドガラスは、4% フッ化水素酸 (VWR Scientific Products Corporation (VWR), West Chester PA) 中でエッチングし、多量の蒸留水で洗浄し、95% エタノール中で 0.05% アミノプロピルシラン (Sigma) を用いてコーティングする。コーティングしたスライドは、110 $^{\circ}$ C のオーブンで硬化させる。 20

【0265】

米国特許第 5,807,522 号に記載されている手順を用いて、コーティングしたガラス基板にアレイエレメント群を付加する。この特許は、引用を以って本明細書の一部とする。平均濃度 100 ng/ μ l のアレイエレメント DNA 1 μ l を、高速ロボット装置 (robotic apparatus) により、開放型キャピラリープリンティングエレメント (open capillary printing element) に充填する。装置はここで、スライド毎に約 5 nl のアレイエレメントサンプルを置く。

【0266】

マイクロアレイを UV 架橋するため、STRATALINKER UV 架橋剤 (Stratagene) を用いる。マイクロアレイの洗浄を、室温において、0.2% SDS で 1 回、蒸留水で 3 回行う。非特異結合部位をブロックするため、マイクロアレイのインキュベートをリン酸緩衝生理食塩水 (PBS) (Tropix, Inc., Bedford MA) 中の 0.2% カゼイン中において 60 $^{\circ}$ C で 30 分間行った後、前に行ったように 0.2% SDS および蒸留水で洗浄する。 30

【0267】

ハイブリダイゼーション

ハイブリダイゼーション反応に用いる 9 μ l のサンプル混合体には、Cy3 または Cy5 で標識した cDNA 合成産物群の各 0.2 μ g を、5 $\times$ SSC, 0.2% SDS ハイブリダイゼーション緩衝液中に含む。サンプル混合体は、65 $^{\circ}$ C まで 5 分間加熱し、マイクロアレイ表面上で等分して 1.8 cm² のカバーガラスで覆う。アレイを、顕微鏡用スライドよりわずかに大きい空洞を有する防水チェンバーに移す。チェンバー内を湿度 100 に保つため、140 μ l の 5 $\times$ SSC をチェンバーの 1 コーナーに加える。アレイを入れたチェンバーは、60 $^{\circ}$ C で約 6.5 時間インキュベートする。アレイは、第 1 洗浄緩衝液中 (1 $\times$ SSC, 0.1% SDS) において 45 $^{\circ}$ C で 10 分間洗浄し、第 2 洗浄緩衝液中 (0.1 $\times$ SSC) において 45 $^{\circ}$ C で 10 分間ずつ 3 回洗浄し、乾燥させる。 40

【0268】

検出

レポーター標識したハイブリダイゼーション複合体を検出するには、Cy3 の励起のために 488 nm、Cy5 の励起のために 632 nm でスペクトル線を発生し得る Innov 50

a 70 混合ガス10Wレーザー (C o h e r e n t , I n c . , S a n t a C l a r a C A) を備えた顕微鏡を用いる。励起レーザー光の焦点をアレイ上に置くため、20×顕微鏡対物レンズ (N i k o n , I n c . , M e l v i l l e N Y) を用いる。アレイを含むスライドを、顕微鏡の、コンピュータ制御のX-Yステージに置き、対物レンズを通してラスタースキャンする。本実施例で用いる1.8cm×1.8cmのアレイは、解像度20μmでスキャンする。

【0269】

異なる2回のスキャンで、混合ガスマルチラインレーザーは2つの蛍光色素を連続的に励起する。発光された光は、波長に基づき分離され、2つの蛍光色素に対応する2つの光電子増倍管検出器 (P M T R 1 4 7 7 , H a m a m a t s u P h o t o n i c s S y s t e m s , B r i d g e w a t e r N J) に送られる。好適なフィルタ群をアレイと光電子増倍管との間に設置して、シグナルをフィルタする。用いる蛍光色素の最大発光の波長は、Cy3では565nm、Cy5では650nmである。装置は両方の蛍光色素からのスペクトルを同時に記録し得るが、レーザー源において好適なフィルタを用いて、蛍光色素1つにつき1度スキャンし、各アレイを通常2度スキャンする。

10

【0270】

通常、各スキャンの感度を較正するため、cDNA対照種を或る既知濃度でサンプル混合体に添加し、対照種が発現するシグナル強度を用いる。アレイ上の或る特定の位置には或る相補的DNA配列が含まれ、その位置におけるシグナルの強度をハイブリダイゼーション種の重量比1:100,000で相関させる。異なる源泉 (例えば代表的な試験細胞および対照細胞など) からの2つのサンプルを、各々異なる蛍光色素で標識し、異なる発現をする遺伝子群を同定するために単一のアレイにハイブリダイズする場合には、較正するcDNAのサンプルを2種の蛍光色素で標識し、ハイブリダイゼーション混合液に各々等量を加えることによって較正を行う。

20

【0271】

光電子増倍管の出力をデジタル化するため、IBMコンパチブルPCコンピュータにインストールした12ビットRTI-835Hアナログデジタル (A/D) 変換ボード (A n a l o g D e v i c e s , I n c . , N o r w o o d M A) を用いる。デジタル化したデータは、青色 (低シグナル) から赤色 (高シグナル) までの擬似カラースケールへのリニア20色変換を用いて、シグナル強度がマッピングされたイメージとして表示される。データは、定量的にも分析される。2つの異なる蛍光色素を同時に励起および測定する場合には、各蛍光体の発光スペクトルを用いて、データは先ず蛍光色素間の光学的クロストーク (発光スペクトルの重なりに起因する) を補正される。

30

【0272】

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

【0273】

11 相補的ポリヌクレオチド

MDDTをコードする配列群或いはその任意の部分に対する相補配列群を用いることにより、天然MDDTの発現が検出、低減、または阻害される。約15~30塩基対を含むオリゴヌクレオチドの使用について記すが、これより小さなあるいは大きな配列の断片の場合でも、本質的に同じ手順を用いる。適切なオリゴヌクレオチドを設計するため、O l i g o 4 . 0 6 ソフトウェア (N a t i o n a l B i o s c i e n c e s) およびMDDTのコーディング配列を用いる。転写を阻害するためには、最も独特な5'配列から相補的オリゴヌクレオチドを設計し、これを用いて、プロモーターがコーディング配列に結合するのを防止する。翻訳を阻害するには、相補的なオリゴヌクレオチドを設計して、MDDTをコードする転写物にリボソームが結合するのを防ぐ。

40

【0274】

50

1.2 MDDTの発現

MDDTの発現および精製は、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。細菌内でMDDTを発現するには、cDNAを好適なベクターにサブクロニングする。ベクターは、抗生物質耐性遺伝子と、cDNA転写レベルを高める誘導性プロモーターとを持つものを用いる。このようなプロモーターとしては、lacオペレーター調節エレメントと併用するT5またはT7バクテリオファージプロモーター、およびtrp-lac(tac)ハイブリッドプロモーターが含まれるが、これらに限定するものではない。組換えベクターを、BL21(DE3)などの好適な細菌宿主に形質転換する。抗生物質耐性細菌にMDDTを発現させるには、イソプロピルβ-Dチオガラクトピラノシド(IPTG)で誘発する。真核細胞でのMDDT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るAutographica californica核多角体病ウイルス(AcMNPV)の組換え型を感染させて行う。バキュロウイルスの非必須ポリヘドリン遺伝子をMDDTをコードするcDNAと置換するには、相同組換えを行うか、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を行う。ウイルスの感染力は維持され、強力な多角体プロモーターによって高レベルのcDNA転写が行われる。組換えバキュロウイルスは、多くの場合は夜蛾の1種Spodoptera frugiperda(Sf9)昆虫細胞への感染に用いるが、ヒト肝細胞への感染に用いることもある。後者の感染の場合は、バキュロウイルスへの更なる遺伝的修飾が必要になる(Engelhard, E. K. 他(1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. 他(1996) Hum. Gene Ther. 7:1937-1945を参照)。

【0275】

殆どの発現系では、MDDTが、例えばグルタチオンSトランスフェラーゼ(GST)と、またはFLAGや6-Hisなどのペプチドエピトープ標識と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を、迅速に1ステップで行い得る。GSTは日本住血吸虫からの26kDaの酵素であり、タンパク質の活性および抗原性を維持した状態で、固定化したグルタチオン上での融合タンパク質の精製を可能とする(Amersham Pharmacia Biotech)。精製の後、GST部分を、特異的に操作した部位でMDDTからタンパク質的に切断できる。FLAGは8アミノ酸のペプチドであり、市販されているモノクローナルおよびポリクローナル抗FLAG抗体(Eastman Kodak)を用いた免疫親和性精製を可能にする。6ヒスチジン残基が連続して伸長した6-Hisは、金属キレート樹脂上での精製を可能にする(QIAGEN)。タンパク質の発現および精製の方法は、前出のAusubel(1995)10章、16章に記載されている。これらの方法で精製したMDDTを直接用いて、以下の実施例16の、適用可能なアッセイを行うことができる。

【0276】

1.3 機能的アッセイ

MDDTの機能は、哺乳動物細胞培養系において生理的に高められたレベルでの、MDDTをコードする配列の発現によって算定する。cDNAを高いレベルで発現する強いプロモーターを持つ哺乳動物発現ベクターにcDNAをサブクロニングする。選択されるベクターとしては、PCMV SPORT(Life Technologies)およびPCR3.1(Invitrogen, Carlsbad CA)があり、どちらもサイトメガロウイルスプロモーターを持つ。リポソーム製剤あるいは電気穿孔法を用いて、5~10μgの組換えベクターをヒト細胞株、例えば内皮由来または造血由来の細胞株に、一過的に形質移入する。更に、標識タンパク質をコードする配列を含む1~2μgのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、組換えベクターからのcDNA発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質(GFP; Clontech)、CD64またはCD64-GFP融合タンパク質から

選択できる。自動化された、レーザ光学に基づく技術であるフローサイトメトリー（FCM）を用いて、GFPまたはCD64-GFPを発現する形質移入された細胞を同定し、それらの細胞のアポトーシス状態や他の細胞特性を評価する。FCMは、細胞死に先行するかあるいは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、プロピジウムヨウ化物によるDNA染色によって計測される核DNA内容物の変化、前方散乱光と90°側方散乱光によって計測される細胞サイズと顆粒性の変化、プロモデオキシウリジンの取込量の低下によって計測されるDNA合成の下方調節、特異抗体との反応性によって計測される細胞表面および細胞内におけるタンパク質の発現の変容、および蛍光複合アネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変容とがある。フローサイトメトリー法については、Ormerod, M. G. (1994) Flow Cytometry, Oxford, New York NYに記述がある。 10

【0277】

遺伝子発現におけるMDDTの影響は、MDDTをコードする配列と、CD64またはCD64-GFPのどちらかまたは形質移入された、高度に精製された細胞集団を用いて評価することができる。CD64またはCD64-GFPは、形質転換された細胞表面で発現し、ヒト免疫グロブリンG (IgG) の保存された複数の領域に結合する。形質転換された細胞と形質転換されない細胞とは、ヒトIgGかCD64に対する抗体のどちらかで被覆された磁気ビーズを用いて効率的に分離することができる (DYNAL, Lake Success NY)。mRNAは、当業者に周知の方法で細胞から精製することができる。MDDTと、目的とする他の遺伝子とをコードするmRNAの発現は、ノーザン分析やマイクロアレイ技術で分析することができる。 20

【0278】

1.4 MDDTに特異的な抗体の作製

実質的に精製されたMDDTを、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. 182: 488-495を参照) または他の精製技術で作製し、これ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。

【0279】

別法では、MDDTアミノ酸配列をLASERGENEソフトウェア (DNASTAR) を用いて解析して免疫原性の高い領域を決定し、対応するオリゴペプチドを合成してこれを用いて当業者に周知の方法で抗体を生産する。C末端付近の、或いは隣接する親水性領域内のエピトープなどの適切なエピトープの選択については、当分野で周知である (例えば、前出のAusubel, 1995, 11章を参照)。 30

【0280】

通常は、長さ約15残基のオリゴペプチドを、FMOCケミストリを用いるABI 431A ペプチドシンセサイザ (Applied Biosystems) を用いて合成し、N-マレイミドベンゾイル-N-ヒドロキシスクシンイミドエステル (MBS) を用いた反応によってKLH (Sigma-Aldrich, St. Louis MO) に結合させて、免疫原性を高める (例えば前出のAusubel, 1995を参照)。完全フロイントアジュバントにおいて、オリゴペプチド-KLH複合体を用いてウサギを免疫化する。得られた抗血清の抗ペプチド活性および抗MDDT活性を検査するには例えば、ペプチドまたはMDDT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、放射性ヨウ素標識したヤギ抗ウサギIgGと反応させる。 40

【0281】

1.5 特異的な抗体を用いる天然MDDTの精製

天然MDDTあるいは組換えMDDTを実質的に精製するため、MDDTに特異的な抗体類を用いるイムノアフィニティー (免疫親和性) クロマトグラフィーを行う。イムノアフィニティーカラムは、CNBr-活性化SEPHAROSE (Amersham Phar 50

macia Biotech) のような活性化クロマトグラフィ用レジンと抗MDDT抗体とを共有結合させることにより形成する。結合後に、製造者の使用説明書に従ってこのレジンをブロックし、洗浄する。

【0282】

MDDTを有する培養液をイムノアフィニティーカラムに通し、MDDT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MDDTとの結合を切るような条件で（例えば、pH2～3のバッファー、或いは高濃度の尿素またはチオシアン酸イオンのようなカオトロップで）溶出させ、MDDTを収集する。

【0283】

10

16 MDDTと相互作用する分子の同定

MDDTまたは生物学的に活性なMDDT断片を、¹²⁵I ボルトンハンター試薬で標識する（例えば Bolton, A. E. および W. M. Hunter (1973) Biochem. J. 133: 529-539を参照）。マルチウェルプレートの各ウェルに予め配列しておいた候補の分子群を、標識したMDDTと共にインキュベートし、洗浄して、標識されたMDDT複合体を有する全てのウェルをアッセイする。様々なMDDT濃度で得られたデータを用いて、候補分子と結合したMDDTの数量、親和性、および会合についての値を計算する。

【0284】

別法では、MDDTと相互作用する分子を、Fields, S. および O. Song (1989, Nature 340: 245-246) に記載の酵母2-ハイブリッドシステム (yeast two-hybrid system) や、MATCHMAKERシステム (Clontech) などの、2-ハイブリッドシステムに基づく市販のキットを用いて分析する。

20

【0285】

MDDT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大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し得る (Nandabalan, K. 他 (2000) 米国特許第6, 057, 101号)。

30

【0286】

当業者には、本発明の要旨および精神から逸脱しない範囲での、記載した本発明の方法およびシステムの、種々の修正および変更の手段は自明であろう。本発明について説明するにあたり幾つかの実施例に関連して説明を行ったが、本発明の請求の範囲が、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。分子生物学または関連分野の専門家には明らかな、本明細書に記載する本発明の実施方法の様々な修正は、明確に特許請求の範囲内にあるものとする。

【0287】

(表の簡単な説明)

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

40

【0288】

表2は、GenBank識別番号と、本発明のポリペプチドに最も近いGenBank相同体の注釈とを示す。また、各ポリペプチドとその相同体(1つ以上)が一致する確率スコアも併せて示す。

【0289】

表3は、予測されるモチーフおよびドメインなど、本発明のポリヌクレオチド配列の構造的特徴を、ポリペプチドの分析に用いる方法、アルゴリズムおよび検索可能なデータベースと共に示す。

【0290】

50

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

【 0 2 9 1 】

表 5 は、本発明のポリヌクレオチドの代表的な c D N A ライブラリを示す。

【 0 2 9 2 】

表 6 は、表 5 に示した c D N A ライブラリの作製に用いた組織およびベクターを説明する付表である。

【 0 2 9 3 】

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

10

【 表 1 】

表 1

IncYTE プロジ エクト ID	ポリペプチド SEQ ID NO.	IncYTE ポリペプチ ド ID	ポリヌクレオチド SEQ ID NO.	IncYTE ポリヌクレオ チド ID
6242606	1	6242606CD1	27	6242606CB1
997105	2	997105CD1	28	997105CB1
1482843	3	1482843CD1	29	1482843CB1
3218219	4	3218219CD1	30	3218219CB1
7371678	5	7371678CD1	31	7371678CB1
7473883	6	7473883CD1	32	7473883CB1
7478662	7	7478662CD1	33	7478662CB1
7650474	8	7650474CD1	34	7650474CB1
8092378	9	8092378CD1	35	8092378CB1
1263178	10	1263178CD1	36	1263178CB1
3535417	11	3535417CD1	37	3535417CB1
7473555	12	7473555CD1	38	7473555CB1
7474985	13	7474985CD1	39	7474985CB1
7475137	14	7475137CD1	40	7475137CB1
5036986	15	5036986CD1	41	5036986CB1
1375644	16	1375644CD1	42	1375644CB1
1356261	17	1356261CD1	43	1356261CB1
1419379	18	1419379CD1	44	1419379CB1
3509272	19	3509272CD1	45	3509272CB1
708425	20	708425CD1	46	708425CB1
7469966	21	7469966CD1	47	7469966CB1
6920129	22	6920129CD1	48	6920129CB1
71514704	23	71514704CD1	49	71514704CB1
7715945	24	7715945CD1	50	7715945CB1
7025368	25	7025368CD1	51	7025368CB1
7500954	26	7500954CD1	52	7500954CB1

20

30

40

【 表 2 】

表 2 - 1

ポリペプチド SEQ ID NO:	Incyte ポリペプ チド ID	GenBank ID NO:	確率スコア	GenBank 相同体
1	6242606CD1	g2706620	1.40E-45	NA14, 核内自己抗原 (14 kDa), コイルドコイル ドメイン在り [ヒト] (Ramos-Morales, F. 他(1998) J. Biol. Chem. 273:1634-1639)
10	1263178CD1	g9896486	2.10E-40	RP42 タンパク質 [ヒト] (Mas, C. 他(2000) Genomics 65:70-74)
11	3535417CD1	g13881165	6.00E-83	[ヒト] 胚移植 CD155 [LAO/AO 輸送系キナーゼ
12	7473555CD1	g4185940	2.20E-22	[ヒト] 内因性レトロウイルス K] エンペロープ蛋白 質
13	7474985CD1	g11096305	3.70E-109	[ヒト] (AF310106) NALP2 (NACHT, LRR, および P4D を含む)
14	7475137CD1	g2072963	8.30E-126	[ヒト] p40
15	5036986CD1	g5070357	6.70E-181	[ウシ] 生体異物/中鎖脂肪酸 CoA リガナーゼ型 XL- III
16	1375644CD1	g10441648	4.50E-56	[マウス] mdgl-1
22	6920129CD1	g7228476	9.40E-16	[ヒト] 象牙質シアロタンパク質前駆体 Gu, K., 他 (2000) Molecular cloning of a human dentin sialoprotein gene. Eur. J. Oral Sci. 108, 35-42
23	71514704CD1	g7580490	1.80E-45	[ヒト] NM23-H8 Postel, E.H. (1998) NM23-NDP kinase. Int. J. Biochem. Cell Biol. 30, 1291-1295
24	7715945CD1	g406263	2.60E-131	[マウス] T10 Halford, S., 他 (1993) Isolation of a gene expressed during early embryogenesis from the region of 22q11 commonly deleted in DiGeorge syndrome. Hum. Mol. Genet. 2, 1577-1582

10

20

30

【表 3】

表 2 - 2

ポリペプチド SEQ ID NO:	Incyte ポリペプ チド ID	GenBank ID No:	確率スコア	GenBank 相同体
25	7025368CD1	g11065721	4.10E-39	【ヒト】 (AJ251832) 28kD インターフェロニン応答 性タンパク質

【表 4】

10

20

30

表 3-1

SEQ ID NO.	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	シグネチャドメイン (paired box domain) シグネチャ	分析方法及びデータベース
1	624260CD1	127	T52 T82 T115	N110	対合ボックスドメイン (paired box domain) シグネチャ ヤ: Q42-S97 核内自己抗原 (14 kDa) PD174194; L2-S127	シグネチャドメイン (paired box domain) シグネチャ	ProfileScan BLAST-PRODOM
2	997105CD1	238	S115 T72 T156 T173	N19 N127	ロイシンジッパーモチーフ: L2-L23, L9-L30		MOTIPS
3	1482843CD1	175	S7 S121 T171	N71	シグナル ペプチド: M1-G25		HMER
4	3218219CD1	1146	S35 S36 S56 S98 S121 S361 S450 S503 S585 S609 S653 S709 S817 S824 S908 S958 S1015 S1100 S1141 T27 T40 T65 T250 T294 T305 T362 T462 T512 T654 T713 T782 T926 T951		膜貫通ドメイン: R338-H366, S895-F910 N 末端はサイトゾル内		TMAP
5	7371678CD1	351	S19 S23 S45 S50 S280 S282 S329 T31 T244 T333 Y304				
6	7473883CD1	286	S13 S67 S164 S187 S201 S232 T17 T120 T144	N199			
7	7478662CD1	205	S9 S26 S81 S143 S149 S191 T68	N74 N139	膜貫通ドメイン: D84-A108 N 末端はサイトゾル内		TMAP
8	7650474CD1	233	S13 S21 S87 S118 S150 S215 T80 T101 T181 T206		核タンパク質 SNF7, 推定上の copia 様転移因子 (transposable element) PD009719; G11-A212		BLAST-PRODOM

【表 5】

表 3 - 2

SEQ ID NO.	Incyte ポリペプチド ID	アミノ酸残基数	構造的リン酸化部位	構造的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
9	8092378CD1	205	S101 S127 T45 T51 T53 T61 T118	N88	シグナル ペプチド ; M11-S35	HMER
9					シグナル ペプチド ; M13-S35	HMER
9					膜貫通ドメイン ; V4-T31 N 末端はサイトゾル内にはない	TMAP
10	1263178CD1	304	S25 S30 S71 S81 S261 S268 S293 T56 T143 T272	N26 N54 N248 N259		
11	3535417CD1	418	S79 S113 S136 S291 S317 S328 S374 T70 T106 T158 T168 T223 T289 T356 T384 Y90		膜貫通ドメイン ; Q248-L264 N 末端はサイトゾル内にはない	TMAP
					ATP 結合タンパク質、輸送、LAO/AO 系、キナーゼ、トランスフェラーゼ PD011981 ; I105-E373	BLAST-PRODDM
					mitH/ExC フアミリ ; DM00105P2728453-225; R145-K315	BLAST-DOMO
					ATP/GTP 結合部位モチーフ A (P-loop) ; G150-S157	MOTIFS
12	7473555CD1	1117	S18 S125 S256 S260 S322 S619 S641 S823 S860 S880 S932 S992 S1061 S1096 S1103 T317 T365 T854	N214 N682 N856 N871	膜貫通ドメイン ; G61-A89, A452-I478, G523-R550 N 末端はサイトゾル内にはない	TMAP

10

20

30

【表 6】

表 3-3

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					env, レトロウイルス関連ポリタンパク質、糖タンパク質コード PD023510: C106-V296	BLAST-PRODBM
					ATP/GTP 結合部位モチーフ A (P-loop): A354-S361	MOTIFS
13	7474985(D1)	920	S75 S144 S289 S341 S384 S488 S548 S582 S730 S759 S880 S903 S908 T29 T112 T335 T340 T438 T482 T659 T681 T764 T791 T859 Y844	N96 N710	ロイシンリッチリピート(LRR): N819-N846, R762-Q789, K734-L757	HMMER-PFAM
					膜貫通ドメイン: V506-S526, C790-A811 N 末端はサイトゾル内にある	TMAP
					受容体、アンギオテンシン/バゾプレッシン AII/AVP PD156095: M412-E594	BLAST-PRODBM
					リボスクレアーゼ阻害因子、ロイシンリッチリピート、アンジオグニン受容体 PD017636: L766-L854, I707-I814, W682-S759	BLAST-PRODBM
					ATP/GTP 結合部位モチーフ A (P-loop): G212-T219	MOTIFS
					ミトコンドリア型エネルギー輸送タンパク質シグネチャ: P536-Q545	MOTIFS

【表 7】

表 3 - 4

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
14	7475137CD1	340	S6 S33 S123 S149 S166 S211 S292 S298 T90 T176 T181 T216 T243 T300 T310	N4 N14 N54 N174	L1 エレメント、p40, p150 PD005182: M232-L340	BLAST-PRODOM
					L1 エレメント、p40, p150 PD005183: T53-R142	BLAST-PRODOM
					L1 エレメント、p40, p150 PD004752: M1-R48	BLAST-PRODOM
					L1 エレメント、p40, p150 PD003272: L143-K231	BLAST-PRODOM
15	5036986CD1	577	S19 S176 S184 S272 S325 S388 S476 S495 S571 T239 T353 T375 T406 T555 Y147 Y235 Y455	N467	膜貫通ドメイン: V111-T137, M257-S276 N 末端はサイトゾル内にはない	TMAP
					AMP 結合酵素: N62-V493	HMMER-PFAM
					推定 E: AMP 結合ドメイン: F220-Y235	BLI:IMPS-BLOCKS
					リガナーゼ、合成酵素、生合成、抗生物質、ホスホペンテアイン、AcyCoA PD000070: D396-I494, W81-A281, T137-T366	BLAST-PRODOM
					SA タンパク質、遺伝子シグナル、腎臓特異的 PD151238: V40-W81	BLAST-PRODOM
					推定 E: AMP 結合ドメイン: DM00073/A61209-65-538: E58-Q393, G411-R564 DM00073/P39062150-555: A68-R564 DM00073/P270951107-644: I218-R560, F83-I266 DM00073/P1491212-533: V72-R564	BLAST-PRODOM

【表 8】

表 3-5

SEQ ID NO.	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
					細胞接着配列: RGD: R105-D107, R438-D440 推定上 AMP 結合ドメイン I218-K229 シグナル ペプチド: M1-A62	MOTIFS MOTIFS
16	137564CD1	119	S22 S97 Y105		シグナル ペプチド: M1-A62 ヘルペスシグナールドメイン: S44-L86 SEC17 QCR1, 遺伝子間領域 PD020820: Y21-E108	SPScan ProfileScan BLAST-PRODOM
17	135626CD1	389	S2 S73 S161 S274 S312 S377 T41 T79 T131 T149 T324 T325 T335 T376 Y342 T3 T13 T30 T81		膜貫通ドメイン: A162-Q177 N 末端はサイトソル内	TMAP
18	1419379CD1	266				
19	3509272CD1	661	S147 S241 S245 S285 S292 S322 S327 S429 S498 S571 S593 S594 S598 S634 S635 T18 T42 T71 T255 T276 T279 T353 T373 T398 T404 T461 T540 T652	N40 N60 N347 N514	シグナル ペプチド: M1-C25	SPSCAN
					膜貫通ドメイン: Y539-C567 N 末端はサイトソル内にはない ニューロメディン U タンパク質シグネチャ BL00967: S580-R590	TMAP BLIMPS_BLOCKS

【表 9】

表 3-6

SEQ ID NO.	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
20	708425CD1	205	S14 S18 S40 S69 S140 S148 S152 S189 T2 T42 T193	N83 N123 N127 N136		MOTIFS
21	746996aCD1	833	S40 S97 S111 S143 S189 S208 S324 S415 S434 S473 S482 S542 S546 S621 S707 S807 T35 T148 T202 T350 T394 T532 T553 T604 T654 T775 Y136 Y149 Y253 Y424	N18 N705		MOTIFS
22	8920129CD1	655	S2 S51 S87 S109 S110 S114 S117 S129 S136 S150 S173 S197 S217 S224 S233 S264 S331 S336 S341 S370 S378 S383 S387 S402 S403 S432 S437 S468 S481 S506 S513 S522 S554 S561 S603 S610 S617 S636 S637 S641 S647 S651 T258 T325 T326 T476 T541 T545 T569 T591 Y40	N77 N220 N262 N589	ユビキチン-カルボキシル末端加水分解酵素ファミリー: Q24-R67	HMMER_PFM
					酸性セリン/スチラスターリピート DM04746[P41777]-386; S303-P627 DM04746[S57757]-646; K320-S647, A70-P586	BLAST_DOMO

【表 10】

10

20

30

表 3-7

SEQ ID NO:	Incyte ポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
23	71514704CD1	330	S3 S15 S25 S135 S156 S264 S297 S301 T30 T108 T191 T196 T198 T233 T244	N13	シグナル ペプチド: M1-A68	SPSCAN
					チオレドキシン ファミリ 蛋白質シグネチャ BL00194: V31-K43	BLIMPS_BLOCKS
					チオレドキシン ファミリ 活性部位: T16-L61	PROFILESKAN
					スクレオキシンドニリン酸キナーゼ活性部位: A257-D304	PROFILESKAN
					チオレドキシン ファミリ シグネチャ W38, W38-S47, R78-A89	BLIMPS_PRINTS
23					キナーゼ ニリン酸 スクレオキシンド転移酵素 NDK NDP ATP 結合タンパク質 I 前駆体 PD001018: T161-A307	BLAST_PRODOM
					スクレオキシンドニリン酸キナーゼ DM00773P505891-136: T161-P296 DM00773P438021-140: T161-K303 DM00773P242331-141: T161-F295 DM00773P152661-143: T161-F295	BLAST_DOMO
					チオレドキシン ファミリ 活性部位: V31-F49	MOTIFS
					染色体濃縮の調節因子(RCC1) シグネチャ 2: M222-L232	MOTIFS
24	7715915CD1	276	S48 S65 S110 S192 S256 S263 T98 T128 T251 T253 Y116		タンパク質領域 R186.I SER/THR リッチ T10 DGCR (ASN2PHB) 遺伝子間 PD024111: M1-I134, D132-L274	BLAST_PRODOM

【表 1 1】

10

20

30

表 3-8

SEQ ID NO.	Incyte ペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
25	7025368CD1	232	T26 T90 T176	N74 N110 N210		MOTIFS
26	7500954CD1	120	S16 S42 T33	N3		MOTIFS

【表 1 2】

10

20

30

【表 1 3】

表 4 - 1

ポリヌクレオチド SEQ ID NO/ Incyte ポリヌクレオチド ID/ 配列長		配列断片
27/6242606CB1/675		1-266, 147-266, 181-352, 181-432, 181-500, 181-665, 219-355, 267-373, 268-467, 269-515, 281-328, 299-664, 300-488, 313-563, 400-675
28/997105CB1/1387		1-584, 364-629, 364-677, 373-999, 384-626, 384-641, 385-627, 386-666, 398-604, 398-677, 400-844, 407-585, 407-1024, 445-677, 446-1071, 452-677, 459-677, 502-677, 637-926, 677-1203, 759-996, 759-1002, 759-1012, 868-1387, 948-1249
29/1482843CB1/1534		1-249, 29-328, 49-711, 66-328, 82-430, 87-658, 109-710, 115-719, 138-768, 245-724, 274-637, 281-521, 281-823, 282-691, 319-672, 341-597, 344-597, 373-508, 404-687, 434-687, 462-724, 468-717, 482-1089, 497-1020, 533-1153, 562-1351, 563-866, 563-1163, 621-1437, 686-1437, 690-1437, 692-1337, 698-843, 713-1337, 720-1337, 732-1437, 748-1437, 749-1330, 773-1437, 789-1426, 789-1437, 829-1330, 836-1437, 838-1426, 843-1335, 859-1163, 860-1437, 872-1385, 883-1437, 935-1534, 962-1532, 970-1534, 972-1534, 983-1534, 994-1147, 1031-1407, 1038-1409, 1047-1469, 1052-1324, 1053-1337, 1089-1396, 1093-1347, 1117-1337, 1120-1407, 1125-1407, 1127-1407, 1130-1407, 1136-1407, 1143-1373, 1164-1407, 1172-1407, 1191-1437, 1210-1337, 1239-1525, 1242-1335, 1280-1414, 1282-1534, 1292-1534, 1302-1534, 1417-1517, 1438-1520, 1442-1495, 1443-1506, 1458-1517

表 4-2

ポリスクレオチド SEQ ID NO/ Incyte が リスクリオチド ID/ 配列長	配列断片
30/3218219CB1/4203	1-660, 98-655, 182-463, 182-790, 184-402, 192-752, 194-777, 197-458, 211-782, 211-808, 216-463, 216-623, 220-554, 221-767, 295-731, 375- 971, 368-971, 389-622, 390-971, 402-891, 413-971, 448-971, 476-971, 487-971, 496-971, 500-971, 501-1246, 508-971, 518-971, 521-971, 528- 971, 542-971, 546-971, 560-971, 567-971, 866-1603, 967-1619, 1115- 1637, 1133-1650, 1161-1398, 1178-1577, 1273-1849, 1456-2083, 1498- 1960, 1506-1988, 1508-1717, 1624-2096, 1704-2352, 1738-2373, 1847- 2352, 1908-2315, 1914-2352, 1952-2200, 2024-2545, 2050-2702, 2127- 2405, 2136-2347, 2378-2571, 2395-2701, 2395-2899, 2401-3960, 2553- 2907, 2558-2812, 2559-2999, 2590-2848, 2590-3131, 2650-3272, 2871- 3105, 2957-3218, 2985-3193, 2985-3248, 2985-3507, 3010-3286, 3026- 3179, 3034-3343, 3111-3281, 3117-3382, 3154-3492, 3157-3313, 3157- 3417, 3160-3327, 3160-3431, 3160-3657, 3220-3653, 3220-3775, 3223- 3508, 3224-3357, 3237-3481, 3261-3531, 3454-3706, 3454-3716, 3490- 3776, 3493-3697, 3512-3766, 3513-3728, 3513-3737, 3513-4033, 3519- 3692, 3602-3889, 3615-3836, 3615-4047, 3616-3859, 3646-3784, 3659- 3912, 3699-3961, 3752-3876, 3752-3999, 3752-4203, 3824-4097, 3829- 4114, 3836-4061, 3918-4169, 3933-4203 31/7371678CB1/2082 1-622, 90-715, 104-575, 156-270, 161-412, 197-787, 197-799, 283-845, 287-809, 345-920, 392-1104, 502-705, 502-1114, 557-793, 557-843, 557-957, 557-1075, 557-1076, 557-1077, 557-1110, 557-1128, 557-1182, 557-1192, 557-1204, 560-1004, 595-1202, 608-1077, 610-788, 618-1115, 621-1015, 682-971, 690-1031, 699-1306, 704-1133, 727-896, 741-1135, 756-1135, 787-1065, 822-1391, 823-935, 841-1212, 891-1018, 898-992, 905-1542, 916-979, 960-1199, 979-1292, 995-1352, 1013-1596, 1023- 1263, 1031-1624, 1046-1641, 1080-1596, 1118-1818, 1153-1748, 1171- 1758, 1177-1637, 1189-1702, 1207-1259, 1246-1901, 1253-1816, 1277- 1587, 1278-1875, 1282-1380, 1303-1837, 1318-2032, 1360-1875, 1377- 1943, 1377-2080, 1402-1733, 1439-1679, 1443-1829, 1445-1749, 1459- 1698, 1468-1778, 1469-1896, 1490-2081, 1493-2082, 1509-2082, 1552- 2082, 1559-2082, 1577-1884, 1584-2082, 1593-1771, 1597-2082, 1611- 2082, 1628-2051, 1632-1814, 1667-1954, 1667-2065, 1686-2027, 1694- 2073, 1707-2010, 1752-2082

10

20

30

表 4-3

ポリスクレオチド SBQ ID NO/ Incyte ポリスクレオチド ID/ 配列長		配列断片
32/7473883CB1/1459		1-154, 1-157, 1-255, 1-487, 5-157, 6-157, 8-157, 9-157, 9-850, 10-157, 11-157, 13-157, 15-157, 17-124, 17-157, 18-109, 18-157, 19-157, 20-157, 25-157, 26-153, 26-157, 27-157, 28-157, 113-730, 157-225, 157-234, 157-241, 157-244, 157-246, 157-288, 157-303, 157-315, 157-316, 157-327, 157-342, 157-354, 157-364, 157-385, 157-400, 157-406, 157-423, 157-449, 157-458, 157-465, 157-486, 157-514, 157-559, 157-749, 258-805, 284-858, 292-853, 297-577, 330-578, 336-599, 340-610, 348-771, 349-643, 349-794, 362-826, 370-1011, 381-855, 422-687, 423-897, 437-671, 437-840, 448-909, 464-555, 465-889, 466-1059, 474-1116, 501-744, 501-810, 514-1092, 555-1128, 565-1085, 577-826, 589-870, 589-878, 589-918, 605-1130, 605-1161, 615-815, 617-839, 621-821, 633-876, 635-1105, 662-997, 662-1241, 690-1459, 753-1021, 760-1294, 862-1164, 888-1035, 903-1297, 907-1196, 931-1069, 996-1269, 1421-1456, 1422-1459, 1423-1456
		1-309, 1-599, 15-557, 20-564, 27-275, 27-282, 32-511, 39-629, 89-710, 270-824, 296-803, 378-642, 378-837, 445-692, 452-993, 461-729, 523-1068, 847-1082, 850-1485, 887-1159, 974-1216, 974-1455, 974-1466, 974-1526, 974-1540, 983-1261, 1023-1532, 1023-1626, 1024-1342, 1049-1337, 1070-1806, 1094-1359, 1094-1365, 1113-1372, 1138-1803, 1185-1802, 1196-1806, 1199-1806, 1208-1805, 1208-1806, 1213-1806, 1224-1803, 1226-1555, 1279-1803, 1283-1803, 1288-1806, 1297-1806, 1370-1806, 1415-1777, 1560-1806, 1589-1806, 1594-1806, 1619-1806, 1646-1806, 1674-1806, 1685-1806, 1702-1806, 1758-1806
33/7478662CB1/1806		

10

20

30

表 4 - 4

ポリヌクレオチド SEQ ID NO/ Incyte ポリヌクレオチド ID/ 配列長	配列断片
34/7650474CB1/1698	1-244, 1-259, 1-292, 1-457, 1-459, 1-477, 1-541, 35-234, 35-542, 35-552, 35-598, 35-608, 59-303, 60-402, 69-307, 80-753, 82-353, 110-740, 113-721, 116-379, 120-336, 134-645, 155-596, 159-880, 165-658, 180-684, 182-852, 233-852, 242-898, 251-537, 251-644, 258-498, 258-791, 261-871, 263-622, 265-840, 287-840, 307-899, 330-899, 336-663, 336-847, 345-967, 371-878, 376-909, 404-976, 454-1058, 499-954, 508-1010, 535-1118, 608-1198, 617-757, 643-1170, 660-930, 692-931, 692-1072, 704-1363, 705-1412, 751-1088, 760-1401, 761-1022, 768-1276, 790-1092, 815-1050, 815-1092, 815-1218, 840-1036, 841-1092, 851-1056, 856-1092, 906-1037, 907-1037, 918-1037, 998-1430, 1059-1120, 1059-1173, 1059-1203, 1089-1634, 1092-1327, 1092-1623, 1118-1399, 1121-1401, 1121-1598, 1121-1682, 1144-1414, 1172-1407, 1206-1695, 1217-1695, 1225-1698, 1228-1698, 1233-1690, 1233-1695, 1240-1695, 1245-1485, 1250-1476, 1253-1513, 1253-1696, 1256-1675, 1274-1697, 1276-1695, 1278-1692, 1280-1694
35/8092378CB1/1753	1-276, 1-608, 2-352, 76-789, 170-673, 170-685, 170-761, 170-842, 212-432, 287-549, 287-560, 287-840, 304-875, 405-330, 407-837, 408-707, 416-652, 626-1291, 712-1419, 718-1333, 741-1407, 776-1378, 809-1403, 925-1436, 925-1710, 988-1210, 988-1434, 988-1457, 990-1169, 997-1682, 1110-1691, 1137-1432, 1150-1399, 1161-1409, 1161-1693, 1161-1753, 1197-1456, 1221-1659, 1237-1616, 1238-1527, 1238-1531, 1258-1464, 1258-1470, 1265-1704, 1307-1555
36/1263178CB1/1489	1-326, 22-297, 22-624, 22-678, 57-97, 171-529, 173-707, 276-1190, 501-776, 523-1414, 607-779, 607-1073, 607-1098, 607-1104, 607-1105, 607-1108, 607-1145, 607-1224, 608-1190, 610-960, 610-1186, 611-746, 612-1157, 693-1358, 704-1125, 751-1400, 832-1488, 849-1489, 854-1425, 857-1425, 906-1173, 906-1289, 922-1489, 933-1410, 955-1489, 983-1105, 1021-1383, 1164-1411, 1329-1481
37/3535417CB1/1383	1-254, 1-262, 1-267, 1-341, 1-387, 1-460, 1-462, 1-464, 1-501, 1-537, 1-561, 1-624, 1-629, 3-164, 3-506, 4-576, 140-334, 226-799, 404-676, 506-658, 581-872, 713-991, 784-1309, 799-1335, 807-1291, 843-1383, 866-1323, 949-1381
38/7473555CB1/3519	1-264, 184-1995, 1278-1431, 1279-1431, 1279-1578, 1279-3519, 1432-1578, 2240-2464, 2324-2605, 2427-3237, 2428-2820, 2428-3089, 2428-3106, 2428-3228, 2429-3154, 2429-3193, 2429-3234, 2465-2746, 2465-2747, 2606-2746, 3098-3429
39/7474985CB1/2763	1-2046, 1-2763

表 4-5

ポリスクレオサド SEQ ID NO/ Inceyte ポ リスクレオサド ID/ 配列長	配列断片
40/7475137CBI/1023	1-1023
41/5036986CBI/2046	1-98, 1-264, 5-437, 7-569, 7-701, 15-608, 27-271, 36-274, 36-497, 42-402, 42-635, 74-590, 78-735, 81-735, 94-602, 135-349, 135-490, 282-475, 282-496, 282-501, 284-682, 351-840, 441-702, 447-1206, 488- 1046, 491-916, 520-702, 566-1000, 629-1206, 637-906, 637-1030, 739- 1396, 773-1429, 852-1115, 875-1114, 875-1388, 890-1205, 905-1490, 907-1171, 918-1470, 919-1243, 933-1277, 939-1486, 939-1509, 958- 1583, 973-1215, 978-1603, 1004-1505, 1008-1555, 1023-1262, 1026- 1604, 1030-1583, 1038-1631, 1077-1706, 1082-1652, 1086-1322, 1086- 1333, 1086-1573, 1088-1709, 1097-1633, 1100-1690, 1100-1710, 1113- 1398, 1113-1704, 1117-1780, 1118-1406, 1119-1719, 1123-1653, 1127- 1576, 1130-1729, 1132-1415, 1133-1721, 1147-1554, 1147-1735, 1150- 1442, 1162-1695, 1164-1841, 1172-1735, 1192-1774, 1204-1405, 1206- 1741, 1209-1741, 1212-1469, 1214-1741, 1232-1569, 1232-1708, 1234- 1849, 1242-1449, 1247-1875, 1255-1841, 1266-1894, 1269-1842, 1273- 1710, 1289-1687, 1299-1877, 1308-1842, 1315-1806, 1325-1722, 1337- 1606, 1339-1852, 1340-1675, 1363-1899, 1365-1854, 1365-1892, 1372- 1940, 1377-1947, 1379-1764, 1379-1917, 1410-1940, 1411-1676, 1411- 1901, 1416-1683, 1416-1925, 1421-1976, 1424-1672, 1433-1984, 1447-1842, 1463-1980, 1464-2015, 1466-1618, 1466-1859, 1466-1881, 1466-1947, 1466-1999, 1466-2005, 1466-2015, 1467-1859, 1467-2011, 1475-2000, 1477-2000, 1478-2011, 1484-2046, 1487-1811, 1488-1735, 1492-1951, 1493-1953, 1494-2011, 1495-2006, 1497-1953, 1498-1982, 1499-2015, 1509-1736, 1519-1785, 1521-1783, 1540-1618, 1540-1953, 1542-2015, 1545-1946, 1550-2012, 1554-1893, 1559-2015, 1564-1879, 1565-1951, 1574-2004, 1579-1998, 1586-1949, 1590-1951, 1599-2015, 1613-1989, 1613-2015, 1614-1844, 1614-2015, 1615-1703, 1615-1995, 1615-2015, 1642-1908, 1655-1923, 1656-2015, 1659-2007, 1665-2006, 1665-2015, 1697-1965, 1702-2015, 1711-1949, 1713-1924, 1734-2007, 1739-1946, 1739-1957, 1744-2007, 1749-2015, 1782-2015, 1788-2015, 1801-2015, 1829-2015, 1845-2007,

10

20

30

表 4 - 6

ポリヌクレオチド SEQ ID NO/ Incyte ポリヌクレオチド ID/ 配列長	配列断片
42/1375644CBI/857	1-255, 1-559, 20-191, 20-204, 20-249, 20-266, 93-283, 93-385, 109-349, 109-377, 110-378, 110-380, 110-439, 111-261, 111-264, 111-367, 111-372, 114-357, 114-600, 116-390, 116-674, 119-251, 119-266, 119-375, 120-249, 123-413, 136-284, 136-358, 136-443, 136-447, 137-407, 140-382, 142-353, 142-358, 142-386, 142-410, 149-340, 151-383, 160-426, 160-446, 164-249, 171-832, 175-372, 177-402, 177-724, 179-287, 187-432, 187-443, 189-448, 189-518, 194-464, 196-825, 204-447, 211-443, 213-495, 218-467, 227-486, 235-394, 235-542, 235-771, 235-825, 241-834, 243-476, 244-691, 248-304, 249-391, 257-829, 258-500, 260-560, 263-522, 291-529, 294-654, 304-455, 304-555, 307-551, 309-575, 310-563, 310-848, 312-537, 312-839, 323-571, 323-599, 331-575, 331-854, 332-447, 332-572, 332-590, 332-857, 350-563, 350-572, 350-578, 350-585, 350-598, 350-630, 353-857, 356-731, 358-445, 358-614, 358-834, 358-836, 358-846, 359-857, 364-443, 367-423, 372-613, 373-844, 377-634, 379-443, 388-447, 406-529, 515-686, 603-857
43/1356261CBI/1485	1-183, 1-443, 1-672, 6-183, 8-183, 11-183, 14-183, 15-183, 17-553, 18-183, 25-175, 34-183, 36-183, 66-183, 122-183, 153-182, 212-363, 212-412, 212-415, 212-416, 213-478, 212-478, 212-617, 212-703, 215-434, 247-506, 249-532, 257-617, 304-401, 304-536, 324-617, 367-536, 376-554, 387-595, 387-596, 387-815, 413-689, 486-736, 488-989, 546-815, 546-1006, 558-818, 576-820, 591-870, 592-798, 592-828, 592-1106, 671-955, 684-968, 795-1265, 830-1059, 837-1098, 872-1436, 873-1209, 892-1060, 931-1437, 959-1124, 975-1213, 981-1441, 984-1440, 991-1439, 1005-1479, 1021-1270, 1035-1473, 1037-1174, 1037-1285, 1038-1435, 1047-1473, 1064-1352, 1067-1485, 1070-1478, 1070-1480, 1081-1473, 1086-1485, 1088-1478, 1102-1424, 1105-1477, 1119-1478, 1122-1249, 1122-1399, 1122-1428, 1122-1473, 1155-1474, 1157-1479, 1230-1473, 1234-1472, 1298-1473, 1300-1469, 1301-1435, 1379-1474
44/1419379CBI/989	1-205, 1-329, 1-278, 1-493, 2-326, 4-253, 23-283, 28-300, 87-233, 162-610, 162-671, 321-844, 457-700, 457-832, 457-841, 471-844, 485-565, 535-668, 540-785, 726-989, 728-989, 842-989, 881-989, 944-989

10

20

30

【表 1 8】

表 4-7

ボリスケレオチド SEQ ID NO/ Incyte ポ リスケレオチド ID/ 配列長	配列断片
45/3509272CBI/3886	1-451, 1-541, 1-581, 1-581, 1-541, 29-537, 110-688, 125-338, 147-524, 158-622, 190-664, 222-513, 234-495, 281-503, 281-650, 295-566, 321-1000, 322-781, 327-817, 347-574, 350-598, 354-782, 356-752, 370-896, 372-868, 375-950, 380-939, 400-934, 402-892, 404-989, 433-1108, 493-1049, 512-919, 531-831, 569-1065, 593-964, 593-1148, 594-886, 614-1080, 642-1128, 642-1282, 691-1234, 712-896, 728-994, 738-1144, 752-1382, 757-1340, 830-1498, 844-1431, 853-1081, 872-1384, 936-1442, 937-1470, 942-1559, 962-1471, 966-1473, 996-1521, 1002-1536, 1023-1366, 1026-1734, 1053-1372, 1059-1691, 1081-1372, 1102-1369, 1105-1154, 1112-1739, 1157-1410, 1161-1417, 1171-1542, 1190-1461, 1202-1454, 1249-1679, 1298-1488, 1302-1875, 1311-1868, 1336-1950, 1344-1868, 1355-1868, 1366-1868, 1428-1868, 1429-2022, 1579-2031, 1579-2181, 1620-1868, 1774-2532, 1791-2425, 1791-2435, 1858-2574, 1922-2553, 1930-2440, 1939-2391, 1955-2608, 1992-2240, 2034-2342, 2041-2191, 2041-2653, 2067-2633, 2096-2764, 2097-2519, 2100-2718, 2101-2371, 2104-2665, 2124-2668, 2185-2397, 2191-2414, 2191-2774, 2317-2526, 2317-2539, 2336-2621, 2392-2908, 2431-3068, 2542-2814, 2549-2840, 2591-2872, 2636-2845, 2636-2958, 2696-2935, 2732-3123, 2867-3063, 2867-3102, 2888-3497, 2939-3194, 2961-3251, 2993-3556, 3007-3689, 3174-3690, 3221-3886, 3299-3546, 3328-3540
46/708425CBI/2677	1-372, 1-439, 1-471, 1-498, 1-524, 1-624, 2-251, 20-653, 28-333, 28-344, 28-435, 31-342, 46-317, 46-592, 60-653, 485-1010, 511-1079, 553-1182, 559-1075, 592-1162, 599-892, 659-1084, 666-1064, 684-1329, 726-1395, 742-1413, 762-1220, 794-1174, 917-1007, 1001-1682, 1063-1644, 1086-1606, 1086-1678, 1093-1670, 1154-1426, 1181-1480, 1277-1566, 1277-1843, 1302-1531, 1302-1724, 1307-1514, 1400-2053, 1454-1788, 1454-1829, 1500-2096, 1501-1755, 1557-2072, 1558-1810, 1589-2016, 1630-1883, 1650-1853, 1662-2020, 1696-1934, 1697-2081, 1710-1953, 1752-1853, 1752-1961, 1760-2024, 1854-2082, 1854-2276, 1854-2338, 1869-2088, 1869-2123, 1869-2432, 1933-2398, 1963-2404, 1974-2385, 1986-2214, 1997-2634, 2030-2656, 2069-2242, 2069-2594, 2094-2850, 2096-2649, 2105-2356, 2111-2650, 2171-2677, 2207-2864, 2207-2866, 2219-2669, 2219-2677, 2254-2658, 2290-2677, 2338-2671, 2346-2604, 2400-2660, 2402-2667, 2432-2677, 2446-2677, 2489-2655, 2489-2674, 2489-2677, 2571-2667

10

20

30

表 4 - 8

ポリヌクレオチド SEQ ID NO/ Incyte ポリヌクレオチド ID/ 配列長	配列断片
47/7469966CB1/3031	1-606, 87-978, 227-601, 227-832, 504-1262, 656-1264, 713-1289, 854-1490, 859-1512, 988-1241, 988-1450, 988-1474, 988-1511, 988-1560, 988-1561, 1008-1562, 1014-1236, 1040-1573, 1047-1667, 1085-1739, 1098-1762, 1112-1442, 1156-1741, 1204-1643, 1242-1816, 1281-1305, 1295-1537, 1298-1555, 1356-1622, 1360-1611, 1376-1728, 1378-1819, 1418-2019, 1425-1949, 1427-1643, 1431-1904, 1450-2051, 1473-1600, 1475-2086, 1480-1739, 1493-2118, 1530-1800, 1557-1815, 1651-1914, 1657-2153, 1695-1967, 1695-2031, 1732-2082, 1735-2376, 1759-1988, 1788-2037, 1788-2286, 1794-2042, 1794-2054, 1794-2083, 1825-2392, 1842-2169, 1850-2490, 1864-2101, 1873-2197, 1900-2147, 1930-2146, 1984-2275, 2102-2267, 2130-2376, 2143-2416, 2156-2472, 2190-2421, 2190-2708, 2370-2633, 2299-2960, 2311-2586, 2353-2582, 2359-3021, 2369-2500, 2369-2949, 2369-2969, 2407-3020, 2409-2987, 2464-3007, 2552-2812, 2552-3003, 2575-2781, 2586-2816, 2586-2823, 2670-3031, 2687-3031, 2688-2934, 2698-3031, 2710-2984, 2710-3031, 2711-2927, 2714-3031, 2735-3031, 2801-3024, 2897-3031, 2902-3009
48/6920129CB1/2415	1-529, 1-530, 5-530, 36-530, 156-530, 398-613, 455-585, 455-645, 455-697, 461-613, 505-878, 545-1160, 614-1309, 641-1223, 688-1233, 760-1254, 771-1335, 772-1233, 780-1246, 814-1374, 866-1240, 877-1544, 885-1251, 890-1552, 902-1158, 903-1242, 919-1463, 948-1172, 953-1217, 1019-1287, 1019-1437, 1024-1463, 1026-1469, 1061-1607, 1088-1219, 1093-1690, 1118-1398, 1190-1748, 1200-1967, 1202-1802, 1223-1707, 1235-1913, 1237-1864, 1243-1736, 1244-1785, 1249-1850, 1260-1874, 1263-1772, 1267-1732, 1320-1763, 1353-1785, 1370-1944, 1371-2023, 1379-1923, 1389-1963, 1415-1972, 1431-2004, 1454-2083, 1461-1851, 1462-1980, 1502-2011, 1518-2071, 1524-1886, 1531-1737, 1531-2038, 1540-2126, 1541-1729, 1542-2014, 1566-1845, 1576-1722, 1578-1993, 1599-2215, 1615-1870, 1621-1921, 1631-2079, 1631-2080, 1658-2003, 1677-2234, 1698-2106, 1708-2375, 1715-2139, 1717-1925, 1739-2274, 1749-2096, 1837-2389, 1853-2362, 1874-2036, 1907-2415, 2040-2390, 2046-2412, 2131-2415, 2168-2373, 2248-2402, 2264-2402, 2277-2402

10

20

30

【表 2 0】

表 4-9

ボリスクレオチド SEQ. ID. No/ Incyte ポ リスクレオチド ID/ 配列長	配列断片
49/71514704CB1/1843	1-282, 1-376, 11-277, 88-371, 88-602, 88-622, 88-629, 88-633, 88-598, 102-836, 103-767, 142-178, 176-412, 258-341, 326-626, 448-928, 857-874, 681-1226, 687-782, 702-1421, 709-1259, 722-1350, 729-1306, 790-1040, 797-1120, 797-1370, 812-1047, 813-1280, 935-1074, 944-1345, 976-1243, 976-1346, 976-1540, 988-1556, 1038-1577, 1057-1580, 1059-1537, 1070-1223, 1096-1582, 1129-1843, 1132-1418, 1285-1353, 1439-1539, 1731-1769
50/7715945CB1/1827	1-610, 16-658, 16-695, 18-652, 136-579, 136-580, 136-614, 136-645, 136-749, 139-691, 169-673, 206-447, 228-635, 232-727, 240-441, 288-560, 304-515, 304-518, 316-642, 324-586, 324-618, 334-620, 337-573, 439-705, 443-944, 575-886, 575-1164, 596-878, 596-1275, 650-1133, 728-775, 781-1028, 892-1285, 915-1441, 1002-1301, 1035-1300, 1063-1640, 1063-1757, 1069-1280, 1069-1700, 1070-1356, 1136-1640, 1294-1574, 1319-1588, 1323-1583, 1390-1827, 1406-1651, 1458-1620, 1532-1827
51/7025368CB1/794	1-564, 1-658, 2-732, 2-745, 3-592, 3-643, 4-570, 116-725, 354-783, 583-794
52/7500954CB1/1539	1-246, 1-248, 1-273, 1-276, 1-608, 1-1539, 2-352, 76-789, 170-685, 171-673, 171-761, 171-842, 212-432, 287-549, 287-560, 287-840, 298-823, 304-875, 407-837, 407-881, 407-883, 407-930, 408-707, 416-652, 536-1182, 578-839, 654-956, 654-959, 723-1182, 1044-1256, 1045-1482, 1051-1489, 1060-1491, 1061-1489, 1063-1486, 1064-1486, 1069-1421, 1069-1489, 1070-1492, 1076-1539, 1087-1489, 1093-1341, 1094-1539, 1096-1484, 1097-1438, 1102-1488, 1102-1489, 1106-1489, 1114-1539, 1127-1220, 1127-1224, 1127-1235, 1127-1241, 1129-1241, 1131-1304, 1137-1485, 1138-1489, 1142-1486, 1142-1489, 1144-1539, 1146-1434, 1197-1484, 1197-1539, 1206-1400, 1211-1482, 1213-1539, 1219-1489, 1226-1493, 1228-1493, 1232-1486, 1331-1487, 1336-1539, 1389-1489, 1437-1539

10

20

30

表 5

ポリスクレオノイド SEQ ID NO:	Incyte プロジェクト ID	代表的ライブラリ
27	6242606CB1	MONOTXN03
28	997105CB1	KIDNTUT01
29	1482843CB1	BRAYDIN03
30	3218219CB1	LUNGNON07
31	7371678CB1	ARTANOT06
32	7473863CB1	HNT2TXC01
33	7478662CB1	PROSDIT01
34	7650474CB1	LUNGFET03
35	8092378CB1	FIBPFEN06
36	1263178CB1	BLADDIT01
37	3535417CB1	CONNTMT01
38	7473555CB1	BRAXDRI12
41	5036986CB1	BRATUT02
42	1375644CB1	LUNGNOT10
43	1356261CB1	ADRETUT06
44	1419379CB1	STOMTUT02
45	3509272CB1	LEUKNOT03
46	708425CB1	ENDANOT01
47	7489866CB1	THPLPLB01
48	6920129CB1	BEPINOT01
49	71514704CB1	BRAIDIT05
50	7715945CB1	PROSNOT19
51	7025368CB1	LIVRDI06
52	7500964CB1	MCLDTXT04

表 6-1

ライブラリ	ペクター	ライブラリの説明
AURETUT06	pINCY	ライブラリは57才の白人女性の片側右副腎摘出時に採取した副腎腫瘍組織から単離したRNAを用いて作製した。病理学は、褐色細胞腫を示した（結節性の増を形成しており、この塊は副腎の髄質を完全に置換していた）。
ARTANOT06	pINCY	ライブラリは48才の白人男子から採取した大動脈外膜組織から単離したRNAを用いて作製した。
BEPTW0101	PSPORT1	ライブラリは54才の白人男子に由来する気管支上皮初代細胞株から単離したRNAを用いて作製した。
BLADDIT01	pINCY	ライブラリは73才の男性の膀胱全摘時に取られた腫瘍膀胱組織から単離したRNAを用いて作製された。病理学は、膀胱癌に程度の慢性炎症を示した。随伴する腫瘍組織の病理は低悪度の腺癌を示した。腫瘍は脆い塊（friable mass）を近位尿道内で、遠位尿道切除線から14cmの位置に形成していた。腫瘍は固有筋の層に浸潤していたが、貫通してはいなかった。
BRAIDIT05	pINCY	ライブラリは2名の白人のブールから採取した腫瘍脂肪組織から単離したRNAを用いて作製した。1人のドナラーは多発性硬化症を有し、白質に急性の炎症性変化と広範な炎症の併発を伴っていた。脳組織は、脳白質と小脳とから得た。他のドナラーは多発性硬化症を有し、慢性的無症候性病変（silent lesions）を伴った。脳組織は、或る不特定領域から得た。
BRAITUT02	PSPORT1	ライブラリは、56才の白人男性の脳腫瘍病変の切除時に前頭葉から採取した脳腫瘍組織から単離したRNAを用いて作製した。病理学検査では、グレート2の低移行性副腎腫がみられた。患者の病歴は、グレート2の腎細胞癌、不眠症および慢性的気道閉塞などであった。家系歴には腎臓の悪性腫瘍が含まれている。
BRAITUR12	PCDNA2.1	このランダムライブラリ法によるライブラリは、胆管癌で死亡した55才白人女性から切除した前頭新皮質組織から単離したRNAを用いて作製した。病理は胆管癌全体に優勢な軽度の腫瘍組織があり、常時免疫白質および巨核に軸索球状体が散在しており、内嗅野および中脳皮質白質領域にいくぶん神経線維変化の散在がみられた。随伴腫瘍組織の病理は、残存腫瘍または再発性腫瘍を伴う高分化型の肝細胞癌組織を示した。患者の病歴は胆管癌、術後パッド、キラーリー症候群、胆汁性腹水、水胸症、脱水症、栄養不良、乏尿および急性腎不全が含まれる。前に受けた手術には胆管切除術および肝臓の85%の摘除がある。
BRAIDIN03	pINCY	精製したライブラリは、或る脳腫瘍ライブラリの870万個の独立性クローンから作製した。開始RNAは脳血管作で死亡した57才の白人男子から採取された脳神経根床下部組織から単離したRNAから作製した。患者の病歴はハンチントン病組織および肺気腫がある。このライブラリは、極めて長時間（48時間/一回）の再アニーリングによるハイブリダイゼーションを用いたことを除いては、Scarses 他、PNAS (1994) 91:9228 および Bonaldi 他、Genome Research (1996) 6:791 を改作した条件を用いて2回にわたってスクリーニングされた。このライブラリは、挿入断片を有するクローン用いて選択するために直線化した。再増強した。

表 6-2

ライブラリ	ベクター	ライブラリの説明
CONINT01	pINCY	ライブラリは33才の女性から採取した子宮軟部組織から単離したRNAを用いて作製した。病理は非腫瘍性軟部組織を示した。一致する腫瘍組織の病理学検査では、複数(3)の平滑筋腫が子宮底の後部と上端とから示された。
ENDANT01	PBLUESCRIPT	ライブラリは、心臓移植時に或る男性から切除した外種心臓からの大動脈内皮細胞組織から単離されたRNAを用いて作製された。
FIBPFEN06	pINCY	標準化された前立腺間質性線維芽細胞組織ライブラリ群は、或る前立腺線維芽細胞ライブラリからの156万個の独立クローンから作製した。開始RNAは妊娠26週間後死亡の男子胎児から採取した前立腺間質の線維芽細胞から作製した。このライブラリは、極めて長時間(48時間/1回)の再アニーリングによるハイブリダイゼーションを用いた他は、Soares 他, PNAS (1994) 91:9228 および Bonaldo 他, Genome Research (1996) 6:791 を改作した条件を用いて2回にわたって標準化した。このライブラリは次に、下記の挿入断片を有するクローン用に変換するために直線化し、再変化した。プラスミドDNAの調製は、約100万のクローンを(由来はノーマライズした前立腺間質性線維芽細胞組織ライブラリ)から行い、ゾフトアガー形質転換後に行った。
INT2TXC01	pINCY	ライブラリは、INT2細胞株(発生の或る早期段階における明確な神経前駆体の特性を示すヒト奇形癌腫に由来する)から単離したRNAを使用してStratageneにて作製した(ST937233)。細胞の処理を、マウスのレプシン(1マイクログラム/ml)と9-cisレチノイン酸(3.3μM)とで6日間おこなった。
KIDNTUT01	PSPORT1	ライブラリは8ヶ月の女児の腎尿管切除時に採取した腎臓癌組織から単離したRNAを用いて作製した。病理学検査では、Wilms腫瘍(腎芽細胞腫)を示した。腫瘍は腎実質の90%に及んでいた。術前に、患者はヘパリン抗凝固療法を受けていた。
LEUKNOT03	pINCY	ライブラリは血液型がA+の27才女性の白血球から単離されたRNAを用いて作製した。供与者は検査でサイトメガロウイルス(CMV)が陰性であった。
LIVRDT06	pINCY	ライブラリは、肝臓の針生検時に35才白人男性から取り除かれた罹患肝臓組織から単離されたRNAを用いて作製された。患者の病歴にはC型肝炎が含まれる。
LUNGFT03	pINCY	ライブラリは妊娠20週で死亡の白人女子胎児から採取した胎組織から単離したRNAを用いて作製された。
LUNGNON07	pINCY	この標準化した肺組織ライブラリは、或る肺組織ライブラリの510万個の独立クローンから作製した。開始RNAの作製は、肺組織から単離したRNAから行った。このライブラリは、極めて長時間(48時間/1回)の再アニーリングによるハイブリダイゼーションを用いた他は、Soares 他, PNAS (1994) 91:9228-9232 および Bonaldo 他, Genome Research (1996) 6:791 を適用した条件を用いて2回にわたってノーマライズ

10

20

30

【表 2 4】

表 6-3

ライブラリ	ベクター	ライブラリの説明
		した。
LENGNOT10	pINCY	ライブラリは妊娠 23 週で死亡した白人男子胎児から採取した肺組織から単離した RNA を用いて作製された。
MCLDTX04	pINCY	ライブラリは、1 男児から採取した脐帯血 CD34+前駆体細胞由来の処理された樹状細胞から単離した RNA を用いて作製した。これらの細胞は顆粒球/マクロファージコロニー刺激因子 (GM-CSF)、腫瘍壊死因子アルファ (TNF アルファ)、および幹細胞因子 (SCF) で処理し、次にホルボミルミリスチン酸アセテート (PMA) およびイノマイシンで処理した。GM-CSF は、時間 0、100 ng/ml で加え、TNF アルファは時間 0、2.5 ng/ml で加え、SCF は時間 0、25 ng/ml で加えた。PMA およびイノマイシンは 13 日に 5 時間加えた。インキュベーション時間は 13 日間であった。
MONOTIN03	pINCY	標準化した処理済み単球組織ライブラリを、或る処理済み単球組織ライブラリからの 760 万個の独立クローンから作製した。開始 RNA は、42 才の女性から得た末梢血の処理済みの単球から単離した RNA から作製した。細胞を抗インターロキン-10 (抗-IL-10) とリポ多糖 (LPS) で処理した。抗-IL-10 を時間 0 に 10 ng/ml 加え、LPS を 1 時間後に 5 ng/ml 加えた。単球をプラスチックに接着させることにより buffy コートから単離した。インキュベーション時間は 24 時間であった。cDNA 合成を開始するには NotI 阻着したオリゴ(dT)プライマーを用いた。これらのライブラリは、極めて長時間 (48 時間/1 回) の再アニーリングによるハイブリダイゼーションを用いた他は、Soares 他, PNAS (1994) 91:9228 および Bonaldi 他, Genome Research (1996 6):791 を改作した条件を用いて 2 回にわたり標準化した。これらのライブラリは次に、下記のように入断断片を有するクローニングに選択するために置換化し、再線状化した。プラスミド DNA の調製は、約 100 万のクローン (由来はノーマライズした処理済み単球組織ライブラリ) から、ソフトアガー形成転換後に行った。DNA は NotI で置換化し、挿入断片を有するクローニング群をサイズ選択するためにアガロースゲル電気泳動を行い、ライゲーションで再環状化した。
PROSDIT01	pINCY	このライブラリは、58 才白人男性の根治的前立腺摘出、所屬リンパ節切除、および前立腺針生検時に除去された罹患前立腺組織から単離された RNA を用いて作製された。病理学検査では腺癌腫増殖症が見られた。随伴する腫瘍組織の病理は、Gleason グレード 3+3 の腺癌を示した。腺癌は主要な塊を形成し、この塊は右側後上節前立腺を侵していた。腫瘍の別の微細的病変が、左側後下部に同定された。腫瘍は膀胱に浸潤していたが、それを越えてはいなかった。この患者は前立腺特異抗原 (PSA) 上昇、夜間多尿、血尿、および硬結を示した。患者の病歴には良性高血圧がある。家族歴には良性高血圧および前立腺癌がある。
PRUSNOT19	pINCY	ライブラリは、59 才白人男性の前立腺全摘出および所屬リンパ節切除時に切除された罹患前立腺組織から単離された RNA を用いて作製された。病理学検査では腺癌腫増殖症が見られた。随伴する腫瘍組織の病理学検査

10

20

30

【表 2 6】

表 6 - 4

ライブラリ	ベクター	ライブラリの説明
		では、腫瘍 (Gleason グレード 3 + 3) が見られた。この患者は前立腺特異抗原 (PSA) の上昇を示した。患者の病歴には結腸癌、アスベスト肺、血栓性静脈炎がある。以前に部分結腸切除を含む手術を受けている。家族歴には良性高血圧、多発性骨腫瘍、高脂血症および慢性関節リウマチが含まれている。
STOMT02	pINCY	ライブラリは、胃部分切除時に 68 オウ人女性から得た胃の癌組織から単離された RNA を用いて作製された。病理検査は、びまん性大細胞型 (diffuse large-cell type) の悪性リンパ腫を示した。以前の手術としては胆嚢摘出術があった。患者の病歴にはサラセミアが含まれる。家族の病歴には急性白血病、食道の悪性新生物、悪性胃新生物、アデノーム性冠状動脈疾患がある。
THP1L801	PBLUESCRIPT	ライブラリは、THP-1 培養細胞から単離した RNA を用いて作製した。培養は、48 時間、100 ng/ml のホルボール エステル (PMA) と共に行った後、4 時間の培養を 1 ug/ml の LPS を有する培地で行った。THP-1 は、急性単球 性白血球の 1 オの芽児の末梢血に由来するヒト前駆球株である (参照: Iml. J. Cancer (1980) 26:171)。

表 7-1

プログラム	説明	参考文献	パラメータ値
ABI FACTURA	ベクター配列を除き、核酸配列のまぎらわしい塩基をマスクするプログラム。	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	アミノ酸または核酸配列の比較、注釈に有用なファーストデータファインダ。	Applied Biosystems, Foster City, CA. Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler	核酸配列の構築を行うプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	アミノ酸および核酸配列の配列類似性検索に有用なベージングローカルアライメント検索ツール (Basic Local Alignment Search Tool)。BLASTには5つの機能がある: blastp、blastn、blastx、tblastnおよびtblastx。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) Nucleic Acids Res. 25:3389-3402.	EST: 確率値=1.0E-8 以下 完全長配列: 確率値=1.0E-10以下
FASTA	一群の同じタイプの配列と問合せ配列との間の類似性について検索するPearson およびLipman アルゴリズム。FASTAは最低5つの機能からなる: fasta、tfasta、fastx、tfastxおよびssearch。	Pearson, W.R. および D. J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; Smith, T.F. および M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	EST: fasta E 値=1.0E-6 構築された EST: fasta 同一性=95%以上 一致した長さ=200塩基以上 fasta E 値=1.0E-8以下 完全長配列: fastx スコア=100以上
BLIMPS	配列をBLOCKS、PRINTS、DOMO、PRODOM およびPFAM データベースの配列と対応させて遺伝子ファミリー、配列相溶性および構造的フィングアングリメント領域を検索するBLOCKS IMPROVED Searcher。	Henikoff, S. および J. G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J. G. および S. Henikoff (1996) Methods Enzymol. 266:88-105; Atwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	確率値=1.0E-3以下
HMMER	問合せ配列をPFAMのようなタンパク質ファミリーコンセンサス配列の隠れマルコフモデル (HMM) に基づいたデータベースに対して検索するアルゴリズム。	Krogh, A. 他 (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. 他 (1998) Nucleic Acids Res. 26:320-322; Durbin, R. 他 (1998) Our World View, in a Nutshell, Cambridge Univ. Press, 1-350 ページ。	PFAM ヒット 確率値=1.0E-3 以下 シグナル ペプチド ヒット: スコア=0 以上

10

20

30

表 7-2

プログラム	説明	参考文献	パラメータ閾値
ProfileScan	Prositeで定義された配列パターンと一致するタンパク質配列の構造的モチーフおよび配列モチーフを検索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4:61-66; Gribskov, M. 他 (1989) Methods Enzymol. 183:146-159; Baroch, A. 他 (1997) Nucleic Acids Res. 25:217-221.	標準化された質スコア≧特定の Prositeモチーフに対するGCO指 定「HIGH」値 通常、スコア=1.4~2.1.
Phred	高度の感受性および確率のある自動シーケンサー トレースを調べる塩基コールングアルゴリズム。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Twine, B. および P. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的実行に基づいたプログラムである。SWATとCrossMatchを含むPhrap改訂構築プログラムで、配列相同性検索やDNA配列の構築に有用である。	Smith, T. F. および M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T. F. および M. S. Waterman (1981) J. Mol. Biol. 147:195-197; Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さ=58以上
Consed	Phrap 構築の表示、編集用グラフィカルツール。	Gordon, D. 他 (1988) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキャンし、分泌シグナルペプチド の存在を調べる重量マトリックス分析プログラム。	Nielson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J. M. および S. Audic (1997) CABIOS 12:431-439.	スコア≧3.5以上
TMAP	重量マトリックスを使ってタンパク質配列上の膜貫 通部分を描写し、方向を決定するプログラム。	Persson, B. および P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. および P. Argos (1996) Protein Sci. 5:363-371.	
THEMMER	隠れマルコフモデル(HMM)を使ってタンパク質配列 上の膜貫通部分を描写し、方向を決定するプログラ ム。	Sonnhammer, E.L. 他 (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow 他, 編集, The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, 175-182-ページ。175-182.	
Motifs	Prositeで定義されたものと一致するバクテリアのアミ ノ酸配列を検索するプログラム。	Baroch, A. 他 (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, バージョン9, MSI-59-ページ, Genetics Computer Group, Madison, WI	

10

20

30

【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
22 August 2002 (22.08.2002)

PCT

(10) International Publication Number
WO 02/064792 A2(51) International Patent Classification: C12N 15/12,
C07K 14/47, C12N 5/10, C07K 16/18, C12Q 1/68, G01N
33/50, 33/53, A61K 38/00, 39/00

(21) International Application Number: PCT/US02/00254

(22) International Filing Date: 4 January 2002 (04.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/260,168 5 January 2001 (05.01.2001) US
60/262,857 19 January 2001 (19.01.2001) US
60/262,736 19 January 2001 (19.01.2001) US(71) Applicant (for all designated States except US): INCYTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(72) Inventors: and

(75) Inventors/Applicants (for US only): BAUGHN, Mariah,
R. [US/US]; 14244 Santiago Road, San Leandro, CA
94577 (US). WARREN, Bridget, A. [US/US]; 10130 Par-
wood Drive #2, Cupertino, CA 95014 (US). HONCHELL,
Cynthia, D. [US/US]; 400 Laurel Street #203, San Carlos,
CA 94070 (US). XU, Yuming [US/US]; 1739 Wal-
nut Drive, Mountain View, CA 94040 (US). WALLA,
Narinder, K. [US/US]; 890 Davis Street #205, San Leand-
ro, CA 94577 (US). RAMKUMAR, Jayalaxmi [IN/US];
34359 Maybird Circle, Fremont, CA 94555 (US). YAO,
Monique, G. [US/US]; 1180 Woodgate Drive, Carmel, IN
46033 (US). LU, Yan [CN/US]; 3885 Corrina Way, Palo
Alto, CA 94303 (US). YUE, Henry [US/US]; 826 Lois
Avenue, Sunnyvale, CA 94087 (US). THANGAVELU,
Kavitha [IN/US]; 1950 Montecito Avenue 23, Mountain
View, CA 94043 (US). TANG, Tom, Y. [CN/US]; 4230
Ranwick Court, San Jose, CA 95118 (US). DING, Li
[CN/US]; 3353 Alma Street, #146, Palo Alto, CA 94306
(US). BOROWSKY, Mark, L. [US/US]; 122 Orchard Av-
enue, Redwood City, CA 94061 (US). HAFALIA, April,
J., A. [US/US]; 2227 Calle de Primavera, Santa Clara,CA 95054 (US). LU, Dyung, Aina, M. [US/US]; 233
Coy Drive, San Jose, CA 95123 (US). AZIMZAI, Yalda
[US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA
94552 (US). TRAN, Bao [US/US]; 750 Saltberg Avenue,
Santa Clara, CA 95051 (US). NGUYEN, Daniel, B.
[US/US]; 1403 Ridgewood Drive, San Jose, CA 95118
(US). BURFORD, Neil [GB/US]; 105 Wildwood Circle,
Durham, CT 06422 (US). ISON, Graig, H. [US/US]; 1242
Weathersfield Way, San Jose, CA 95118 (US). GURURA-
JAN, Rajagopal [US/US]; 5591 Dent Avenue, San Jose,
CA 95118 (US). GANDHI, Ameena, R. [US/US]; 705
5th Avenue, San Francisco, CA 94118 (US). ELLIOTT,
Vicki, S. [US/US]; 3770 Politen Place Way, San Jose, CA
95121 (US). TRAN, Uyen, K. [US/US]; 2638 Mabury
Square, San Jose, CA 95133 (US).(74) Agents: HAMLET-COX, Diana et al.; Incyte Pharma-
ceuticals, Inc., 3160 Porter Drive, Palo Alto, CA 94304
(US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GI,
GM, GR, GU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MY, NZ, PA, PE, PG, PH, PL, PT, RO, RU, SD, SE,
SI, SK, SL, ST, SV, TM, TR, TT, TZ, UA, UG, US, UZ, VN,
YU, ZA, ZW.(84) Designated States (regional): ARIPO patent (GL, GM,
KL, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NI, SN, TD, TG).

Published:

— without international search report and to be republished
upon receipt of that reportFor two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

WO 02/064792 A2

(54) Title: MOLECULES FOR DISEASE DETECTION AND TREATMENT

(57) Abstract: The invention provides human molecules for disease detection and treatment (MSST) and polynucleotides which identify and encode MDDT. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of MDDT.

WO 02/064792

PCT/US02/00254

MOLECULES FOR DISEASE DETECTION AND TREATMENT

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of molecules for disease
5 detection and treatment and to the use of these sequences in the diagnosis, treatment, and prevention
of cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders, and in the
assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid
sequences of molecules for disease detection and treatment.

10 BACKGROUND OF THE INVENTION

It is estimated that only 2% of mammalian DNA encodes proteins, and only a small fraction of
the genes that encode proteins are actually expressed in a particular cell at any time. The various
types of cells in a multicellular organism differ dramatically both in structure and function, and the
identity of a particular cell is conferred by its unique pattern of gene expression. In addition, different
15 cell types express overlapping but distinctive sets of genes throughout development. Cell growth and
proliferation, cell differentiation, the immune response, apoptosis, and other processes that contribute
to organismal development and survival are governed by regulation of gene expression. Appropriate
gene regulation also ensures that cells function efficiently by expressing only those genes whose
functions are required at a given time. Factors that influence gene expression include extracellular
20 signals that mediate cell-cell communication and coordinate the activities of different cell types. Gene
expression is regulated at the level of DNA and RNA transcription, and at the level of mRNA
translation.

Aberrant expression or mutations in genes and their products may cause, or increase
susceptibility to, a variety of human diseases such as cancer and other cell proliferative disorders. The
25 identification of these genes and their products is the basis of an ever-expanding effort to find markers
for early detection of diseases and targets for their prevention and treatment. For example, cancer
represents a type of cell proliferative disorder that affects nearly every tissue in the body. The
development of cancer, or oncogenesis, is often correlated with the conversion of a normal gene into a
cancer-causing gene, or oncogene, through abnormal expression or mutation. Oncoproteins, the
30 products of oncogenes, include a variety of molecules that influence cell proliferation, such as growth
factors, growth factor receptors, intracellular signal transducers, nuclear transcription factors, and
cell-cycle control proteins. In contrast, tumor-suppressor genes are involved in inhibiting cell
proliferation. Mutations which reduce or abrogate the function of tumor-suppressor genes result in

WO 02/064792

PCT/US02/00254

aberrant cell proliferation and cancer. Thus a wide variety of genes and their products have been found that are associated with cell proliferative disorders such as cancer, but many more may exist that are yet to be discovered.

DNA-based arrays can provide an efficient, high-throughput method to examine gene
5 expression and genetic variability. For example, SNPs, or single nucleotide polymorphisms, are the most common type of human genetic variation. DNA-based arrays can dramatically accelerate the discovery of SNPs in hundreds and even thousands of genes. Likewise, such arrays can be used for SNP genotyping in which DNA samples from individuals or populations are assayed for the presence of selected SNPs. These approaches will ultimately lead to the systematic identification of all genetic
10 variations in the human genome and the correlation of certain genetic variations with disease susceptibility, responsiveness to drug treatments, and other medically relevant information. (See, for example, Wang, D.G. et al. (1998) Science 280:1077-1082.)

DNA-based array technology is especially important for the rapid analysis of global gene expression patterns. For example, genetic predisposition, disease, or therapeutic treatment may
15 directly or indirectly affect the expression of a large number of genes in a given tissue. In this case, it is useful to develop a profile, or transcript image, of all the genes that are expressed and the levels at which they are expressed in that particular tissue. A profile generated from an individual or population affected with a certain disease or undergoing a particular therapy may be compared with a profile generated from a control individual or population. Such analysis does not require knowledge of gene
20 function, as the expression profiles can be subjected to mathematical analyses which simply treat each gene as a marker. Furthermore, gene expression profiles may help dissect biological pathways by identifying all the genes expressed, for example, at a certain developmental stage, in a particular tissue, or in response to disease or treatment. (See, for example, Lander, E.S. et al. (1996) Science 274:536-539.)

25 Certain genes are known to be associated with diseases because of their chromosomal location, such as the genes in the myotonic dystrophy (DM) regions of mouse and human. The mutation underlying DM has been localized to a gene encoding the DM-kinase protein, but another active gene, DMR-N9, is in close proximity to the DM-kinase gene (Jansen, G. et al. (1992) Nat. Genet. 1:261-266). DMR-N9 encodes a 650 amino acid protein that contains WD repeats, motifs
30 found in cell signaling proteins. DMR-N9 is expressed in all neural tissues and in the testis, suggesting a role for DMR-N9 in the manifestation of mental and testicular symptoms in severe cases of DM (Jansen, G. et al. (1995) Hum. Mol. Genet. 4:843-852).

Other genes are identified based upon their expression patterns or association with disease

WO 02/064792

PCT/US02/00254

syndromes. For example, autoantibodies to subcellular organelles are found in patients with systemic rheumatic diseases. A recently identified protein, golgin-67, belongs to a family of Golgi autoantigens having alpha-helical coiled-coil domains (Eystathiou, T. et al. (2000) J. Autoimmun. 14:179-187). The Stac gene was identified as a brain specific, developmentally regulated gene. The Stac protein
 5 contains an SH3 domain, and is thought to be involved in neuron-specific signal transduction (Suzuki, H. et al. (1996) Biochem. Biophys. Res. Commun. 229:902-909).

The discovery of new molecules for disease detection and treatment, 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, autoimmune/inflammatory, developmental,
 10 and neurological disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of molecules for disease detection and treatment.

SUMMARY OF THE INVENTION

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

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

WO 02/064792

PCT/US02/00254

polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-26. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:27-52.

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

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

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

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

WO 02/064792

PCT/US02/00254

90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

5 Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, 10 c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said 15 probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a 20 polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said 25 target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected 30 from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an immunogenic fragment of

WO 02/064792

PCT/US02/00254

a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. The invention additionally provides a method of treating a disease or condition associated with decreased expression of

5 functional MDDT, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected

10 from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the

15 invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional MDDT, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as

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

25 and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of

30 treating a disease or condition associated with overexpression of functional MDDT, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid

WO 02/064792

PCT/US02/00254

sequence selected from the group consisting of SEQ ID NO:1-26, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

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

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

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52, ii) a

WO 02/064792

PCT/US02/00254

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

BRIEF DESCRIPTION OF THE TABLES

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

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability scores for the matches between each polypeptide and its homolog(s) are also shown.

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

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

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

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

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and

WO 02/064792

PCT/US02/00254

polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

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

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

Unless defined otherwise, all technical and scientific terms used herein have the same
15 meanings as commonly understood by one of ordinary skill in the art to which this invention belongs.
Although any machines, materials, and methods similar or equivalent to those described herein can be
used to practice or test the present invention, the preferred machines, materials and methods are now
described. All publications mentioned herein are cited for the purpose of describing and disclosing the
cell lines, protocols, reagents and vectors which are reported in the publications and which might be
20 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

"MDDT" refers to the amino acid sequences of substantially purified MDDT obtained from
any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and
25 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
MDDT. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other
compound or composition which modulates the activity of MDDT either by directly interacting with
MDDT or by acting on components of the biological pathway in which MDDT participates.

30 An "allelic variant" is an alternative form of the gene encoding MDDT. 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

WO 02/064792

PCT/US02/00254

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 MDDT include those sequences with deletions, 5 insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as MDDT or a polypeptide with at least one functional characteristic of MDDT. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding MDDT, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding MDDT. 10 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 MDDT. 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 MDDT is retained. For example, negatively charged amino acids may 15 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 hydrophobicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

20 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.

25 “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 MDDT. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small 30 molecules, or any other compound or composition which modulates the activity of MDDT either by directly interacting with MDDT or by acting on components of the biological pathway in which MDDT participates.

The term “antibody” refers to intact immunoglobulin molecules as well as to fragments

WO 02/064792

PCT/US02/00254

thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind MDDT polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "aptamer" refers to a nucleic acid or oligonucleotide molecule that binds to a specific molecular target. Aptamers are derived from an *in vitro* evolutionary process (e.g., SELEX (Systematic Evolution of Ligands by EXponential Enrichment), described in U.S. Patent No. 5,270,163), which selects for target-specific aptamer sequences from large combinatorial libraries. Aptamer compositions may be double-stranded or single-stranded, and may include deoxyribonucleotides, ribonucleotides, nucleotide derivatives, or other nucleotide-like molecules. The nucleotide components of an aptamer may have modified sugar groups (e.g., the 2'-OH group of a ribonucleotide may be replaced by 2'-F or 2'-NH₂), which may improve a desired property, e.g., resistance to nucleases or longer lifetime in blood. Aptamers may be conjugated to other molecules, e.g., a high molecular weight carrier to slow clearance of the aptamer from the circulatory system. Aptamers may be specifically cross-linked to their cognate ligands, e.g., by photo-activation of a cross-linker. (See, e.g., Brody, E.N. and L. Gold (2000) J. Biotechnol. 74:5-13.)

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

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

The term "antisense" refers to any composition capable of base-pairing with the "sense"

WO 02/064792

PCT/US02/00254

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

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

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

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

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

WO 02/064792

PCT/US02/00254

“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

5 conservative amino acid substitutions.

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

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation,

(b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

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

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

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

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

WO 02/064792

PCT/US02/00254

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

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

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

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

A fragment of SEQ ID NO:1-26 is encoded by a fragment of SEQ ID NO:27-52. A fragment of SEQ ID NO:1-26 comprises a region of unique amino acid sequence that specifically
30 identifies SEQ ID NO:1-26. For example, a fragment of SEQ ID NO:1-26 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-26. The precise length of a fragment of SEQ ID NO:1-26 and the region of SEQ ID NO:1-26 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the

WO 02/064792

PCT/US02/00254

intended purpose for the fragment.

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

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

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

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of
15 molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent
20 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
25 <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The
30 "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

WO 02/064792

PCT/US02/00254

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

Percent identity may be measured over the length of an entire defined sequence, for example,
 10 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
 15 length over which percentage identity may be measured.

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

20 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
 25 substitution, thus preserving the structure (and therefore function) of the polypeptide.

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

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

WO 02/064792

PCT/US02/00254

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

Matrix: BLOSUM62

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

WO 02/064792

PCT/US02/00254

1% (w/v) SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

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

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

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

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

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

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of MDDT which is

WO 02/064792

PCT/US02/00254

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 MDDT which is useful in any of the antibody production methods disclosed herein or known in the art.

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

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

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

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

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

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

"Post-translational modification" of an MDDT 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 MDDT.

"Probe" refers to nucleic acid sequences encoding MDDT, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target

WO 02/064792

PCT/US02/00254

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

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

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

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

WO 02/064792

PCT/US02/00254

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

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

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be 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, translation, or RNA stability.

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

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

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

WO 02/064792

PCT/US02/00254

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

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

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

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

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

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

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

WO 02/064792

PCT/US02/00254

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
5 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
10 the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an
15 "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to
20 another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a
25 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
30 least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length of one of the polypeptides.

WO 02/064792

PCT/US02/00254

THE INVENTION

The invention is based on the discovery of new human molecules for disease detection and treatment (MDDT), the polynucleotides encoding MDDT, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders.

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

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

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

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are molecules for disease detection and treatment. For example, SEQ ID NO:11 is 47% identical to *Escherichia coli* putative nucleotide-binding protein (GenBank ID g1789285) as determined by the Basic Local Alignment Search Tool (BLAST). (See

WO 02/064792

PCT/US02/00254

Table 2.) The BLAST probability score is $1.0e-63$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. Data from MOTIFS and BLAST analyses provide further corroborative evidence that SEQ ID NO:11 is a nucleotide-binding protein. In a further example, SEQ ID NO:15 is 57% identical to bovine xenobiotic/medium-chain fatty acid:CoA ligase form XL-III (GenBank ID g5070357) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is $6.7e-181$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:15 also contains an AMP-binding enzyme domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and BLAST analyses provide further corroborative evidence that SEQ ID NO:15 is an acyl CoA ligase with an AMP binding domain. In yet a further example, SEQ ID NO:23 is 40% identical from residues 1 to 268 to human NM23-H8 (GenBank ID g7580490) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The nm23 genes were discovered on the basis of their reduced expression in highly metastatic cell lines. Nm23 proteins from various species display nucleoside diphosphate (NDP) kinase activity and are involved in tissue development and differentiation (Freije, J.M. et al. (1998) Biochem. Soc. Symp. 63:261-271). The BLAST probability score is $1.8e-45$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:23 also contains a thioredoxin family active site and a nucleoside diphosphate kinases active site as determined by PROFILESCAN analysis. (See Table 3.) Data from BLIMPS, and MOTIFS analyses provide further corroborative evidence that SEQ ID NO:23 is an nm23 family NDP kinase. SEQ ID NO:1-10, SEQ ID NO:12-14, SEQ ID NO:16-22, and SEQ ID NO:24-26 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-26 are described in Table 7.

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

WO 02/064792

PCT/US02/00254

The polynucleotide fragments described in Column 2 of Table 4 may refer specifically, for example, to Incyte cDNAs derived from tissue-specific cDNA libraries or from pooled cDNA libraries. Alternatively, the polynucleotide fragments described in column 2 may refer to GenBank cDNAs or ESTs which contributed to the assembly of the full length polynucleotide sequences. In addition, the polynucleotide fragments described in column 2 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the polynucleotide fragments described in column 2 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the polynucleotide fragments described in column 2 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, a polynucleotide sequence identified as FL_XXXXXX_N₁_N₂_YYYY_N₃_N₄ represents a "stitched" sequence in which XXXXXX is the identification number of the cluster of sequences to which the algorithm was applied, and YYYY is the number of the prediction generated by the algorithm, and N_{1,2,3,4} if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the polynucleotide fragments in column 2 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, a polynucleotide sequence identified as FLXXXXXX_gAAAAA_gBBBBB_1_N is a "stretched" sequence, with XXXXXX being the Incyte project identification number, gAAAAA being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, gBBBBB being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and N referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, gBBBBB).

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

Prefix	Type of analysis and/or examples of programs
GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).

WO 02/064792

PCT/US02/00254

GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).
INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

5 In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in Table 4 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

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

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

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

The invention also encompasses a variant of a polynucleotide sequence encoding MDDT. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least
25 about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding MDDT. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:27-52 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ
30 ID NO:27-52. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of MDDT.

WO 02/064792

PCT/US02/00254

In addition, or in the alternative, a polynucleotide variant of the invention is a splice variant of a polynucleotide sequence encoding MDDT. A splice variant may have portions which have significant sequence identity to the polynucleotide sequence encoding MDDT, but will generally have a greater or lesser number of polynucleotides due to additions or deletions of blocks of sequence arising from alternate splicing of exons during mRNA processing. A splice variant may have less than about 70%, or alternatively less than about 60%, or alternatively less than about 50% polynucleotide sequence identity to the polynucleotide sequence encoding MDDT over its entire length; however, portions of the splice variant will have at least about 70%, or alternatively at least about 85%, or alternatively at least about 95%, or alternatively 100% polynucleotide sequence identity to portions of the polynucleotide sequence encoding MDDT. For example, a polynucleotide comprising a sequence of SEQ ID NO:52 is a splice variant of a polynucleotide comprising a sequence of SEQ ID NO:35. Any one of the splice variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of MDDT.

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

Although nucleotide sequences which encode MDDT and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring MDDT under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding MDDT 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 MDDT 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 MDDT and MDDT derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the

WO 02/064792

PCT/US02/00254

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 MDDT or any fragment thereof.

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

Methods for DNA sequencing are well known in the art and may be used to practice any of
10 the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is
15 automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences
20 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 MDDT may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences,
25 such as promoters and regulatory elements. For example, one method which may be employed,

WO 02/064792

PCT/US02/00254

ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

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

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

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

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter MDDT-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

WO 02/064792

PCT/US02/00254

oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent No. 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of MDDT, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding MDDT may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, MDDT itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of MDDT, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

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

WO 02/064792

PCT/US02/00254

In order to express a biologically active MDDT, the nucleotide sequences encoding MDDT or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding MDDT. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding MDDT. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding MDDT and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

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

A variety of expression vector/host systems may be utilized to contain and express sequences encoding MDDT. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509; Engelhard, E.K. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945; Takamatsu, N. (1987) *EMBO J.* 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; and

WO 02/064792

PCT/US02/00254

Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding MDDT. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding MDDT 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 MDDT 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 MDDT are needed, e.g. for the production of antibodies, vectors which direct high level expression of MDDT may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

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

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

WO 02/064792

PCT/US02/00254

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

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of MDDT in cell lines is preferred. For example, sequences encoding MDDT 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

WO 02/064792

PCT/US02/00254

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

5 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 MDDT is inserted within a marker gene sequence, transformed cells

10 containing sequences encoding MDDT can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding MDDT under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding MDDT and that express

15 MDDT 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 MDDT using either

20 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 MDDT is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See,

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

A wide variety of labels and conjugation techniques are known by those skilled in the art and

30 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 MDDT include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding MDDT, or any fragments thereof, may be cloned into a vector

WO 02/064792

PCT/US02/00254

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
 5 (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 MDDT may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein
 10 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 MDDT may be designed to contain signal sequences which direct secretion of MDDT through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the
 15 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
 20 (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding MDDT may be ligated to a heterologous sequence resulting in translation of a
 25 fusion protein in any of the aforementioned host systems. For example, a chimeric MDDT protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of MDDT 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
 30 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

WO 02/064792

PCT/US02/00254

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 MDDT encoding sequence and the heterologous protein sequence, so that MDDT may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, *supra*, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled MDDT 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.

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

In one embodiment, the compound thus identified is closely related to the natural ligand of MDDT, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) *Current Protocols in Immunology* 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which MDDT 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 MDDT, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E. coli*. Cells expressing MDDT or cell membrane fractions which contain MDDT are then contacted with a test compound and binding, stimulation, or inhibition of activity of either MDDT 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 MDDT, either in solution or affixed to a solid support, and detecting the binding of MDDT 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.

WO 02/064792

PCT/US02/00254

MDDT of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of MDDT. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for MDDT activity, wherein MDDT is combined with at least one test compound, and the activity of MDDT in the presence of a test compound is compared with the activity of MDDT in the absence of the test compound. A change in the activity of MDDT in the presence of the test compound is indicative of a compound that modulates the activity of MDDT. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising MDDT under conditions suitable for MDDT activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of MDDT 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 MDDT or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (*neo*; Capecchi, M.R. (1989) *Science* 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. et al. (1997) *Nucleic Acids Res.* 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding MDDT 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 MDDT 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

WO 02/064792

PCT/US02/00254

of a polynucleotide encoding MDDT 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 MDDT, e.g., by secreting MDDT 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 MDDT and molecules for disease detection and treatment. In addition, examples of tissues expressing MDDT can be found in Table 6. Therefore, MDDT appears to play a role in cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders. In the treatment of disorders associated with increased MDDT expression or activity, it is desirable to decrease the expression or activity of MDDT. In the treatment of disorders associated with decreased MDDT expression or activity, it is desirable to increase the expression or activity of MDDT.

Therefore, in one embodiment, MDDT 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 MDDT. 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; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma,

WO 02/064792

PCT/US02/00254

Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia.

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

In a further embodiment, a composition comprising a substantially purified MDDT in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent

WO 02/064792

PCT/US02/00254

a disorder associated with decreased expression or activity of MDDT including, but not limited to, those provided above.

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

In a further embodiment, an antagonist of MDDT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of MDDT. Examples of such disorders include, but are not limited to, those cell proliferative, autoimmune/inflammatory, developmental, and neurological disorders described above. In one aspect, an antibody which specifically binds MDDT 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 MDDT.

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

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

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

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with MDDT 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,

WO 02/064792

PCT/US02/00254

polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to MDDT have an amino acid sequence consisting of at least about 5 amino acids, and generally will
 5 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 MDDT amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

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

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

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population
 25 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 MDDT may also be generated. For example, such fragments include, but are not limited to, F(ab')₂ fragments produced by pepsin
 30 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.)

WO 02/064792

PCT/US02/00254

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 MDDT and its
 5 specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering MDDT epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for MDDT. Affinity is expressed as an association
 10 constant, K_a , which is defined as the molar concentration of MDDT-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 MDDT epitopes, represents the average affinity, or avidity, of the antibodies for MDDT. The
 15 K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular MDDT 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 MDDT-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 MDDT, preferably in active form, from the
 20 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
 25 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 MDDT-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*.)

30 In another embodiment of the invention, the polynucleotides encoding MDDT, 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

WO 02/064792

PCT/US02/00254

MDDT. 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 MDDT. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

- 5 In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral
- 10 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.)
- 15 In another embodiment of the invention, polynucleotides encoding MDDT 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
- 20 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
- 25 hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:11395-11399), hepatitis
- 30 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 MDDT expression or regulation causes disease, the expression of

WO 02/064792

PCT/US02/00254

MDDT 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 MDDT are treated by constructing mammalian expression vectors encoding MDDT and introducing these vectors by mechanical means into MDDT-deficient cells. Mechanical transfer technologies for use with cells *in vivo* or *ex vitro* include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J.-L. and H. R  c  pon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of MDDT include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). MDDT 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 MDDT 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 MDDT expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding MDDT under the control of an independent promoter or the retrovirus long

WO 02/064792

PCT/US02/00254

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

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding MDDT to cells which have one or more genetic abnormalities with respect to the expression of MDDT. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent No. 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding MDDT to target cells which have one or more genetic abnormalities with respect to the expression of MDDT. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing MDDT 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

WO 02/064792

PCT/US02/00254

been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent No. 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent No. 5,804,413 teaches the use of recombinant HSV d92
 5 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned
 10 herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to
 15 deliver polynucleotides encoding MDDT to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA,
 20 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 MDDT into the alphavirus genome in place of the capsid-coding region results in the production of a large number of MDDT-coding RNAs and the synthesis of high levels of MDDT in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a
 25 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 MDDT 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
 30 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

WO 02/064792

PCT/US02/00254

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
 5 been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of
 10 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 MDDT.

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

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA
 25 sequences encoding MDDT. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible
 30 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,

WO 02/064792

PCT/US02/00254

and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

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

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

WO 02/064792

PCT/US02/00254

human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

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

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

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

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

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

Compositions suitable for use in the invention include compositions wherein the active

WO 02/064792

PCT/US02/00254

ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

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

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

15 A therapeutically effective dose refers to that amount of active ingredient, for example MDDT or fragments thereof, antibodies of MDDT, and agonists, antagonists or inhibitors of MDDT, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED₅₀ (the dose therapeutically effective in 50% of the population) or LD₅₀ (the dose
20 lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no toxicity.
25 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
30 severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

WO 02/064792

PCT/US02/00254

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

DIAGNOSTICS

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

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

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding MDDT or closely related molecules may be used to identify nucleic acid sequences which encode MDDT. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a

WO 02/064792

PCT/US02/00254

conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding MDDT, allelic variants, or related sequences.

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

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

Polynucleotide sequences encoding MDDT may be used for the diagnosis of disorders associated with expression of MDDT. 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; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus,

WO 02/064792

PCT/US02/00254

systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucopolysaccharidosis, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; and a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Strausler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebellar hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, cataplexy, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia. The polynucleotide sequences encoding MDDT 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 MDDT expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding MDDT may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide

WO 02/064792

PCT/US02/00254

sequences encoding MDDT 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 MDDT in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of MDDT, 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 MDDT, 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 MDDT 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 MDDT, or a fragment of a polynucleotide complementary to the polynucleotide encoding MDDT, and will be employed under optimized conditions for identification of a specific gene or

WO 02/064792

PCT/US02/00254

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 MDDT 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 MDDT are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplicons in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isNP), 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 MDDT include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) *J. Immunol. Methods* 159:235-244; Duplax, C. et al. (1993) *Anal. Biochem.* 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor

WO 02/064792

PCT/US02/00254

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

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

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

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression

WO 02/064792

PCT/US02/00254

provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

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

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for

WO 02/064792

PCT/US02/00254

example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein
5 identification.

A proteomic profile may also be generated using antibodies specific for MDDT to quantify the levels of MDDT expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem.
10 270:103-111; Mendoe, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and
15 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
20 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.
25 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
30 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

WO 02/064792

PCT/US02/00254

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

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

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

WO 02/064792

PCT/US02/00254

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

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

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

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

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

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

The disclosures of all patents, applications and publications, mentioned above and below, including U.S. Ser. No. 60/260,168, U.S. Ser. No. 60/262,857, and U.S. Ser. No. 60/262,736, are

WO 02/064792

PCT/US02/00254

expressly incorporated by reference herein.

EXAMPLES

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

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

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPTMT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTMA plasmid (Invitrogen), PCMV-ICIS plasmid (Stratagene), pIGEN (Incyte Genomics, Palo Alto CA), pRARE (Incyte Genomics), or pINCY (Incyte Genomics), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

WO 02/064792

PCT/US02/00254

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, 5 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 10 cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

15 Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared 20 using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI 25 protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing 30 vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and

WO 02/064792

PCT/US02/00254

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

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

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID

WO 02/064792

PCT/US02/00254

NO:27-52. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 2.

IV. Identification and Editing of Coding Sequences from Genomic DNA

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

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

- Partial cDNA sequences were extended with exons predicted by the Genscan gene
- 30 identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information,

WO 02/064792

PCT/US02/00254

generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpi public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

15 "Stretched" Sequences

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

VI. Chromosomal Mapping of MDDT Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:27-52 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:27-52 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available

WO 02/064792

PCT/US02/00254

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.

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

15 In this manner, SEQ ID NO:28 was mapped to chromosome 1 within the interval from 137.6 to 149.0 centiMorgans; SEQ ID NO:30 was mapped to chromosome 9 within the interval from 50.3 to 75.8 centiMorgans; SEQ ID NO:35 was mapped to chromosome 4 within the interval from 61.0 to 62.7 centiMorgans; SEQ ID NO:37 was mapped to chromosome 4 within the interval from 145.3 to 170.9 centiMorgans; SEQ ID NO:38 was mapped to chromosome 4 within the interval from 81.9 to
20 115.1 centiMorgans; and SEQ ID NO:40 was mapped to chromosome 7 within the interval from 47.9 to 62.8 centiMorgans.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs
25 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
30 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:

WO 02/064792

PCT/US02/00254

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

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.

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

VIII. Extension of MDDT Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate

WO 02/064792

PCT/US02/00254

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

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

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

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

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector

WO 02/064792

PCT/US02/00254

(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.

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

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

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:27-52 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base
20 pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 µCi of [γ -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a
25 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 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
30 membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and

WO 02/064792

PCT/US02/00254

compared.

X. Microarrays

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

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

25 Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/ μ l oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/ μ l RNase inhibitor, 500 μ M dATP, 500 μ M dGTP, 500 μ M dTTP, 40 μ M dCTP, 40 μ M dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by *in vitro* transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one

WO 02/064792

PCT/US02/00254

with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85° C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μ l 5X SSC/0.2% SDS.

Microarray Preparation

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

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

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

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

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

Hybridization

Hybridization reactions contain 9 μ l of sample mixture consisting of 0.2 μ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65° C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly

WO 02/064792

PCT/US02/00254

larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

5 **Detection**

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

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

The sensitivity of the scans is typically calibrated using the signal intensity generated by a
20 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,
25 are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC
30 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

WO 02/064792

PCT/US02/00254

spectra) between the fluorophores using each fluorophore's emission spectrum.

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

XI. Complementary Polynucleotides

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

XII. Expression of MDDT

Expression and purification of MDDT is achieved using bacterial or virus-based expression systems. For expression of MDDT 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 MDDT upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of MDDT 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 MDDT by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect *Spodoptera frugiperda* (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, MDDT is synthesized as a fusion protein with, e.g., glutathione S-

WO 02/064792

PCT/US02/00254

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 MDDT at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified MDDT obtained by these methods can be used directly in the assays shown in Example XVI where applicable.

XIII. Functional Assays

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

WO 02/064792

PCT/US02/00254

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

XIV. Production of MDDT Specific Antibodies

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

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-MDDT activity by, for example, binding the peptide or MDDT to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring MDDT Using Specific Antibodies

Naturally occurring or recombinant MDDT is substantially purified by immunoaffinity chromatography using antibodies specific for MDDT. An immunoaffinity column is constructed by covalently coupling anti-MDDT 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 MDDT are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of MDDT (e.g., high ionic strength

WO 02/064792

PCT/US02/00254

buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/MDDT 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 MDDT is collected.

XVI. Identification of Molecules Which Interact with MDDT

5 MDDT, or biologically active fragments thereof, are labeled with ¹²⁵I Bolton-Hunter reagent. (See, e.g., Bolton, A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled MDDT, washed, and any wells with labeled MDDT complex are assayed. Data obtained using different concentrations of MDDT are used to calculate values for the number, affinity, and association of MDDT with the
10 candidate molecules.

Alternatively, molecules interacting with MDDT 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).

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

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention.

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

WO 02/064792

PCT/US02/00254

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
6242606	1	6242606CD1	27	6242606CB1
997105	2	997105CD1	28	997105CB1
1482843	3	1482843CD1	29	1482843CB1
3218219	4	3218219CD1	30	3218219CB1
7371678	5	7371678CD1	31	7371678CB1
7471883	6	7471883CD1	32	7471883CB1
7478662	7	7478662CD1	33	7478662CB1
7650474	8	7650474CD1	34	7650474CB1
8092378	9	8092378CD1	35	8092378CB1
1263178	10	1263178CD1	36	1263178CB1
3535417	11	3535417CD1	37	3535417CB1
7473555	12	7473555CD1	38	7473555CB1
7474985	13	7474985CD1	39	7474985CB1
7475137	14	7475137CD1	40	7475137CB1
5036986	15	5036986CD1	41	5036986CB1
1375644	16	1375644CD1	42	1375644CB1
1356261	17	1356261CD1	43	1356261CB1
1419379	18	1419379CD1	44	1419379CB1
3509272	19	3509272CD1	45	3509272CB1
708425	20	708425CD1	46	708425CB1
7469966	21	7469966CD1	47	7469966CB1
6920129	22	6920129CD1	48	6920129CB1
71514704	23	71514704CD1	49	71514704CB1
7715945	24	7715945CD1	50	7715945CB1
7025368	25	7025368CD1	51	7025368CB1
7500954	26	7500954CD1	52	7500954CB1

WO 02/064792

PCT/US02/00254

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
1	6243606CD1	g2706620	1.40E-45	NA14, nuclear autoantigen of 14 kDa, with coiled-coil domain [Homo sapiens] (Ramos-Morales, F. et al. (1998) J. Biol. Chem. 273:1634-1639)
10	1263178CD1	g9896486	2.10E-40	RF42 protein [Homo sapiens] (Maas, C. et al. (2000) Genomics 65:70-74)
11	3535417CD1	g13881165	6.00E-83	[H]Mycobacterium tuberculosis CDC1551 [LAO/VO transport system kinase
12	7473555CD1	g4185940	2.20E-22	[Human endogenous retrovirus K] envelope protein
13	7474955CD1	g11096305	3.70E-109	[Homo sapiens] (AF131010) NALP2 (contains NACHT, LRR, and P&D)
14	7475117CD1	g2072963	8.50E-126	[Homo sapiens] J140
15	5056986CD1	g5070337	6.50E-181	[Bos taurus] xenotropic/medium-chain fatty acid-CoA ligase form XI-L-III
16	1373644CD1	g1041648	4.50E-56	[Mus musculus] mdgl-1
22	6920129CD1	g7228476	9.40E-16	[Homo sapiens] dentin sialophosphoprotein precursor Gu, K., et al. (2000) Molecular cloning of a human dentin sialophosphoprotein gene. Eur. J. Oral Sci. 108, 35-42
23	71514704CD1	g7586490	1.80E-45	[Homo sapiens] NM23-H8 Postel, E.H. (1998) NM23-NDP kinase. Int. J. Biochem. Cell Biol. 30, 1291-1295
24	7715945CD1	g406263	2.60E-131	[Mus musculus] T10 Halford, S., et al. (1993) Isolation of a gene expressed during early embryogenesis from the region of 22q11 commonly deleted in DiGeorge syndrome. Hum. Mol. Genet. 2, 1577-1582
25	7025368CD1	g11065721	4.10E-39	[Homo sapiens] (AJ251832) 28kD interferon responsive protein

WO 02/064792

PCT/US02/00254

Table 3

SEQ ID NO.	Incite Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	6242660CD1	127	T52 T82 T115	N110	Paired box domain signature: Q42-S97 Nuclear autoantigen of 14 kDa PD174194; L2-S127	ProfileScan BLAST-PRODOM
2	997108CD1	258	S115 T72 T156 T173	N19 N127	Leucine zipper pattern: L2-L23, L9-L30	MOTIFS
3	1482843CD1	175	S7 S121 T171	N71	Signal peptide: M1-G25	HMMER
4	3218219CD1	1146	S35 S36 S56 S98 S121 S361 S450 S563 S585 S609 S653 S709 S817 S824 S908 S958 S1015 S1100 S1141 T27 T40 T65 T250 T294 T305 T362 T462 T512 T654 T713 T782 T926 T951 S19 S23 S45 S50 S280 S282 S329 T31 T244 T333 Y304 S13 S67 S164 S187 S201 S232 T17 T120 T144 S9 S26 S81 S143 S149 S191 T68 S13 S21 S87 S118 S150 S215 T80 T101 T181 T206 S101 S127 T45 T51 T53 T61 T118	Transmembrane domain: R338-H366, S895-P910 N-terminus is cytosolic	HMMER TMAP	
5	7371678CD1	351				
6	7473883CD1	286		N199		
7	7478662CD1	205		N74 N139	Transmembrane domain: D84-A108 N-terminus is non-cytosolic	TMAP
8	7650474CD1	233			Nuclear protein SNF7, putative copia-like transposable element PD009719; G11-A212	BLAST-PRODOM
9	8092378CD1	205		N88	Signal peptide: M11-S35	HMMER
					Signal peptide: M13-S35	HMMER

WO 02/064792

PCT/US02/00254

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
9					Transmembrane domain: V4-T31 N-terminus is non-cytosolic	TMAP
10	1263178CD1	304	S25 S30 S71 S81 S261 S268 S293 T56 T143 T272	N26 N54 N248 N259		
11	3535417CD1	418	S79 S113 S136 S291 S317 S328 S374 T70 T106 T158 T168 T223 T289 T356 T384 Y90		Transmembrane domain: Q248-L264 N-terminus is non-cytosolic	TMAP
					ATP-binding protein, transport, LAO/AO system, kinase, transference PD011981: I105 E373 mFL/frxC family: DM00103P2725453-225: R145-K315 ATP/GTP-binding site motif A (P-loop): G150 S157	BLAST-PRODOM
						BLAST-DBOM
						MOTIFS
12	7473555CD1	1117	S18 S125 S256 S260 S522 S619 S641 S823 S860 S880 S932 S992 S1061 S1096 S1103 T317 T365 T854	N214 N682 N856 N871	Transmembrane domains: G61-A89 A452-L478 G523-K550 N-terminus is non-cytosolic	TMAP
					env. retrovirus-related polypeptide, glycoprotein coat PD023510: C106 V296	BLAST-PRODOM
					ATP/GTP-binding site motif A (P-loop): A334 S361	MOTIFS

WO 02/064792

PCT/US02/00254

Table 3 (cont.)

SEQ ID NO.	Incye Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
13	7474983CD1	920	S75 S144 S289 S341 S384 S488 S548 S582 S730 S759 S880 S903 S908 T29 T112 T335 T340 T438 T482 T659 T681 T764 T791 T859 Y844	N96 N710	Leucine Rich Repeat (LRR): N819-N846, R762-Q789, K734-L757	HMMER-PFAM
					Transmembrane domains: V506-S526, C790-A811 N-terminus is non-cytosolic	TMAP
					Receptor, angiotensin/vasopressin AII/AVP PD156095: M412-E594	BLAST-PRODOM
					Ribonuclease inhibitor, leucine rich repeat, angiotensin receptor PD017636: L766-L854, I707-I814, W682-S759	BLAST-PRODOM
					ATP/GTP-binding site motif A (P-loop): G212-T219	MOTIFS
					Mitochondrial energy transfer proteins signature: P536-Q545	MOTIFS
14	7475137CD1	340	S6 S33 S123 S149 S166 S211 S292 S298 T90 T176 T181 T216 T243 T300 T310	N4 N14 N54 N174	L1 element, p40, p150 PD005182: M232-L340	BLAST-PRODOM
					L1 element, p40, p150 PD005183: T53-R142	BLAST-PRODOM
					L1 element, p40, p150 PD004752: M1-R48	BLAST-PRODOM
					L1 element, p40, p150 PD003772: L143-K231	BLAST-PRODOM

WO 02/064792

PCT/US02/00254

Table 3 (cont.)

SEQ ID NO:	Incye Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
15	5036986CD1	577	S19 S176 S184 S272 S325 S388 S476 S495 S571 T259 T353 T375 T406 T555 Y147 Y235 Y455	N467	Transmembrane domains: V111-T137, M257-S276 N-terminus is non-cytosolic	TMAP
					AMP-binding enzyme: N82-V493 Putative AMP-binding domain: F220-Y235	HMMER-PFAM BLIMPS-BLOCKS
					Ligase, synthetase, biosynthesis, antibiotic, phosphopantetheine, AcylCoA PD000070: D396-I494, W81-A281, T137-T366	BLAST-PRODOR
					SA protein, gene signal, kidney specific PD151238: V40-W81	BLAST-PRODOR
					Putative AMP-binding domain: DM00073IA6120965-538: E58-Q393, G411-R564 DM00073IP3906250-555: A68-R564 DM00073IP27095107-644: I218-R560, F83-I266 DM00073IP1491232-533: V72-R564	BLAST-DOMO
					Cell attachment sequence RGD: R105-D107, R438-D440	MOTIFS
					Putative AMP-binding domain signature: I218-K229	MOTIFS
16	1375644CD1	119	S22 S97 Y105		Signal peptide: M1-A62 Peroxidase signature: S441-86 SEC17 OCR1, intergenic region PD020820: Y21-E106	SPScan ProSiteScan BLAST-PRODOR

WO 02/064792

PCT/US02/00254

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
17	135626(CD1)	389	S2 S73 S161 S274 S312 S377 T41 T79 T131 T149 T324 T325 T335 T376 Y342 T3 T13 T30 T81		Transmembrane domain: A162-Q177 N-terminus is cytosolic	TMAP
18	1419379(CD1)	266	S147 S241 S245 S285 S292 S322 S327 S429 S498 S571 S593 S594 S598 S634 S635 T18 T42 T71 T255 T276 T279 T333 T373 T398 T404 T461 T540 T652	N40 N60 N347 N514	Signal peptide: M1-G25	SPSCAN
					Transmembrane domain: Y539-C567 N-terminus is non-cytosolic	TMAP
20	708425(CD1)	205	S14 S18 S40 S69 S140 S148 S152 S189 T2 T42 T193	N83 N123 N127 N136	Neuroconin U proteins signature BL00967: S580-R590	BLIMPS BLOCKS MOTIFS
21	746996(CD1)	833	S40 S97 S111 S143 S189 S208 S324 S415 S434 S473 S482 S542 S545 S621 S707 S807 T38 T148 T202 T350 T394 T532 T551 T604 T634 T775 Y136 Y149 Y253 Y424	N18 N705		MOTIFS

WO 02/064792

PCT/US02/00254

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
22	6920129CD1	655	S2 S51 S87 S109 S110 S114 S117 S129 S136 S150 S173 S197 S217 S224 S233 S264 S331 S336 S341 S370 S378 S383 S387 S402 S403 S432 S437 S468 S481 S506 S513 S522 S554 S561 S603 S610 S617 S636 S637 S641 S647 S651 T258 T325 T326 T476 T541 T545 T569 T591 Y40	N77 N220 N262 N589	Ubiquitin carboxyl-terminal hydrolase family: G34-R67	HMMER, PFAM
23	71514704CD1	330	S3 S15 S25 S135 S156 S264 S297 S301 T30 T108 T191 T196 T198 T233 T244	N13	ACIDIC SERINE CLUSTER REPEAT DM04746[P4177]1-386: S303-P627 DM04746[S5757]1-646: K320-S647, A70-P586 Signal peptide: M1-A68	BLAST, DOMO SPSCAN
					Thioredoxin family protein signature BL00194: V31-K43 Thioredoxin family active site: T16-L61	BLIMPS, BLOCKS PROFILES CAN
					Nucleoside diphosphate kinases active site: A257-D304 Thioredoxin family signature PR00421: T30-W38, W38-S47, R38-A89	PROFILES CAN BLIMPS, PRINTS

WO 02/064792

PCT/US02/00254

Table 3 (cont.)

SEQ ID NO.	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
23					KINASE DIPHOSPHATE NUCLEOSIDE TRANSFERASE NDK NDP ATP BINDING PROTEIN I PRECURSOR PD001018: T161-A307 NUCLEOSIDE DIPHOSPHATE KINASES DM00773P505891-136; T161-P296 DM00773P438021-140; T161-K303 DM00773P242331-141; T161-P295 DM00773P152661-143; T161-P295 Thioredoxin family active site: Y31-P49 Regulator of chromosome condensation (RCC1) signature M222-L232	BLAST_PRODOM
24	7715945CD1	276	S48 S65 S110 S192 S246 S263 T98 T128 T231 T233 Y116		PROTEIN REGION F186.I SER/THR/RICH T10 DCCR ASNPTRH1 INTERGENIC PD024111: M11134, D132-L274	BLAST_PRODOM
25	7025368CD1	232	T26 T90 T176	N74 N110 N210		MOTIFS
26	7500954CD1	120	S16 S42 T33	N3		MOTIFS

WO 02/064792

PCT/US02/00254

Table 4

Polynucleotide SEQ ID NO./ Insert ID/ Sequence Length	Sequence Fragments
27/624266CB1/1675	1-266, 147-266, 181-352, 181-432, 181-500, 181-665, 219-353, 267-373, 268-467, 269-315, 281-328, 299-664, 300-488, 313-563, 400-675
28/997105CB1/1387	1-584, 364-629, 364-677, 373-999, 384-626, 384-641, 385-627, 386-666, 398-604, 398-677, 400-844, 407-585, 407-1024, 445-677, 446-1071, 452-677, 459-677, 502-677, 637-926, 677-1203, 759-996, 759-1002, 759-1012, 868-1387, 948-1249
29/1482843CB1/1534	1-249, 29-328, 49-711, 66-328, 82-430, 87-638, 109-710, 115-719, 138-768, 245-724, 274-637, 281-521, 281-823, 282-691, 319-672, 341-597, 344-597, 373-508, 404-687, 434-687, 462-724, 468-717, 482-1089, 497-1020, 533-1153, 562-1351, 563-866, 563-1163, 621-1437, 686-1437, 690-1437, 692-1337, 698-843, 713-1337, 720-1337, 732-1437, 748-1437, 749-1330, 773-1437, 789-1426, 789-1437, 829-1330, 836-1437, 838-1426, 843-1335, 859-1163, 860-1437, 872-1385, 883-1437, 935-1534, 962-1532, 970-1534, 972-1534, 983-1534, 994-1147, 1031-1407, 1038-1409, 1047-1469, 1052-1324, 1053-1337, 1089-1396, 1093-1347, 1117-1337, 1120-1407, 1125-1407, 1127-1407, 1130-1407, 1136-1407, 1145-1373, 1164-1407, 1172-1407, 1191-1437, 1210-1337, 1239-1525, 1242-1335, 1286-1414, 1282-1534, 1292-1534, 1302-1534, 1417-1517, 1438-1520, 1442-1495, 1443-1506, 1458-1517
30/3218219CB1/4203	1-660, 98-655, 182-463, 182-790, 184-402, 192-752, 194-727, 197-458, 211-782, 211-808, 216-463, 216-623, 220-554, 221-767, 295-731, 375-971, 388-971, 389-622, 390-971, 402-891, 413-971, 448-971, 476-971, 496-971, 500-971, 501-1246, 505-971, 518-971, 521-971, 528-971, 542-971, 546-971, 560-971, 697-971, 866-1603, 967-1619, 1115-1637, 1133-1650, 1161-1398, 1178-1677, 1273-1849, 1456-2083, 1498-1960, 1506-1988, 1508-1717, 1624-2096, 1704-2352, 1738-2373, 1847-2352, 1908-2315, 1914-2352, 1952-2200, 2024-2545, 2050-2702, 2127-2405, 2136-2347, 2378-2571, 2395-2701, 2395-2899, 2401-2960, 2553-2907, 2558-2812, 2559-2999, 2590-2848, 2590-3131, 2650-3272, 2871-3105, 2957-3218, 2985-3193, 2985-3248, 2985-3507, 3010-3286, 3026-3179, 3034-3343, 3111-3281, 3117-3382, 3154-3492, 3157-3313, 3157-3417, 3160-3227, 3160-3431, 3160-3637, 3220-3653, 3220-3775, 3222-3508, 3224-3357, 3237-3481, 3261-3531, 3454-3706, 3454-3716, 3490-3776, 3493-3697, 3512-3766, 3513-3728, 3513-3737, 3513-4033, 3519-3692, 3602-3889, 3615-3836, 3615-4047, 3616-3859, 3646-3784, 3659-3912, 3699-3961, 3752-3876, 3752-3999, 3752-4203, 3824-4097, 3829-4114, 3836-4061, 3918-4169, 3933-4203

WO 02/064792

PCT/US02/00254

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
31/7371678CB/2082	1-622, 90-715, 104-575, 156-270, 161-412, 197-787, 197-799, 283-845, 287-805, 345-920, 392-1104, 502-705, 502-1114, 557-793, 557-843, 557-957, 557-1076, 557-1077, 557-1182, 557-1182, 557-1192, 557-1204, 560-1004, 595-1202, 608-1077, 610-768, 618-1115, 621-1015, 662-971, 690-1306, 704-1133, 727-896, 741-1135, 756-1135, 787-1065, 822-1391, 823-935, 841-1212, 891-1018, 898-992, 905-1542, 916-979, 960-1098, 979-1292, 995-1352, 1013-1596, 1023-1263, 1031-1624, 1046-1641, 1080-1596, 1118-1818, 1153-1748, 1171-1758, 1177-1637, 1189-1702, 1207-1259, 1246-1901, 1253-1816, 1277-1587, 1278-1875, 1282-1380, 1303-1837, 1318-2032, 1360-1875, 1377-1943, 1377-2080, 1402-1733, 1439-1679, 1443-1829, 1445-1749, 1459-1698, 1468-1778, 1469-1896, 1490-2081, 1493-2082, 1509-2082, 1552-2082, 1559-2082, 1577-1884, 1584-2082, 1593-1771, 1597-2082, 1611-2082, 1628-2051, 1652-1814, 1667-1954, 1667-2065, 1686-2027, 1694-2073, 1707-2010, 1752-2082
32/7473883CB/1459	1-154, 1-157, 1-255, 1-487, 5-157, 6-157, 8-157, 9-157, 9-650, 10-157, 11-157, 13-157, 15-157, 17-124, 17-157, 18-109, 18-157, 19-157, 20-157, 25-157, 26-153, 26-157, 27-157, 28-157, 113-730, 157-225, 157-234, 157-241, 157-244, 157-246, 157-288, 157-303, 157-315, 157-316, 157-327, 157-342, 157-354, 157-385, 157-400, 157-406, 157-423, 157-449, 157-458, 157-465, 157-486, 157-514, 157-559, 157-749, 258-805, 284-858, 292-853, 297-577, 330-578, 336-599, 340-610, 348-771, 349-643, 349-794, 362-826, 370-1011, 381-855, 422-687, 423-897, 437-671, 437-840, 448-909, 464-555, 465-889, 466-1059, 474-1116, 501-744, 501-810, 514-1092, 555-1128, 565-1085, 577-826, 589-870, 589-878, 589-918, 605-1130, 605-1161, 615-815, 617-839, 621-821, 633-876, 635-1105, 662-997, 662-1241, 690-1459, 753-1021, 760-1294, 802-1164, 888-1035, 903-1297, 907-1196, 931-1069, 996-1269, 1421-1456, 1422-1459, 1423-1456
33/7478662CB/1806	1-309, 1-599, 15-557, 20-564, 27-275, 27-282, 32-511, 39-639, 89-710, 270-824, 296-803, 378-642, 378-937, 445-692, 452-993, 461-729, 523-1068, 847-1082, 850-1485, 887-1159, 974-1216, 974-1455, 974-1466, 974-1526, 974-1540, 983-1261, 1023-1532, 1023-1626, 1024-1342, 1049-1337, 1070-1806, 1094-1359, 1094-1363, 1119-1372, 1138-1803, 1185-1802, 1196-1806, 1199-1806, 1208-1805, 1208-1806, 1213-1806, 1224-1803, 1226-1553, 1229-1803, 1283-1803, 1288-1806, 1297-1806, 1370-1806, 1415-1777, 1560-1806, 1589-1806, 1594-1806, 1619-1806, 1646-1806, 1674-1806, 1685-1806, 1702-1806, 1758-1806

WO 02/064792

PCT/US02/00254

Table 4

Polynucleotide SEQ ID NO./ Inventor/Sequence Length	Sequence Fragments
347/630474CB/1/1698	1-244, 1-259, 1-292, 1-457, 1-459, 1-477, 1-541, 35-234, 35-542, 35-552, 35-598, 35-608, 59-303, 60-402, 60-307, 80-753, 82-353, 110-740, 113-721, 116-579, 120-536, 134-645, 135-598, 159-397, 160-880, 165-658, 180-684, 182-898, 233-852, 242-898, 251-537, 251-644, 258-498, 258-791, 261-871, 263-622, 265-840, 287-840, 307-899, 330-899, 336-663, 336-847, 345-967, 371-878, 376-909, 404-976, 454-1058, 499-954, 508-1010, 535-1118, 608-1198, 617-757, 643-1170, 660-930, 692-931, 692-1072, 704-1263, 705-1412, 751-1088, 760-1401, 761-1022, 788-1276, 790-1092, 815-1050, 815-1092, 815-1218, 840-1036, 841-1092, 851-1056, 856-1092, 906-1037, 907-1037, 918-1037, 998-1430, 1059-1120, 1059-1173, 1059-1203, 1089-1634, 1092-1327, 1092-1623, 1118-1399, 1121-1401, 1121-1598, 1121-1682, 1144-1414, 1172-1407, 1206-1695, 1217-1695, 1225-1698, 1228-1698, 1233-1690, 1239-1695, 1240-1695, 1245-1485, 1250-1476, 1253-1513, 1253-1696, 1256-1675, 1274-1697, 1276-1695, 1278-1692, 1280-1694
35/8092378CB/1/1753	1-276, 1-608, 2-352, 76-789, 170-673, 170-685, 170-761, 170-842, 212-432, 287-549, 287-560, 287-840, 304-875, 405-930, 407-837, 408-707, 416-652, 626-1291, 712-1419, 718-1333, 741-1407, 776-1378, 809-1403, 925-1436, 925-1710, 988-1210, 988-1434, 988-1457, 990-1169, 997-1682, 1110-1691, 1137-1432, 1150-1399, 1161-1409, 1161-1693, 1161-1753, 1197-1456, 1221-1659, 1237-1616, 1238-1527, 1238-1531, 1258-1464, 1258-1470, 1265-1704, 1307-1555
36/1263178CB/1/1489	1-326, 22-297, 22-624, 22-678, 57-97, 171-529, 173-707, 276-1190, 501-776, 523-1414, 607-779, 607-1073, 607-1098, 607-1104, 607-1105, 607-1108, 607-1145, 607-1224, 608-1190, 610-960, 610-1186, 611-746, 612-1157, 693-1358, 704-1125, 751-1400, 832-1488, 849-1489, 854-1425, 857-1425, 906-1173, 906-1285, 922-1489, 933-1410, 955-1489, 983-1105, 1021-1383, 1164-1411, 1329-1481
37/3535417CB/1/1383	1-254, 1-262, 1-267, 1-341, 1-387, 1-460, 1-462, 1-464, 1-501, 1-537, 1-561, 1-624, 1-629, 3-164, 3-506, 4-576, 140-334, 226-799, 404-676, 506-658, 581-872, 713-991, 784-1309, 799-1333, 807-1291, 843-1383, 886-1323, 949-1381
38/7473555CB/1/3519	1-264, 184-1995, 1278-1431, 1279-1431, 1279-1578, 1279-3519, 1432-1578, 2240-2464, 2324-2605, 2427-3237, 2428-2870, 2428-3089, 2428-3106, 2428-3228, 2429-3154, 2429-3193, 2429-3234, 2465-2746, 2465-2747, 2606-2746, 3098-3429
39/7474985CB/1/2763	1-2046, 1-2763
40/7475137CB/1/1023	1-1023

WO 02/064792

PCT/US02/00254

Table 4

Polynucleotide SEQ ID NO./ Inventor ID/Sequence Length	Sequence Fragments
41/5036986CB/2046	1-98, 1-264, 5-437, 7-569, 7-701, 15-608, 27-271, 36-274, 36-497, 42-402, 42-635, 74-590, 78-735, 81-735, 94-602, 135-349, 135-490, 282-475, 282-486, 282-501, 284-682, 351-840, 441-702, 447-1206, 488-1046, 491-916, 520-702, 566-1000, 629-1206, 637-906, 637-1030, 739-1396, 773-1429, 832-1115, 875-1388, 890-1205, 905-1490, 907-1171, 918-1470, 919-1243, 933-1277, 939-1486, 939-1509, 938-1583, 973-1215, 978-1603, 1004-1505, 1008-1555, 1023-1262, 1026-1604, 1030-1583, 1038-1631, 1077-1706, 1082-1652, 1086-1322, 1086-1333, 1086-1573, 1088-1709, 1097-1633, 1100-1690, 1100-1710, 1113-1398, 1113-1704, 1117-1780, 1118-1406, 1119-1719, 1123-1653, 1127-1576, 1130-1729, 1132-1415, 1133-1721, 1147-1554, 1147-1735, 1150-1442, 1162-1695, 1164-1841, 1172-1735, 1192-1774, 1204-1405, 1206-1741, 1209-1741, 1212-1469, 1214-1741, 1232-1569, 1232-1708, 1234-1849, 1242-1449, 1247-1875, 1255-1841, 1266-1894, 1269-1842, 1273-1710, 1289-1687, 1299-1877, 1308-1842, 1315-1806, 1325-1722, 1337-1606, 1339-1852, 1340-1675, 1363-1899, 1365-1854, 1365-1892, 1372-1940, 1377-1947, 1379-1764, 1379-1917, 1410-1940, 1411-1676, 1411-1901, 1416-1683, 1416-1923, 1421-1976, 1424-1672, 1433-1984, 1447-1842, 1463-1980, 1464-2015, 1466-1618, 1466-1859, 1466-1881, 1466-1947, 1466-1999, 1466-2005, 1466-2015, 1467-1859, 1467-2011, 1475-2000, 1477-2000, 1478-2011, 1484-2046, 1487-1811, 1488-1735, 1492-1951, 1493-1953, 1494-2011, 1495-2006, 1497-1953, 1498-1982, 1499-2015, 1509-1726, 1519-1785, 1521-1783, 1540-1618, 1540-1953, 1542-2015, 1545-1946, 1550-2012, 1554-1893, 1559-2015, 1564-1879, 1565-1951, 1574-2004, 1579-1998, 1586-1949, 1590-1951, 1599-2015, 1613-1989, 1613-2015, 1614-1844, 1614-2015, 1615-1703, 1615-1995, 1615-2015, 1642-1908, 1655-1973, 1656-2015, 1659-2007, 1665-2006, 1665-2015, 1697-1965, 1702-2015, 1711-1949, 1713-1924, 1734-2007, 1739-1946, 1739-1957, 1744-2007, 1749-2015, 1782-2015, 1788-2015, 1801-2015, 1829-2015, 1845-2007

WO 02/064792

PCT/US02/00254

Table 4

Polymer SEQ ID NO./ Inventor/Sequence Length	Sequence Fragments
42/1375644CB/857	1-255, 1-559, 20-191, 20-204, 20-249, 20-266, 93-283, 93-385, 109-349, 109-377, 110-378, 110-380, 110-439, 111-261, 111-264, 111-367, 111-372, 114-357, 114-600, 116-390, 116-674, 119-251, 119-266, 119-375, 120-249, 123-413, 136-284, 136-338, 136-443, 136-447, 137-407, 140-382, 142-353, 142-354, 142-358, 142-410, 149-340, 151-383, 160-426, 160-446, 164-249, 171-832, 175-372, 177-402, 177-724, 179-287, 187-432, 187-443, 189-448, 189-518, 194-464, 196-825, 204-447, 211-443, 213-495, 218-467, 227-486, 235-394, 235-542, 235-771, 235-825, 241-834, 243-476, 244-691, 248-304, 249-391, 257-599, 258-500, 260-560, 263-522, 291-529, 294-654, 304-455, 304-555, 309-575, 310-563, 310-848, 312-537, 312-839, 323-571, 323-599, 331-575, 331-854, 332-447, 332-572, 332-590, 332-857, 350-563, 350-572, 350-578, 350-585, 350-598, 350-630, 353-857, 356-731, 358-445, 358-614, 358-834, 358-846, 359-857, 364-443, 367-423, 372-613, 373-844, 377-634, 379-443, 388-447, 406-529, 515-666, 603-857
43/1356261CB/1485	1-183, 1-443, 1-672, 6-183, 8-183, 11-183, 14-183, 15-183, 17-553, 19-183, 25-175, 34-183, 36-183, 66-183, 122-183, 153-182, 212-363, 212-412, 212-415, 212-416, 212-478, 212-617, 212-703, 215-424, 247-506, 249-532, 257-617, 304-401, 304-536, 324-617, 367-536, 376-554, 387-595, 387-596, 387-815, 413-689, 488-736, 488-989, 546-815, 546-1006, 558-818, 576-820, 591-870, 592-798, 592-828, 592-1106, 671-955, 684-968, 795-1265, 830-1059, 837-1098, 872-1436, 873-1209, 892-1060, 931-1437, 959-1124, 975-1213, 981-1441, 984-1440, 991-1439, 1005-1479, 1021-1270, 1035-1473, 1037-1174, 1037-1295, 1038-1425, 1047-1473, 1064-1392, 1067-1485, 1070-1478, 1070-1480, 1081-1473, 1086-1485, 1088-1478, 1102-1424, 1105-1477, 1119-1478, 1122-1249, 1122-1399, 1122-1428, 1122-1473, 1155-1474, 1157-1479, 1230-1473, 1234-1472, 1298-1473, 1300-1469, 1301-1435, 1379-1474
44/1419379CB/1989	1-205, 1-229, 1-278, 1-493, 2-326, 4-253, 23-283, 28-300, 87-233, 162-610, 162-671, 321-844, 457-693, 457-700, 457-832, 457-841, 471-844, 485-565, 535-668, 540-785, 726-989, 728-989, 842-989, 881-989, 944-989

WO 02/064792

PCT/US02/00254

Table 4

Polynucleotide SEQ ID NO./ Incyte ID/ Sequence Length	Sequence Fragments
457509272CB1/3886	1-451, 1-541, 1-581, 13-541, 29-537, 110-688, 125-338, 147-524, 158-622, 190-664, 222-513, 234-495, 281-503, 281-650, 295-566, 321-1000, 323-781, 327-817, 347-574, 350-598, 354-782, 356-752, 370-896, 372-868, 375-930, 380-939, 400-934, 402-892, 404-989, 433-1108, 493-1049, 512-919, 531-831, 569-1065, 593-964, 593-1148, 594-886, 614-1080, 642-1128, 642-1282, 691-1234, 712-896, 728-992, 738-1144, 752-1382, 757-1340, 830-1498, 844-1431, 853-1081, 872-1384, 936-1442, 937-1470, 942-1559, 962-1471, 966-1473, 996-1521, 1002-1536, 1023-1366, 1026-1734, 1053-1372, 1058-1691, 1081-1372, 1102-1569, 1105-1154, 1112-1739, 1157-1410, 1161-1417, 1171-1542, 1190-1461, 1202-1454, 1249-1679, 1298-1488, 1302-1875, 1311-1868, 1336-1950, 1344-1868, 1355-1868, 1366-1868, 1428-1868, 1429-2022, 1579-2031, 1579-2181, 1620-1868, 1774-2532, 1791-2425, 1791-2435, 1858-2574, 1922-2533, 1930-2440, 1939-2391, 1955-2608, 1992-2240, 2034-2342, 2041-2191, 2041-2653, 2067-2633, 2096-2764, 2097-2519, 2100-2718, 2101-2371, 2104-2665, 2124-2668, 2155-2397, 2191-2414, 2191-2774, 2317-2526, 2317-2539, 2336-2621, 2392-2908, 2431-3068, 2542-2814, 2549-2840, 2591-2872, 2636-2845, 2636-2958, 2696-2935, 2732-3123, 2867-3063, 2867-3102, 2888-3497, 2939-3194, 2961-3251, 2993-3556, 3007-3689, 3174-3690, 3221-3886, 3299-3546, 3328-3540
46708425CB1/2677	1-372, 1-439, 1-471, 1-498, 1-524, 1-624, 2-251, 20-653, 28-333, 28-344, 28-435, 31-342, 46-317, 46-592, 60-653, 485-794, 1174, 917-1007, 1001-1682, 1083-1644, 1086-1606, 1086-1678, 1093-1670, 1154-1426, 1181-1480, 1277-1506, 1277-1843, 1302-1531, 1302-1724, 1307-1514, 1400-2053, 1454-1788, 1454-1829, 1500-2096, 1501-1755, 1557-2072, 1558-1810, 1589-2016, 1630-1883, 1650-1853, 1662-2020, 1696-1934, 1697-2081, 1710-1953, 1752-1853, 1752-1961, 1760-2024, 1854-2082, 1854-2276, 1854-2338, 1869-2088, 1869-2123, 1869-2432, 1933-2398, 1963-2404, 1974-2385, 1986-2214, 1997-2634, 2030-2656, 2069-2242, 2069-2594, 2094-2650, 2096-2649, 2105-2356, 2111-2650, 2171-2677, 2207-2664, 2207-2666, 2219-2669, 2219-2677, 2254-2658, 2290-2677, 2338-2671, 2346-2604, 2400-2660, 2402-2667, 2432-2677, 2446-2677, 2489-2653, 2489-2674, 2489-2677, 2571-2667

WO 02/064792

PCT/US02/00254

Table 4

Polynucleotide SEQ ID NO./ Inve ID/ Sequence Length	Sequence Fragments
477/46996GCB/13031	1-606, 87-878, 227-601, 227-832, 504-1262, 656-1264, 713-1298, 854-1490, 839-1512, 988-1241, 988-1460, 988-1474, 988-1511, 988-1560, 988-1561, 1008-1562, 1014-1236, 1040-1573, 1047-1667, 1085-1739, 1098-1762, 1112-1442, 1156-1741, 1204-1643, 1242-1816, 1281-1905, 1295-1537, 1298-1555, 1356-1622, 1360-1611, 1376-1728, 1378-1819, 1418-2019, 1425-1949, 1427-1643, 1431-1904, 1450-2031, 1473-1600, 1475-2086, 1480-1739, 1493-2118, 1530-1800, 1557-1815, 1651-1914, 1657-2153, 1695-1967, 1695-2031, 1732-2082, 1735-2376, 1759-1988, 1788-2037, 1788-2386, 1794-2042, 1794-2054, 1794-2083, 1825-2392, 1842-2169, 1850-2490, 1864-2101, 1873-2197, 1900-2147, 1930-2146, 1984-2275, 2002-2267, 2130-2376, 2143-2416, 2156-2472, 2190-2421, 2190-2708, 2270-2633, 2299-2960, 2311-2586, 2353-2582, 2359-3021, 2369-2500, 2369-2949, 2369-2969, 2407-3020, 2409-2987, 2464-3007, 2552-2812, 2552-3003, 2575-2781, 2586-2816, 2586-2823, 2670-3031, 2687-3031, 2688-2934, 2698-3031, 2710-2984, 2710-3031, 2711-2927, 2714-3031, 2735-3031, 2801-3024, 2897-3031, 2902-3009
48/6920129CBI/2415	1-529, 1-530, 5-530, 36-530, 398-613, 455-585, 455-645, 455-697, 461-613, 505-878, 545-1160, 614-1309, 641-1223, 688-1233, 760-1254, 771-1335, 772-1233, 780-1246, 814-1374, 866-1240, 877-1544, 885-1251, 890-1552, 902-1588, 902-1242, 919-1463, 948-1172, 953-1217, 1019-1287, 1019-1437, 1024-1463, 1026-1469, 1061-1607, 1088-1219, 1093-1690, 1118-1398, 1190-1748, 1200-1967, 1202-1802, 1223-1707, 1235-1913, 1237-1864, 1243-1736, 1244-1785, 1249-1850, 1260-1874, 1263-1772, 1267-1732, 1320-1763, 1353-1785, 1370-1944, 1371-2023, 1379-1923, 1389-1953, 1415-1972, 1431-2004, 1454-2083, 1461-1851, 1462-1980, 1502-2011, 1518-2071, 1524-1886, 1531-1737, 1531-2038, 1540-2126, 1541-1729, 1542-2014, 1566-1845, 1576-1722, 1578-1993, 1599-2215, 1615-1870, 1621-1921, 1631-2079, 1631-2080, 1638-2003, 1677-2234, 1698-2106, 1708-2375, 1715-2139, 1717-1923, 1739-2274, 1749-2096, 1837-2389, 1853-2362, 1874-2036, 1907-2415, 2040-2390, 2046-2412, 2131-2415, 2168-2373, 2248-2402, 2264-2402, 2277-2402
497/151470CBI/1843	1-282, 1-376, 11-277, 88-602, 88-602, 88-629, 88-633, 88-698, 102-836, 103-767, 142-178, 176-412, 258-341, 326-626, 448-928, 657-874, 681-1226, 687-782, 702-1421, 709-1259, 722-1350, 729-1306, 790-1040, 797-1120, 797-1370, 812-1047, 813-1280, 935-1074, 944-1345, 976-1243, 976-1346, 976-1540, 988-1556, 1038-1577, 1057-1580, 1059-1537, 1070-1223, 1096-1582, 1129-1843, 1132-1418, 1285-1353, 1439-1539, 1731-1769

WO 02/064792

PCT/US02/00254

Table 4

Polynucleotide SEQ ID NO./ Insite ID/ Sequence Length	Sequence Fragments
50/7715945CB/1827	1-610, 16-658, 16-695, 18-652, 136-579, 136-580, 136-614, 136-646, 136-749, 139-691, 169-673, 206-447, 228-635, 232-727, 240-441, 288-560, 304-515, 304-518, 316-642, 324-586, 324-618, 334-620, 337-573, 439-705, 443-944, 575-886, 575-1164, 596-878, 596-1275, 650-1133, 728-775, 781-1028, 892-1265, 915-1441, 1002-1301, 1035-1300, 1063-1640, 1063-1757, 1069-1280, 1069-1700, 1070-1356, 1136-1640, 1294-1574, 1319-1588, 1323-1583, 1390-1827, 1406-1651, 1458-1620, 1532-1827
51/7025368CB/1794	1-564, 1-658, 2-732, 2-745, 3-592, 3-643, 4-570, 116-725, 354-783, 583-794
52/7500954CB/1539	1-246, 1-248, 1-273, 1-276, 1-1539, 2-352, 76-789, 170-685, 171-673, 171-761, 171-842, 212-432, 287-549, 287-560, 287-840, 298-823, 304-875, 407-837, 407-881, 407-883, 407-930, 408-707, 416-652, 536-1182, 578-839, 654-956, 654-959, 723-1182, 1044-1256, 1045-1482, 1051-1489, 1060-1491, 1061-1489, 1063-1486, 1064-1486, 1069-1421, 1069-1489, 1070-1492, 1076-1539, 1087-1489, 1093-1341, 1094-1539, 1096-1484, 1097-1438, 1102-1488, 1102-1489, 1106-1489, 1114-1539, 1127-1220, 1127-1224, 1127-1233, 1127-1241, 1129-1241, 1131-1304, 1137-1483, 1139-1489, 1142-1486, 1142-1489, 1144-1539, 1146-1489, 1148-1434, 1197-1484, 1197-1539, 1206-1400, 1211-1482, 1213-1539, 1219-1489, 1226-1493, 1228-1493, 1232-1486, 1331-1487, 1336-1539, 1389-1489, 1437-1539

WO 02/064792

PCT/US02/00254

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID:	Representative Library
27	6242606CB1	MONOTXN03
28	997105CB1	KIDNTUT01
29	1482843CB1	BRAYDIN03
30	3218219CB1	LUNGNON07
31	7371678CB1	ARTANOT06
32	7473883CB1	HNT2TXC01
33	7478662CB1	PROSDIT01
34	7650474CB1	LUNGFET03
35	8092378CB1	FIBPFEN06
36	1263178CB1	BLADDIT01
37	3535417CB1	CONNTMT01
38	7473555CB1	BRAXTDR12
41	5036986CB1	BRAITUT02
42	1375644CB1	LUNGNOT10
43	1356261CB1	ADRETUT06
44	1419379CB1	STOMTUT02
45	3509272CB1	LEUKNOT03
46	708425CB1	ENDANOT01
47	7469966CB1	THP1PLB01
48	6920129CB1	BEPINOT01
49	71514704CB1	BRAIDIT05
50	7715945CB1	PROSNOT19
51	7025368CB1	LIVRDIT06
52	7500954CB1	MCLDITX04

WO 02/064792

PCT/US02/00254

Table 6

Library	Vector	Library Description
ADRETUT06	pINCY	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 57-year-old Caucasian female during a unilateral right adrenalectomy. Pathology indicated pheochromocytoma, forming a nodular mass completely replacing the medulla of the adrenal gland.
ARTANGUT06	pINCY	Library was constructed using RNA isolated from aortic adventitia tissue removed from a 48-year-old Caucasian male.
BEPNOT01	PSPORT1	Library was constructed using RNA isolated from a bronchial epithelium primary cell line derived from a 34-year-old Caucasian male.
BLADDIT01	pINCY	Library was constructed using RNA isolated from diseased bladder tissue removed from a 73-year-old male during a total cystectomy. Pathology indicated the bladder mucosa showed mild chronic cystitis. Pathology for the associated tumor tissue indicated invasive grade 3 adenocarcinoma, which formed a friable mass situated within the proximal urethra, 14 cm from the distal urethral resection margin. The tumor invaded superficially into, but not through, muscularis propria.
BRAIDT05	pINCY	Library was constructed using RNA isolated from diseased brain tissue removed from a pool of two donors. One donor had multiple sclerosis with acute, active lesions and widespread inflammatory involvement in the white matter, and the brain tissue was obtained from cerebral white matter and cerebellum. The other donor had multiple sclerosis with chronic, silent lesions, and the brain tissue was obtained from an unspecified region.
BRATUT02	PSPORT1	Library was constructed using RNA isolated from brain tumor tissue removed from the frontal lobe of a 58-year-old Caucasian male during excision of a cerebral meningioma. Pathology indicated a grade 2 metastatic hypernephroma. Patient history included a grade 2 renal cell carcinoma, insomnia, and chronic airway obstruction. Family history included a malignant neoplasm of the kidney.
BRAXTDR12	PCDNA2.1	This random primed library was constructed using RNA isolated from frontal neocortex tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominantly over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the periaqueductal gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.

WO 02/064792

PCT/US02/00254

Table 6

Library	Vector	Library Description
BRAYDIN03	pINCY	This normalized library was constructed from 6.7 million independent clones from a brain tissue library. Starting RNA was made from RNA isolated from diseased hypothalamus tissue removed from a 57-year-old Caucasian male who died from a cerebrovascular accident. Patient history included Huntington's disease and emphysema. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48-hours/round) reannealing hybridization was used. The library was linearized and recircularized to select for insert containing clones.
CONNTMT01	pINCY	Library was constructed using RNA isolated from uterine soft tissue removed from a 33-year-old female. Pathology indicated non tumorous soft tissue. Pathology for the matched tumor tissue indicated multiple (3) leiomyomas from the posterior and the top of the fundus.
ENDANO701	PBLUESCRIPT	Library was constructed using RNA isolated from aortic endothelial cell tissue from an explanted heart removed from a male during a heart transplant.
FIBPFEN06	pINCY	The normalized prostate stromal fibroblast tissue libraries were constructed from 1.56 million independent clones from a prostate fibroblast library. Starting RNA was made from fibroblasts of prostate stroma removed from a male fetus, who died after 26 weeks' gestation. The libraries were normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48-hours/round) reannealing hybridization was used. The library was then linearized and recircularized to select for insert containing clones as follows: plasmid DNA was prepared from approximately 1 million clones from the normalized prostate stromal fibroblast tissue libraries following soft agar transformation.
HNT2TXC01	pINCY	Library was constructed at Stratagene (STR07233), using RNA isolated from the hNT2 cell line, which was derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor at an early stage of development. Cells were treated with mouse leptin at 1 microgram/ml and 9-cis retinoic acid at 3.3 microM for 6 days.
KIDNTUT01	PSFORT1	Library was constructed using RNA isolated from the kidney tumor tissue removed from an 8-month-old female during nephroureterectomy. Pathology indicated Wilms' tumor (nephroblastoma), which involved 90 percent of the renal parenchyma. Prior to surgery, the patient was receiving heparin anticoagulant therapy.
LEUKNO703	pINCY	Library was constructed using RNA isolated from white blood cells of a 27-year-old female with blood type A+. The donor tested negative for cytomegalovirus (CMV).
LIVRDT06	pINCY	This library was constructed using RNA isolated from diseased liver tissue removed from a 35-year-old Caucasian male during needle biopsy of the liver. Patient history included hepatitis C.

WO 02/064792

PCT/US02/00254

Table 6

Library	Vector	Library Description
LUNGFE103	pINCY	Library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation.
LUNGNO07	pINCY	This normalized lung tissue library was constructed from 5.1 million independent clones from a lung tissue library. Starting RNA was made from RNA isolated from lung tissue. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
LUNGNOT10	pINCY	Library was constructed using RNA isolated from the lung tissue of a Caucasian male fetus, who died at 23 weeks' gestation.
MCLDTXT04	pINCY	Library was constructed using RNA isolated from treated, derived dendritic cells from umbilical cord blood CD34+ precursor cells removed from a male. The cells were derived with granulocyte/macrophage colony stimulating factor (GM-CSF), tumor necrosis factor alpha (TNF alpha), and stem cell factor (SCF), then treated with phorbol myristate acetate (PMA), and Ionomycin. The GM-CSF was added at time 0 at 100 ng/ml, the TNF alpha was added at time 0 at 2.5 ng/ml, and the SCF was added at time 0 at 25 ng/ml. The PMA and Ionomycin were added at 13 days for five hours. Incubation time was 13 days.
MONOTXN03	pINCY	Normalized, treated monocyte tissue library was constructed from 7.6 million independent clones from a treated monocyte library. Starting RNA was made from RNA isolated from treated monocytes from peripheral blood obtained from a 42-year old female. The cells were treated with anti-interleukin-10 (anti-IL-10) and lipopolysaccharide (LPS). The anti-IL-10 was added at time 0 at 10 ng/ml and LPS was added at 1 hour at 5ng/ml. The monocytes were isolated from buffy coat by adherence to plastic. Incubation time was 24 hours. cDNA synthesis was initiated using a NotI-anchored oligo(dT) primer. The libraries were normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used. The libraries were then linearized and recircularized to select for insert containing clones as follows: plasmid DNA was prepped from approximately 1 million clones from the normalized, treated monocyte tissue libraries following soft agar transformation. The DNA was linearized with NotI and insert containing clones were size-selected by agarose gel electrophoresis and recircularized by ligation.

WO 02/064792

PCT/US02/00254

Table 6

Library	Vector	Library Description
PROSDT01	pINCY	The library was constructed using RNA isolated from diseased prostate tissue removed from a 58-year-old Caucasian male during radical prostatectomy, regional lymph node excision, and prostate needle biopsy. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated adenocarcinoma Gleason grade 3+3, which formed a predominant mass involving the right posterior superior prostate. Another microscopic focus of tumor was identified in the left posterior inferior. The tumor invaded the capsule but did not extend beyond it. The patient presented with elevated prostate specific antigen (PSA), nocturia, hematuria, and incontinence. Patient history included benign hypertension. Family history included benign hypertension and prostate cancer.
PROSNOT19	pINCY	Library was constructed using RNA isolated from diseased prostate tissue removed from a 59-year-old Caucasian male during a radical prostatectomy with regional lymph node excision. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma (Gleason grade 3+3). The patient presented with elevated prostate-specific antigen (PSA). Patient history included colon diverticuli, asbestosis, and thrombophlebitis. Previous surgeries included a partial colectomy. Family history included benign hypertension, multiple myeloma, hyperlipidemia and rheumatoid arthritis.
STOMTUT02	pINCY	Library was constructed using RNA isolated from stomach tumor tissue obtained from a 68-year-old Caucasian female during a partial gastrectomy. Pathology indicated a malignant lymphoma of diffuse large-cell type. Previous surgeries included cholecystectomy. Patient history included thalassemia. Family history included acute leukemia, malignant neoplasm of the esophagus, malignant stomach neoplasm, andtherosclerotic coronary artery disease.
THP1PLB01	PBLUESCRIPT	Library was constructed using RNA isolated from THP-1 cells cultured for 48 hours with 100 ng/ml phorbol ester (PMA), followed by a 4-hour culture in media containing 1 ug/ml LPS. THP-1 is a human promonocyte line derived from the peripheral blood of a 1-year-old male with acute monocytic leukemia (ref: Int. J. Cancer (1980) 26:171).

WO 02/064792

PCT/US02/00254

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL PDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, fastn, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.00E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fasta E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLOCKS IMPROVED Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Altwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMIMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nusselt, Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater

WO 02/064792

PCT/US02/00254

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribkov, M. et al. (1988) CABIOS 4:61-66; Gribkov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores; GCG-specified "HICP" value for that particular Prosite motif Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequence traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green. (1998) Genome Res. 8:186-194.	
Phrap	A Phrap Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consef	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:165-202.	
SFScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielsen, H. et al. (1997) Protein Engineering 10:1-6; Clavette, J.M. and S. Audic. (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sommadossi, E.L. et al. (1998) Proc. Sixth Int. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59; Genetics Computer Group, Madison, WI.	

WO 02/064792

PCT/US02/00254

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26,
 - b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-26,
 - c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, and
 - d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.
2. An isolated polypeptide of claim 1 comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.
3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method of producing a polypeptide of claim 1, the method comprising:
 - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant

WO 02/064792

PCT/US02/00254

polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

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

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

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

- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:27-52,
- b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ
- 15 ID NO:27-52,
- c) a polynucleotide complementary to a polynucleotide of a),
- d) a polynucleotide complementary to a polynucleotide of b), and
- e) an RNA equivalent of a)-d).

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

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

- 25 a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
- 30 b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

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

WO 02/064792

PCT/US02/00254

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

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

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

10

18. A composition of claim 17, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.

19. A method for treating a disease or condition associated with decreased expression of functional MDDT, comprising administering to a patient in need of such treatment the composition of claim 17.

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

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

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

25

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

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

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

WO 02/064792

PCT/US02/00254

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

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

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

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

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

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

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

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

WO 02/064792

PCT/US02/00254

29. A method of assessing toxicity of a test compound, the method comprising:
- a) treating a biological sample containing nucleic acids with the test compound,
 - b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 12 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 12 or fragment thereof,
 - c) quantifying the amount of hybridization complex, and
 - d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.
30. A diagnostic test for a condition or disease associated with the expression of MDDT in a biological sample, the method comprising:
- a) combining the biological sample with an antibody of claim 11, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and
 - b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.
31. The antibody of claim 11, wherein the antibody is:
- a) a chimeric antibody,
 - b) a single chain antibody,
 - c) a Fab fragment,
 - d) a F(ab')₂ fragment, or
 - e) a humanized antibody.
32. A composition comprising an antibody of claim 11 and an acceptable excipient.
33. A method of diagnosing a condition or disease associated with the expression of MDDT in a subject, comprising administering to said subject an effective amount of the composition of claim 32.

WO 02/064792

PCT/US02/00254

34. A composition of claim 32, wherein the antibody is labeled.

35. A method of diagnosing a condition or disease associated with the expression of MDDT in a subject, comprising administering to said subject an effective amount of the composition of claim 34.

36. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibodies from said animal, and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.

37. A polyclonal antibody produced by a method of claim 36.

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

39. A method of making a monoclonal antibody with the specificity of the antibody of claim 11, the method comprising:

- a) immunizing an animal with a polypeptide consisting of an amino acid sequence selected from the group consisting of SEQ ID NO:1-26, or an immunogenic fragment thereof, under conditions to elicit an antibody response,
- b) isolating antibody producing cells from the animal,
- c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells,
- d) culturing the hybridoma cells, and
- e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.

WO 02/064792

PCT/US02/00254

40. A monoclonal antibody produced by a method of claim 39.
41. A composition comprising the monoclonal antibody of claim 40 and a suitable carrier.
- 5 42. The antibody of claim 11, wherein the antibody is produced by screening a Fab expression library.
43. The antibody of claim 11, wherein the antibody is produced by screening a recombinant immunoglobulin library.
- 10 44. A method of detecting a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26 in a sample, the method comprising:
- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
 - 15 b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26 in the sample.
45. A method of purifying a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26 from a sample, the method comprising:
- a) incubating the antibody of claim 11 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
 - 20 b) separating the antibody from the sample and obtaining the purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-26.
- 25 46. A microarray wherein at least one element of the microarray is a polynucleotide of claim 13.
- 30 47. A method of generating an expression profile of a sample which contains polynucleotides, the method comprising:
- a) labeling the polynucleotides of the sample,
 - b) contacting the elements of the microarray of claim 46 with the labeled polynucleotides

WO 02/064792

PCT/US02/00254

of the sample under conditions suitable for the formation of a hybridization complex,
and

c) quantifying the expression of the polynucleotides in the sample.

5 48. An array comprising different nucleotide molecules affixed in distinct physical locations
on a solid substrate, wherein at least one of said nucleotide molecules comprises a first oligonucleotide
or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of a target
polynucleotide, and wherein said target polynucleotide is a polynucleotide of claim 12.

10 49. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is
completely complementary to at least 30 contiguous nucleotides of said target polynucleotide.

50. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is
completely complementary to at least 60 contiguous nucleotides of said target polynucleotide.

15 51. An array of claim 48, wherein said first oligonucleotide or polynucleotide sequence is
completely complementary to said target polynucleotide.

52. An array of claim 48, which is a microarray.

20 53. An array of claim 48, further comprising said target polynucleotide hybridized to a
nucleotide molecule comprising said first oligonucleotide or polynucleotide sequence.

54. An array of claim 48, wherein a linker joins at least one of said nucleotide molecules to
25 said solid substrate.

55. An array of claim 48, wherein each distinct physical location on the substrate contains
multiple nucleotide molecules, and the multiple nucleotide molecules at any single distinct physical
location have the same sequence, and each distinct physical location on the substrate contains
30 nucleotide molecules having a sequence which differs from the sequence of nucleotide molecules at
another distinct physical location on the substrate.

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

WO 02/064792

PCT/US02/00254

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 2
58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 3
- 5 59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 4
60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 5
61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 6
- 10 62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 7
63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 8
- 15 64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 9
65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 10
66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 11
- 20 67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 12
68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 13
- 25 69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 14
70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 15
71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 16
- 30 72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 17
73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 18

WO 02/064792

PCT/US02/00254

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 19
75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 20
- 5 76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 21
77. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 22
78. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 23
- 10 79. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 24
80. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 25
- 15 81. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO: 26
82. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 27
83. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 28
- 20 84. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 29
85. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 30
- 25 86. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 31
87. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 32
88. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 33
- 30 89. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 34
90. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 35

WO 02/064792

PCT/US02/00254

91. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 36
92. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 37
- 5 93. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 38
94. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 39
95. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 40
- 10 96. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 41
97. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 42
- 15 98. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 43
99. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO: 44
- 20 45 100. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:
- 46 101. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:
- 25 47 102. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:
- 48 103. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:
- 30 49 104. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:

WO 02/064792

PCT/US02/00254

105. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:
50

106. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:
5 51

107. A polynucleotide of claim 12, comprising the polynucleotide sequence of SEQ ID NO:
52

WO 02/064792

PCT/US02/00254

<110> INCYTE GENOMICS, INC.
BAUGHN, Mariah R.
WARREN, Bridget A.
HONCHELL, Cynthia D.
XU, Yuming
WALLA, Narinder K.
RAMKUMAR, Jayalaxmi
YAO, Monique G.
LU, Yan
YUE, Henry
THANGAVELU, Kavitha
TANG, Y. Tom
DING, Li
BOROWSKY, Mark L.
HAFALIA, April J.A.
LU, Dyung Aina M.
AZIMZAI, Yalda
TRAN, Bao
NGUYEN, Dannel B.
BURFORD, Neil
TSON, Craig H.
CURURAJAN, Rajagopal
GANDHI, Ameena R.
ELLIOTT, Vicki S.
TRAN, Uyen K.

<120> MOLECULES FOR DISEASE DETECTION AND TREATMENT

<130> PF-0874 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/260,168; 60/262,857; 60/262,736

<151> 2001-01-05; 2001-01-19; 2001-01-19

<160> 52

<170> PERL Program

<210> 1

<211> 127

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 6242606CD1

<400> 1

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

WO 02/064792

PCT/US02/00254

95 100 105
Arg Glu Ala Gly Asn Leu Thr Lys Ala Thr Ala Ser Asp Gln Lys
110 115 120
Ser Ser Gly Gly Arg Asp Ser
125

<210> 2

<211> 258

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 997105CD1

<400> 2

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

<210> 3

<211> 175

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1482843CD1

WO 02/064792

PCT/US02/00254

```

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

```

<210> 4

<211> 1146

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3218219CD1

```

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

```

WO 02/064792

PCT/US02/00254

```

Asn Gln Arg Tyr Phe Arg Lys Arg Ala Leu Tyr Leu Ala His Leu
200 205 210
Ala His His Leu Ala Gln Asp Pro Leu Phe Gly Ser Val Cys Phe
215 220 225
Ser Tyr Thr Asn Gly Cys His Leu Lys Pro Ser Leu Leu Leu Arg
230 235 240
Pro Arg Gly Lys Asp Glu Arg Leu Val Thr Val Arg Leu His Pro
245 250 255
Cys Pro Pro Pro Asp Phe Phe Arg Pro Cys Arg Leu Leu Pro Thr
260 265 270
Lys Asn Asn Val Arg Ser Ala Trp Tyr Arg Gly Gln Ser Pro Ala
275 280 285
Gly Asp Gly Ser Pro Glu Pro Pro Thr Pro Arg Tyr Asn Thr Trp
290 295 300
Val Leu Gln Asp Thr Val Leu Glu Ser His Leu Gln Leu Leu Ser
305 310 315
Thr Ile Leu Ser Ser Ala Gln Gly Leu Lys Asp Gly Val Ala Leu
320 325 330
Leu Lys Val Trp Leu Arg Gln Arg Glu Leu Asp Lys Gly Gln Gly
335 340 345
Gly Phe Thr Gly Phe Leu Val Ser Met Leu Val Val Phe Leu Val
350 355 360
Ser Thr Arg Lys Ile His Thr Thr Met Ser Gly Tyr Gln Val Leu
365 370 375
Arg Ser Val Leu Gln Phe Leu Ala Thr Thr Asp Leu Thr Val Asn
380 385 390
Gly Ile Ser Leu Cys Leu Ser Ser Asp Pro Ser Leu Pro Ala Leu
395 400 405
Ala Asp Phe His Gln Ala Phe Ser Val Val Phe Leu Asp Ser Ser
410 415 420
Gly His Leu Asn Leu Cys Ala Asp Val Thr Ala Ser Thr Tyr His
425 430 435
Gln Val Gln His Glu Ala Arg Leu Ser Met Met Leu Leu Asp Ser
440 445 450
Arg Ala Asp Asp Gly Phe His Leu Leu Leu Met Thr Pro Lys Pro
455 460 465
Met Ile Arg Ala Phe Asp His Val Leu His Leu Arg Pro Leu Ser
470 475 480
Arg Leu Gln Ala Ala Cys His Arg Leu Lys Leu Trp Pro Glu Leu
485 490 495
Gln Asp Asn Gly Gly Asp Tyr Val Ser Ala Ala Leu Gly Pro Leu
500 505 510
Thr Thr Leu Leu Glu Gln Gly Leu Gly Ala Arg Leu Asn Leu Leu
515 520 525
Ala His Ser Arg Pro Pro Val Pro Glu Trp Asp Ile Ser Gln Asp
530 535 540
Pro Pro Lys His Lys Asp Ser Gly Thr Leu Thr Leu Gly Leu Leu
545 550 555
Leu Arg Pro Glu Gly Leu Thr Ser Val Leu Glu Leu Gly Pro Glu
560 565 570
Ala Asp Gln Pro Glu Ala Ala Lys Phe Arg Gln Phe Trp Gly Ser
575 580 585
Arg Ser Glu Leu Arg Arg Phe Gln Asp Gly Ala Ile Arg Glu Ala
590 595 600
Val Val Trp Glu Ala Ala Ser Met Ser Gln Lys Arg Leu Ile Pro
605 610 615
His Gln Val Val Thr His Leu Leu Ala Leu His Ala Asp Ile Pro
620 625 630
Glu Thr Cys Val His Tyr Val Gly Gly Pro Leu Asp Ala Leu Ile
635 640 645
Gln Gly Leu Lys Glu Thr Ser Ser Thr Gly Glu Glu Ala Leu Val
650 655 660

```

WO 02/064792

PCT/US02/00254

Ala Ala Val Arg Cys Tyr Asp Asp Leu Ser Arg Leu Leu Trp Gly
 665 670 675
 Leu Glu Gly Leu Pro Leu Thr Val Ser Ala Val Gln Gly Ala His
 680 685 690
 Pro Val Leu Arg Tyr Thr Glu Val Phe Pro Pro Thr Pro Val Arg
 695 700 705
 Pro Ala Phe Ser Phe Tyr Glu Thr Leu Arg Glu Arg Ser Ser Leu
 710 715 720
 Leu Pro Arg Leu Asp Lys Pro Cys Pro Ala Tyr Val Glu Pro Met
 725 730 735
 Thr Val Val Cys His Leu Glu Gly Ser Gly Gln Trp Pro Gln Asp
 740 745 750
 Ala Glu Ala Val Gln Arg Val Arg Ala Ala Phe Gln Leu Arg Leu
 755 760 765
 Ala Glu Leu Leu Thr Gln Gln His Gly Leu Gln Cys Arg Ala Thr
 770 775 780
 Ala Thr His Thr Asp Val Leu Lys Asp Gly Phe Val Phe Arg Ile
 785 790 795
 Arg Val Ala Tyr Gln Arg Glu Pro Gln Ile Leu Lys Glu Val Gln
 800 805 810
 Ser Pro Glu Gly Met Ile Ser Leu Arg Asp Thr Ala Ala Ser Leu
 815 820 825
 Arg Leu Glu Arg Asp Thr Arg Gln Leu Pro Leu Leu Thr Ser Ala
 830 835 840
 Leu His Gly Leu Gln Gln Gln His Pro Ala Phe Ser Gly Val Ala
 845 850 855
 Arg Leu Ala Lys Arg Trp Val Arg Ala Gln Leu Leu Gly Glu Gly
 860 865 870
 Phe Ala Asp Glu Ser Leu Asp Leu Val Ala Ala Ala Leu Phe Leu
 875 880 885
 His Pro Glu Pro Phe Thr Pro Pro Ser Ser Pro Gln Val Gly Phe
 890 895 900
 Leu Arg Phe Leu Phe Leu Val Ser Thr Phe Asp Trp Lys Asn Asn
 905 910 915
 Pro Leu Phe Val Asn Leu Asn Asn Glu Leu Thr Val Glu Glu Gln
 920 925 930
 Val Glu Ile Arg Ser Gly Phe Leu Ala Ala Arg Ala Gln Leu Pro
 935 940 945
 Val Met Val Ile Val Thr Pro Gln Asp Arg Lys Asn Ser Val Trp
 950 955 960
 Thr Gln Asp Gly Pro Ser Ala Gln Ile Leu Gln Gln Leu Val Val
 965 970 975
 Leu Ala Ala Glu Ala Leu Pro Met Leu Glu Lys Gln Leu Met Asp
 980 985 990
 Pro Arg Gly Pro Gly Asp Ile Arg Thr Val Phe Arg Pro Pro Leu
 995 1000 1005
 Asp Ile Tyr Asp Val Leu Ile Arg Leu Ser Pro Arg His Ile Pro
 1010 1015 1020
 Arg His Arg Gln Ala Val Asp Ser Pro Ala Ala Ser Phe Cys Arg
 1025 1030 1035
 Gly Leu Leu Ser Gln Pro Gly Pro Ser Ser Leu Met Pro Val Leu
 1040 1045 1050
 Gly Tyr Asp Pro Pro Gln Leu Tyr Leu Thr Gln Leu Arg Glu Ala
 1055 1060 1065
 Phe Gly Asp Leu Ala Leu Phe Phe Tyr Asp Gln His Gly Gly Glu
 1070 1075 1080
 Val Ile Gly Val Leu Trp Lys Pro Thr Ser Phe Gln Pro Gln Pro
 1085 1090 1095
 Phe Lys Ala Ser Ser Thr Lys Gly Arg Met Val Met Ser Arg Gly
 1100 1105 1110
 Gly Glu Leu Val Met Val Pro Asn Val Glu Ala Ile Leu Glu Asp
 1115 1120 1125

WO 02/064792

PCT/US02/00254

Phe Ala Val Leu Gly Glu Gly Leu Val Gln Thr Val Glu Ala Arg
 1130 1135 1140
 Ser Glu Arg Trp Thr Val
 1145

<210> 5
 <211> 351
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7371678CD1

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

WO 02/064792

PCT/US02/00254

350

<210> 6
<211> 286
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7473883CD1

<400> 6
Met Ser Leu Pro Ile Gly Ile Tyr Arg Arg Ala Val Ser Tyr Asp
1 5 10 15
Asp Thr Leu Glu Asp Pro Ala Pro Met Thr Pro Pro Pro Ser Asp
20 25 30
Met Gly Ser Val Pro Trp Lys Pro Val Ile Pro Glu Arg Lys Tyr
35 40 45
Gln His Leu Ala Lys Val Glu Glu Gly Glu Ala Ser Leu Pro Ser
50 55 60
Pro Ala Met Thr Leu Ser Ser Ala Ile Asp Ser Val Asp Lys Val
65 70 75
Pro Val Val Lys Ala Lys Ala Thr His Val Ile Met Asn Ser Leu
80 85 90
Ile Thr Lys Gln Thr Gln Glu Ser Ile Gln His Phe Glu Arg Gln
95 100 105
Ala Gly Leu Arg Asp Ala Gly Tyr Thr Pro His Lys Gly Leu Thr
110 115 120
Thr Glu Glu Thr Lys Tyr Leu Arg Val Ala Glu Ala Leu His Lys
125 130 135
Leu Lys Leu Gln Ser Gly Glu Val Thr Lys Glu Glu Arg Gln Pro
140 145 150
Ala Ser Ala Gln Ser Thr Pro Ser Thr Thr Pro His Ser Ser Pro
155 160 165
Lys Gln Arg Pro Arg Gly Trp Phe Thr Ser Gly Ser Ser Thr Ala
170 175 180
Leu Pro Gly Pro Asn Pro Ser Thr Met Asp Ser Gly Ser Gly Asp
185 190 195
Lys Asp Arg Asn Leu Ser Asp Lys Trp Ser Leu Phe Gly Pro Arg
200 205 210
Ser Leu Gln Lys Tyr Asp Ser Gly Ser Phe Ala Thr Gln Ala Tyr
215 220 225
Arg Gly Ala Gln Lys Pro Ser Pro Leu Glu Leu Ile Arg Ala Gln
230 235 240
Ala Asn Arg Met Ala Glu Asp Pro Ala Ala Leu Lys Pro Pro Lys
245 250 255
Met Asp Ile Pro Val Met Glu Gly Lys Lys Gln Pro Pro Arg Ala
260 265 270
His Asn Leu Lys Pro Arg Asp Leu Asn Val Leu Thr Pro Thr Gly
275 280 285
Phe

<210> 7
<211> 205
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7478662CD1

WO 02/064792

PCT/US02/00254

```

<400> 7
Met Pro Asp Thr Ala Phe Pro Pro Ser Ser Arg Ala Gly Gln Thr
1      5      10      15
Gly Pro Val Val Ser Gly Ala Gln Val Ser Ser Trp Arg Glu Arg
20     25     30
Gln Pro Cys Ser Gly Ser Arg Gly Pro Ser His Ile Leu Gly Thr
35     40     45
Asp Ala Asn Ala Gly Thr Ala Gly Lys Leu Gly Leu Val Ser Phe
50     55     60
Pro Gly Phe Arg Lys Lys Pro Thr Arg Ser Val Leu Gln Asn Glu
65     70     75
Thr Phe Val Ser Arg Ser Thr Ala Asp Cys Lys Trp Leu Pro Phe
80     85     90
Val Ala Val Ser Ile Leu Ser Ile Leu Arg Leu Asp Trp Phe Leu
95     100    105
Leu Leu Ala Leu Pro Cys Pro Thr Asp His Lys Gly Glu Gln Gln
110    115    120
Glu Ala Pro Ser Lys His Pro Gln Met Ala Leu Asp Ile Ser His
125    130    135
Ile Leu Arg Asn Met Ser Cys Ser Gly Arg Ala Lys Ala Ser Ser
140    145    150
Lys Ala Cys Gly Ala Gly Gly Ser Gln Ala Arg Trp Gly Asn Pro
155    160    165
Glu Pro Trp Gly Leu Pro Met Gly Ile Gly Arg Ser Gln Gly Arg
170    175    180
Val Arg Gly Ser Thr Gly Gly Val Arg Glu Ser Pro Arg Ala Leu
185    190    195
Leu Ala Gln Gly Ala Val Asn Thr Met Asp
200    205

```

```

<210> 8
<211> 233
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7650474CD1

```

```

<400> 8
Met Ser Lys Leu Gly Lys Phe Phe Lys Gly Gly Gly Ser Ser Lys
1      5      10      15
Ser Arg Ala Ala Pro Ser Pro Gln Glu Ala Leu Val Arg Leu Arg
20     25     30
Glu Thr Glu Glu Met Leu Gly Lys Lys Gln Glu Tyr Leu Glu Asn
35     40     45
Arg Ile Gln Arg Glu Ile Ala Leu Ala Lys Lys His Gly Thr Gln
50     55     60
Asn Lys Arg Ala Ala Leu Gln Ala Leu Lys Arg Lys Lys Arg Phe
65     70     75
Glu Lys Gln Leu Thr Gln Ile Asp Gly Thr Leu Ser Thr Ile Glu
80     85     90
Phe Gln Arg Glu Ala Leu Glu Asn Ser His Thr Asn Thr Glu Val
95     100    105
Leu Arg Asn Met Gly Phe Ala Ala Lys Ala Met Lys Ser Val His
110    115    120
Glu Asn Met Asp Leu Asn Lys Ile Asp Asp Leu Met Gln Glu Ile
125    130    135
Thr Glu Gln Gln Asp Ile Ala Gln Glu Ile Ser Glu Ala Phe Ser
140    145    150
Gln Arg Val Gly Phe Gly Asp Asp Phe Asp Glu Asp Glu Leu Met
155    160    165

```

WO 02/064792

PCT/US02/00254

Ala Glu Leu Glu Glu Leu Glu Gln Glu Glu Leu Asn Lys Lys Met
 170 175 180
 Thr Asn Ile Arg Leu Pro Asn Val Pro Ser Ser Ser Leu Pro Ala
 185 190 195
 Gln Pro Asn Arg Lys Pro Gly Met Ser Ser Thr Ala Arg Arg Ser
 200 205 210
 Arg Ala Ala Ser Ser Gln Arg Ala Glu Glu Glu Asp Asp Asp Ile
 215 220 225
 Lys Gln Leu Ala Ala Trp Ala Thr
 230

<210> 9
 <211> 205
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 8092378CD1

<400> 9
 Met Ile Val Val Leu His Val His Phe His Met Ala Met Leu Pro
 1 10 15
 Phe Pro Ile Phe Leu Val Leu Leu Arg Gly Leu Val Leu Trp
 20 25 30
 Thr Pro Ala Ser Ser Gly Thr Ile Met Pro Glu Glu Arg Lys Thr
 35 40 45
 Glu Ile Glu Arg Glu Thr Glu Thr Glu Ser Glu Thr Val Ile Gly
 50 55 60
 Thr Glu Lys Glu Asn Ala Pro Glu Arg Glu Arg Gly Ser Val Ile
 65 70 75
 Thr Val Leu His Gln Val Phe Ser Thr Ala Met Lys Asn Asp Thr
 80 85 90
 Asp Thr Gly Asn Met Gln Lys Glu Val Met Ser Val Thr Glu Gln
 95 100 105
 Val Glu Lys Lys Lys Asn Asp Ile Glu Lys Asp Asp Thr Gly Arg
 110 115 120
 Lys Arg Lys Pro Asp Ile Ser Leu Leu Glu Val Ile Val Asp Val
 125 130 135
 Ala Met Lys Val Lys Lys Glu Ile Val Thr Gly Asp Thr Asn Thr
 140 145 150
 Lys Asn Leu Lys Glu Ala Lys Lys Glu Lys Lys Arg Ala Val Ser
 155 160 165
 Leu Pro Leu Asn Arg Arg Ala Pro Lys Leu His Leu Gln Asn Arg
 170 175 180
 His Gly Phe Gly Leu Leu Cys Ile Leu Val Pro Glu Val Asp Thr
 185 190 195
 Ile Asn Leu Val Ile Phe Leu Asp Asn Val
 200 205

<210> 10
 <211> 304
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1263178CD1

<400> 10
 Met Gly Gln Cys Val Thr Lys Cys Lys Asn Pro Ser Ser Thr Leu
 1 5 10 15

WO 02/064792

PCT/US02/00254

```

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

```

<210> 11

<211> 418

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3535417CD1

<400> 11

```

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

```

WO 02/064792

PCT/US02/00254

```

      80      85      90
Thr Gly Leu Ile Gln Gly Gln Arg Ala Cys Leu Ala Glu Ala Ile
      95      100      105
Thr Leu Val Glu Ser Thr His Ser Arg Lys Lys Glu Leu Ala Gln
      110      115      120
Val Leu Leu Gln Lys Val Leu Leu Tyr His Arg Glu Gln Glu Gln
      125      130      135
Ser Asn Lys Gly Lys Pro Leu Ala Phe Arg Val Gly Leu Ser Gly
      140      145      150
Pro Pro Gly Ala Gly Lys Ser Thr Phe Ile Glu Tyr Phe Gly Lys
      155      160      165
Met Leu Thr Glu Arg Gly His Lys Leu Ser Val Leu Ala Val Asp
      170      175      180
Pro Ser Ser Cys Thr Ser Gly Gly Ser Leu Leu Gly Asp Lys Thr
      185      190      195
Arg Met Thr Glu Leu Ser Arg Asp Met Asn Ala Tyr Ile Arg Pro
      200      205      210
Ser Pro Thr Arg Gly Thr Leu Gly Gly Val Thr Arg Thr Thr Asn
      215      220      225
Glu Ala Ile Leu Leu Cys Glu Gly Ala Gly Tyr Asp Ile Ile Leu
      230      235      240
Ile Glu Thr Val Gly Val Gly Gln Ser Glu Phe Ala Val Ala Asp
      245      250      255
Met Val Asp Met Phe Val Leu Leu Leu Pro Pro Ala Gly Gly Asp
      260      265      270
Glu Leu Gln Gly Ile Lys Arg Gly Ile Ile Glu Met Ala Asp Leu
      275      280      285
Val Ala Val Thr Lys Ser Asp Gly Asp Leu Ile Val Pro Ala Arg
      290      295      300
Arg Ile Gln Ala Glu Tyr Val Ser Ala Leu Lys Leu Leu Arg Lys
      305      310      315
Arg Ser Gln Val Trp Lys Pro Lys Val Ile Arg Ile Ser Ala Arg
      320      325      330
Ser Gly Glu Gly Ile Ser Glu Met Trp Asp Lys Met Lys Asp Phe
      335      340      345
Gln Asp Leu Met Leu Ala Ser Gly Glu Leu Thr Ala Lys Arg Arg
      350      355      360
Lys Gln Gln Lys Val Trp Met Trp Asn Leu Ile Gln Glu Ser Val
      365      370      375
Leu Glu His Phe Arg Thr His Pro Thr Val Arg Glu Gln Ile Pro
      380      385      390
Leu Leu Glu Gln Lys Val Leu Ile Gly Ala Leu Ser Pro Gly Leu
      395      400      405
Ala Ala Asp Phe Leu Leu Lys Ala Phe Lys Ser Arg Asp
      410      415

```

<210> 12
 <211> 1117
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7473555CD1

```

<400> 12
Met Phe Gln Thr Glu Leu Met Ala Leu Pro His Thr Tyr Ser Phe
      1      5      10      15
Leu Tyr Ser Val Ser Asp Ile Thr Cys Pro Pro Arg Gln Ala Pro
      20      25      30
Gly Ser His Leu Val Thr Asn Pro Val Thr Val Val Ala Pro Val
      35      40      45

```

WO 02/064792

PCT/US02/00254

```

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

```

WO 02/064792

PCT/US02/00254

Gly Tyr Ser Thr Ser Gln Leu Pro Ser Val Val Ser Gly Arg Leu
 515 520 525
 Cys His Phe Phe Thr Val Leu Leu His Leu Leu Trp Gly Ser Leu
 530 535 540
 Thr Ser Glu Asn Cys Pro Cys Tyr Ser Arg Cys Gly His Thr Lys
 545 550 555
 Met Ser Ala Ser Val Leu His Ala Thr His Thr Val Glu Ala Val
 560 565 570
 Ile His Arg Pro Ala Val Pro Pro Ala Glu Ile Gln Thr Arg Phe
 575 580 585
 Cys Ser Leu Gly Asn Val Asn Thr Arg Ala Ala Pro His Ser Thr
 590 595 600
 Ile Leu Thr Ser Pro Pro Thr Arg Ala Lys Cys Asn Arg Arg Leu
 605 610 615
 Lys Arg Cys Ser Cys Arg Pro Cys Trp Met Glu Gln Asp Leu Phe
 620 625 630
 Ser Ser Ala Val Pro Thr Ile Pro Phe Leu Ser His Lys Ala Tyr
 635 640 645
 Ser Glu Asp Thr Leu Leu Ser Trp Ala Val Gly Val Cys Ile Tyr
 650 655 660
 Asn Leu Arg Leu Leu Pro Pro Ser Leu Ile Ser Ser Trp Asn Ser
 665 670 675
 Arg Gly Leu His Thr Leu Asn Gly Ser Ala Gln Gln Gln Glu Ser
 680 685 690
 Lys Ala Ala Asp Arg Val Leu Ile Asn Glu Leu Leu Gly Leu Arg
 695 700 705
 Val Asp Arg Lys Glu Asp Asn Leu Met Gln Thr Phe Phe Leu Glu
 710 715 720
 Cys Asp Trp Ser Cys Ser Ala Cys Ser Gly Pro Leu Lys Thr Asp
 725 730 735
 Cys Leu Gln Cys Met Asp Gly Tyr Val Leu Gln Asp Gly Ala Cys
 740 745 750
 Val Glu Gln Cys Leu Ser Ser Phe Tyr Gln Asp Ser Gly Leu Cys
 755 760 765
 Lys Asn Cys Asp Ser Tyr Cys Leu Gln Cys Gln Gly Pro His Glu
 770 775 780
 Cys Thr Arg Cys Lys Gly Pro Phe Leu Leu Leu Glu Ala Gln Cys
 785 790 795
 Val Gln Glu Cys Gly Lys Gly Tyr Phe Ala Asp His Ala Lys His
 800 805 810
 Lys Cys Thr Ala Cys Pro Gln Gly Cys Leu Gln Cys Ser His Arg
 815 820 825
 Asp Arg Cys His Leu Cys Asp His Gly Phe Phe Leu Lys Ser Gly
 830 835 840
 Leu Cys Val Tyr Asn Cys Val Pro Gly Phe Ser Val His Thr Ser
 845 850 855
 Asn Glu Thr Cys Ser Gly Lys Ile His Thr Pro Ser Leu His Val
 860 865 870
 Asn Gly Ser Leu Ile Leu Pro Ile Gly Ser Ile Lys Pro Leu Asp
 875 880 885
 Phe Ser Leu Leu Asn Val Gln Asp Gln Glu Gly Arg Val Lys Asp
 890 895 900
 Leu Leu Phe His Val Val Ser Thr Pro Thr Asn Gly Gln Leu Val
 905 910 915
 Leu Ser Arg Asn Gly Lys Glu Val Gln Leu Asp Lys Ala Gly Arg
 920 925 930
 Phe Ser Trp Lys Asp Val Asn Glu Lys Lys Val Arg Phe Val His
 935 940 945
 Ser Lys Glu Lys Leu Arg Lys Gly Tyr Leu Phe Leu Lys Ile Ser
 950 955 960
 Asp Gln Gln Phe Phe Ser Glu Pro Gln Leu Ile Asn Ile Gln Ala
 965 970 975

WO 02/064792

PCT/US02/00254

```

Phe Ser Thr Gln Ala Pro Tyr Val Leu Arg Asn Glu Val Leu His
980          985          990
Ile Ser Arg Gly Glu Arg Ala Thr Ile Thr Thr Gln Met Leu Asp
995          1000          1005
Ile Arg Asp Asp Asp Asn Pro Gln Asp Val Val Ile Glu Ile Ile
1010          1015          1020
Asp Pro Pro Leu His Gly Gln Leu Leu Gln Thr Leu Gln Ser Pro
1025          1030          1035
Ala Thr Pro Ile Tyr Gln Phe Gln Leu Asp Glu Leu Ser Arg Gly
1040          1045          1050
Leu Leu His Tyr Ala His Asp Gly Ser Asp Ser Thr Ser Asp Val
1055          1060          1065
Ala Val Leu Gln Ala Asn Asp Gly His Ser Phe His Asn Ile Leu
1070          1075          1080
Phe Gln Val Lys Thr Val Pro Gln Gly Gln Ser Tyr Asp Gly Met
1085          1090          1095
Ser Ser Leu Asp Pro Ala Pro Ser Ser Arg Val Leu Arg Ala Phe
1100          1105          1110
Ile Phe Leu Met Leu Val Pro
1115

```

<210> 13

<211> 920

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7474985CD1

<400> 13

```

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

```

WO 02/064792

PCT/US02/00254

```

230      235      240
Tyr Val Phe Tyr Leu Asn Gly Arg Glu Ile Asn Gln Leu Lys Glu
245      250      255
Arg Ser Phe Ala Gln Leu Ile Ser Lys Asp Trp Pro Ser Thr Glu
260      265      270
Gly Pro Ile Glu Glu Ile Met Tyr Gln Pro Ser Ser Leu Leu Phe
275      280      285
Ile Ile Asp Ser Phe Asp Glu Leu Asn Phe Ala Phe Glu Glu Pro
290      295      300
Glu Phe Ala Leu Cys Glu Asp Trp Thr Gln Glu His Pro Val Ser
305      310      315
Phe Leu Met Ser Ser Leu Leu Arg Lys Val Met Leu Pro Glu Ala
320      325      330
Ser Leu Leu Val Thr Thr Arg Leu Thr Thr Ser Lys Arg Leu Lys
335      340      345
Gln Leu Leu Lys Asn His His Tyr Val Glu Leu Leu Gly Met Ser
350      355      360
Glu Asp Ala Arg Glu Glu Tyr Ile Tyr Gln Phe Phe Glu Asp Lys
365      370      375
Arg Trp Ala Met Lys Val Phe Ser Ser Leu Lys Ser Asn Glu Met
380      385      390
Leu Phe Ser Met Cys Gln Val Pro Leu Val Cys Trp Ala Ala Cys
395      400      405
Thr Cys Leu Lys Gln Gln Met Glu Lys Gly Gly Asp Val Thr Leu
410      415      420
Thr Cys Gln Thr Thr Thr Ala Leu Phe Thr Cys Tyr Ile Ser Ser
425      430      435
Leu Phe Thr Pro Val Asp Gly Gly Ser Pro Ser Leu Pro Asn Gln
440      445      450
Ala Gln Leu Arg Arg Leu Cys Gln Val Ala Ala Lys Gly Ile Trp
455      460      465
Thr Met Thr Tyr Val Phe Tyr Arg Glu Asn Leu Arg Arg Leu Gly
470      475      480
Leu Thr Gln Ser Asp Val Ser Ser Phe Met Asp Ser Asn Ile Ile
485      490      495
Gln Lys Asp Ala Glu Tyr Glu Asn Cys Tyr Val Phe Thr His Leu
500      505      510
His Val Gln Glu Phe Phe Ala Ala Met Phe Tyr Met Leu Lys Gly
515      520      525
Ser Trp Glu Ala Gly Asn Pro Ser Cys Gln Pro Phe Glu Asp Leu
530      535      540
Lys Ser Leu Leu Gln Ser Thr Ser Tyr Lys Asp Pro His Leu Thr
545      550      555
Gln Met Lys Cys Phe Leu Phe Gly Leu Leu Asn Glu Asp Arg Val
560      565      570
Lys Gln Leu Glu Arg Thr Phe Asn Cys Lys Met Ser Leu Lys Ile
575      580      585
Lys Ser Lys Leu Leu Gln Cys Met Glu Val Leu Gly Asn Ser Asp
590      595      600
Tyr Ser Pro Ser Gln Leu Gly Phe Leu Glu Leu Phe His Cys Leu
605      610      615
Tyr Glu Thr Gln Asp Lys Ala Phe Ile Ser Gln Ala Met Arg Cys
620      625      630
Phe Pro Lys Val Ala Ile Asn Ile Cys Glu Lys Ile His Leu Leu
635      640      645
Val Ser Ser Phe Cys Leu Lys His Cys Arg Cys Leu Arg Thr Ile
650      655      660
Arg Leu Ser Val Thr Val Val Phe Glu Lys Lys Ile Leu Lys Thr
665      670      675
Ser Leu Pro Thr Asn Thr Trp Leu Glu Tyr Cys Gly Leu Thr Ser
680      685      690
Leu Cys Cys Gln Asp Leu Ser Ser Ala Leu Ile Cys Asn Lys Arg

```

WO 02/064792

PCT/US02/00254

```

695          700          705
Leu Ile Lys Met Asn Leu Thr Gln Asn Thr Leu Gly Tyr Glu Gly
710          715          720
Ile Val Lys Leu Tyr Lys Val Leu Lys Ser Pro Lys Cys Lys Leu
725          730          735
Gln Val Leu Gly Leu Glu Ser Cys Gly Leu Thr Glu Ala Gly Cys
740          745          750
Glu Tyr Leu Ser Leu Ala Leu Ile Ser Asn Lys Arg Leu Thr His
755          760          765
Leu Cys Leu Ala Asp Asn Val Leu Gly Asp Gly Gly Val Lys Leu
770          775          780
Met Ser Asp Ala Leu Gln His Ala Gln Cys Thr Leu Lys Ser Leu
785          790          795
Val Leu Met Gly Cys Val Leu Thr Asn Ala Cys Cys Leu Asp Leu
800          805          810
Ala Ser Val Ile Leu Asn Asn Pro Asn Leu Arg Ser Leu Asp Leu
815          820          825
Gly Asn Asn Asp Leu Gln Asp Asp Gly Val Lys Ile Leu Cys Asp
830          835          840
Ala Leu Arg Tyr Pro Asn Cys Asn Ile Gln Arg Leu Gly Leu Leu
845          850          855
Ile Thr Asn Thr Ile Leu Glu Asn Gly Pro Glu Leu Gln Leu Ser
860          865          870
Asp Gln Pro Arg Arg Leu Arg Lys Lys Ser Met Glu Met Val Leu
875          880          885
Phe Gln Gly Ala Cys Trp Leu Pro His Thr Ile Asp Leu Ser Gly
890          895          900
Asp Ser Ser Leu Ala Asp Glu Ser Glu Arg Ile Ser Ser His Thr
905          910          915
Val Leu His Ile Val
920

```

<210> 14

<211> 340

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7475137CD1

<400> 14

```

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

```

WO 02/064792

PCT/US02/00254

Gln Glu Ile Trp Tyr Tyr Val Lys Arg Pro Asn Leu His Leu Ile
 155 160 165
 Ser Val Pro Glu Thr Asp Gly Ala Asn Arg Thr Lys Leu Glu His
 170 175 180
 Thr Leu His Asp Ile Ile Gln Asn Leu Pro Lys Ile Ala Arg Gln
 185 190 195
 Ala Asn Thr Gln Ile Gln Glu Ile Gln Arg Thr Pro Gln Arg Tyr
 200 205 210
 Ser Ser Arg Ile Ala Thr Pro Arg His Val Val Arg Phe Thr Lys
 215 220 225
 Val Lys Met Lys Glu Lys Met Leu Thr Ala Ala Arg Glu Lys Gly
 230 235 240
 Gln Val Thr His Lys Gly Lys Pro Ile Arg Leu Thr Ala Asp Leu
 245 250 255
 Leu Ala Glu Thr Leu Gln Ala Arg Arg Gln Trp Gly Pro Ile Phe
 260 265 270
 Asn Ile Leu Lys Glu Lys Asn Phe Gln Ser Arg Ile Pro Tyr Pro
 275 280 285
 Ala Lys Leu Ser Phe Ile Ser Glu Gly Glu Ile Lys Ser Phe Thr
 290 295 300
 Glu Lys Gln Met Leu Arg Asp Phe Val Thr Thr Arg Pro Ala Leu
 305 310 315
 Gln Glu Leu Leu Lys Glu Ala Leu Asn Met Glu Arg Asn Asn Gln
 320 325 330
 Tyr Gln Pro Leu Gln Lys His Ala Lys Leu
 335 340

<210> 15

<211> 577

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5036986CD1

<400> 15

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

WO 02/064792

PCT/US02/00254

```

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

```

<210> 16

<211> 119

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1375644CD1

WO 02/064792

PCT/US02/00254

```

<400> 16
Met Val Arg Ile Ser Lys Pro Lys Thr Phe Gln Ala Tyr Leu Asp
1      5      10      15
Asp Cys His Arg Arg Tyr Ser Cys Ala His Cys Arg Ala His Leu
20     25     30
Ala Asn His Asp Asp Leu Ile Ser Lys Ser Phe Gln Gly Ser Gln
35     40     45
Gly Arg Ala Tyr Leu Phe Asn Ser Val Val Asn Val Gly Cys Gly
50     55     60
Pro Ala Glu Glu Arg Val Leu Leu Thr Gly Leu His Ala Val Ala
65     70     75
Asp Ile His Cys Glu Asn Cys Lys Thr Thr Leu Gly Trp Lys Tyr
80     85     90
Glu Gln Ala Phe Glu Ser Ser Gln Lys Tyr Lys Glu Gly Lys Tyr
95     100    105
Ile Ile Glu Leu Asn His Met Ile Lys Asp Asn Gly Trp Asp
110    115

```

```

<210> 17
<211> 389
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1356261CD1

```

```

<400> 17
Met Ser Val Arg Thr Leu Pro Leu Leu Phe Leu Asn Leu Gly Gly
1      5      10      15
Glu Met Leu Tyr Ile Leu Asp Gln Arg Leu Arg Ala Gln Asn Ile
20     25     30
Pro Gly Asp Lys Ala Arg Lys Asp Glu Trp Thr Glu Val Asp Arg
35     40     45
Lys Arg Val Leu Asn Asp Ile Ile Ser Thr Met Phe Asn Arg Lys
50     55     60
Phe Met Glu Glu Leu Phe Lys Pro Gln Glu Leu Tyr Ser Lys Lys
65     70     75
Ala Leu Arg Thr Val Tyr Glu Arg Leu Ala His Ala Ser Ile Met
80     85     90
Lys Leu Asn Gln Ala Ser Met Asp Lys Leu Tyr Asp Leu Met Thr
95     100    105
Met Ala Phe Lys Tyr Gln Val Leu Leu Cys Pro Arg Pro Lys Asp
110    115    120
Val Leu Leu Val Thr Phe Asn His Leu Asp Thr Ile Lys Gly Phe
125    130    135
Ile Arg Asp Ser Pro Ala Ile Leu Gln Gln Val Asp Glu Thr Leu
140    145    150
Arg Gln Leu Thr Glu Ile Tyr Gly Gly Leu Ser Ala Gly Glu Phe
155    160    165
Gln Leu Ile Arg Gln Thr Leu Leu Ile Phe Phe Gln Asp Leu His
170    175    180
Ile Arg Val Ser Met Phe Leu Lys Asp Lys Val Gln Asn Asn Asn
185    190    195
Gly Arg Phe Val Leu Pro Val Ser Gly Pro Val Pro Trp Gly Thr
200    205    210
Glu Val Pro Gly Leu Ile Arg Met Phe Asn Asn Lys Gly Glu Glu
215    220    225
Val Lys Arg Ile Glu Phe Lys His Gly Gly Asn Tyr Val Pro Ala
230    235    240
Pro Lys Glu Gly Ser Phe Glu Leu Tyr Gly Asp Arg Val Leu Lys
245    250    255

```

WO 02/064792

PCT/US02/00254

```

Leu Gly Thr Asn Met Tyr Ser Val Asn Gln Pro Val Glu Thr His
260 265 270
Val Ser Gly Ser Ser Lys Asn Leu Ala Ser Trp Thr Gln Glu Ser
275 280 285
Ile Ala Pro Asn Pro Leu Ala Lys Glu Glu Leu Asn Phe Leu Ala
290 295 300
Arg Leu Met Gly Gly Met Glu Ile Lys Lys Pro Ser Gly Pro Glu
305 310 315
Pro Arg Phe Arg Leu Asn Leu Phe Thr Thr Asp Glu Glu Glu Glu
320 325 330
Gln Ala Ala Leu Thr Arg Pro Glu Glu Leu Ser Tyr Glu Val Ile
335 340 345
Asn Ile Gln Ala Thr Gln Asp Gln Gln Arg Ser Glu Glu Leu Ala
350 355 360
Arg Ile Met Gly Glu Phe Glu Ile Thr Glu Gln Pro Arg Leu Ser
365 370 375
Thr Ser Lys Gly Asp Asp Leu Leu Ala Met Met Asp Glu Leu
380 385

```

```

<210> 18
<211> 266
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1419379CD1

```

```

<400> 18
Met Val Thr Glu Gln Glu Val Asp Ala Ile Gly Gln Thr Leu Val
1 5 10 15
Asp Pro Lys Gln Pro Leu Gln Ala Arg Phe Arg Ala Leu Phe Thr
20 25 30
Leu Arg Gly Leu Gly Gly Pro Gly Ala Ile Ala Trp Ile Ser Gln
35 40 45
Ala Phe Asp Asp Asp Ser Ala Leu Leu Lys His Glu Leu Ala Tyr
50 55 60
Cys Leu Gly Gln Met Gln Asp Ala Arg Ala Ile Pro Met Leu Val
65 70 75
Asp Val Leu Gln Asp Thr Arg Gln Glu Pro Met Val Arg His Glu
80 85 90
Ala Gly Glu Ala Leu Gly Ala Ile Gly Asp Pro Glu Val Leu Glu
95 100 105
Ile Leu Lys Gln Tyr Ser Ser Asp Pro Val Ile Glu Val Ala Glu
110 115 120
Thr Cys Gln Leu Ala Val Arg Arg Leu Glu Trp Leu Gln Gln His
125 130 135
Gly Gly Glu Pro Ala Ala Gly Pro Tyr Leu Ser Val Asp Pro Ala
140 145 150
Pro Pro Ala Glu Glu Arg Asp Val Gly Arg Leu Arg Glu Ala Leu
155 160 165
Leu Asp Glu Ser Arg Pro Leu Phe Glu Arg Tyr Arg Ala Met Phe
170 175 180
Ala Leu Arg Asn Ala Gly Gly Glu Glu Ala Ala Leu Ala Leu Ala
185 190 195
Glu Gly Leu His Cys Gly Ser Ala Leu Phe Arg His Glu Val Gly
200 205 210
Tyr Val Leu Gly Gln Leu Gln His Glu Ala Glu Leu Gly Leu Trp
215 220 225
Glu Glu Ala Asp Val Met Val Ala Ser Ala Ser Leu Gly Met Gly
230 235 240
His Arg Val Trp Gly Pro Leu Thr Pro Pro Pro Arg Gly Leu Gly

```

WO 02/064792

PCT/US02/00254

245 250 255
 Arg Asp Arg Val Val Val Pro Glu Trp Val Arg
 260 265

<210> 19
 <211> 661
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 3509272CD1

<400> 19
 Met His Pro Ile Pro Ser Ser Phe Met Ile Lys Ala Val Ser Ser
 1 5 10 15
 Phe Leu Thr Ala Glu Glu Ala Ser Val Gly Asn Pro Glu Gly Ala
 20 25 30
 Phe Met Lys Val Leu Gln Ala Arg Lys Asn Tyr Thr Ser Thr Glu
 35 40 45
 Leu Ile Val Glu Pro Glu Glu Pro Ser Asp Ser Ser Gly Ile Asn
 50 55 60
 Leu Ser Gly Phe Gly Ser Glu Gln Leu Asp Thr Asn Asp Glu Ser
 65 70 75
 Asp Phe Ile Ser Thr Leu Ser Tyr Ile Leu Pro Tyr Phe Ser Ala
 80 85 90
 Val Asn Leu Asp Val Lys Ser Leu Leu Leu Pro Leu Ile Lys Leu
 95 100 105
 Pro Thr Thr Gly Asn Ser Leu Ala Lys Ile Gln Thr Val Gly Gln
 110 115 120
 Asn Arg Gln Arg Val Lys Arg Val Leu Met Gly Pro Arg Ser Ile
 125 130 135
 Gln Lys Arg His Phe Lys Glu Val Gly Arg Gln Ser Ile Arg Arg
 140 145 150
 Glu Gln Gly Ala Gln Ala Ser Val Glu Asn Ala Ala Glu Glu Lys
 155 160 165
 Arg Leu Gly Ser Pro Ala Pro Arg Glu Val Glu Gln Pro His Thr
 170 175 180
 Gln Gln Gly Pro Glu Lys Leu Ala Gly Asn Ala Val Tyr Thr Lys
 185 190 195
 Pro Ser Phe Thr Gln Glu His Lys Ala Ala Val Ser Val Leu Lys
 200 205 210
 Pro Phe Ser Lys Gly Ala Pro Ser Thr Ser Ser Pro Ala Lys Ala
 215 220 225
 Leu Pro Gln Val Arg Asp Arg Trp Lys Asp Leu Thr His Ala Ile
 230 235 240
 Ser Ile Leu Glu Ser Ala Lys Ala Arg Val Thr Asn Thr Lys Thr
 245 250 255
 Ser Lys Pro Ile Val His Ala Arg Lys Lys Tyr Arg Phe His Lys
 260 265 270
 Thr Arg Ser His Val Thr His Arg Thr Pro Lys Val Lys Lys Ser
 275 280 285
 Pro Lys Val Arg Lys Lys Ser Tyr Leu Ser Arg Leu Met Leu Ala
 290 295 300
 Asn Arg Leu Pro Phe Ser Ala Ala Lys Ser Leu Ile Asn Ser Pro
 305 310 315
 Ser Gln Gly Ala Phe Ser Ser Leu Gly Asp Leu Ser Pro Gln Glu
 320 325 330
 Asn Pro Phe Leu Glu Val Ser Ala Pro Ser Glu His Phe Ile Glu
 335 340 345
 Lys Asn Asn Thr Lys His Thr Thr Ala Arg Asn Ala Phe Glu Glu
 350 355 360

WO 02/064792

PCT/US02/00254

Asn Asp Phe Met Glu Asn Thr Asn Met Pro Glu Gly Thr Ile Ser
365 370 375
Glu Asn Thr Asn Tyr Asn His Pro Pro Glu Ala Asp Ser Ala Gly
380 385 390
Thr Ala Phe Asn Leu Gly Pro Thr Val Lys Gln Thr Glu Thr Lys
395 400 405
Trp Glu Tyr Asn Asn Val Gly Thr Asp Leu Ser Pro Glu Pro Lys
410 415 420
Ser Phe Asn Tyr Pro Leu Leu Ser Ser Pro Gly Asp Gln Phe Glu
425 430 435
Ile Gln Leu Thr Gln Gln Leu Gln Ser Leu Ile Pro Asn Asn Asn
440 445 450
Val Arg Arg Leu Ile Ala His Val Ile Arg Thr Leu Lys Met Asp
455 460 465
Cys Ser Gly Ala His Val Gln Val Thr Cys Ala Lys Leu Ile Ser
470 475 480
Arg Thr Gly His Leu Met Lys Leu Leu Ser Gly Gln Gln Glu Val
485 490 495
Lys Ala Ser Lys Ile Glu Trp Asp Thr Asp Gln Trp Lys Ile Glu
500 505 510
Asn Tyr Ile Asn Glu Ser Thr Glu Ala Gln Ser Glu Gln Lys Glu
515 520 525
Lys Ser Leu Glu Leu Lys Lys Glu Val Pro Gly Tyr Gly Tyr Thr
530 535 540
Asp Lys Leu Ile Leu Ala Leu Ile Val Thr Gly Ile Leu Thr Ile
545 550 555
Leu Ile Ile Leu Phe Cys Leu Ile Val Ile Cys Cys His Arg Arg
560 565 570
Ser Leu Gln Glu Asp Glu Glu Gly Phe Ser Arg Gly Ile Phe Arg
575 580 585
Phe Leu Pro Arg Arg Gly Cys Ser Ser Arg Arg Glu Ser Gln Asp
590 595 600
Gly Leu Ser Ser Phe Gly Gln Pro Leu Trp Phe Lys Asp Met Tyr
605 610 615
Lys Pro Leu Ser Ala Thr Arg Ile Asn Asn His Ala Trp Lys Leu
620 625 630
His Lys Lys Ser Ser Asn Glu Asp Lys Ile Leu Asn Arg Asp Pro
635 640 645
Gly Asp Ser Glu Ala Pro Thr Glu Glu Glu Glu Ser Glu Ala Leu
650 655 660
Pro

<210> 20
<211> 205
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 708425CD1

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

WO 02/064792

PCT/US02/00254

```

      65      70      75
Arg Glu Ala Trp Lys Arg Ser Asn Phe Ser Cys Ala Ser Thr Ser
      80      85      90
Gly Val Ser Gly Ala Ser Ala Ser Ala Ser Ser Ser Ala Ser
      95     100     105
Met Val Ser Ser Ala Ser Ser Ser Gly Ser Ser Val Gly Asn Ser
     110     115     120
Ala Ser Asn Ser Ser Ala Asn Met Ser Arg Ala His Ser Asp Ser
     125     130     135
Asn Leu Ser Ala Ser Ala Ala Glu Arg Ile Arg Asp Ser Lys Lys
     140     145     150
Arg Ser Lys Gln Arg Lys Leu Gln Gln Lys Ala Phe Arg Lys Arg
     155     160     165
Gln Leu Lys Glu Gln Arg Gln Ala Arg Lys Glu Arg Leu Ser Gly
     170     175     180
Leu Phe Leu Asn Glu Glu Val Leu Ser Leu Lys Val Thr Glu Glu
     185     190     195
Asp His Glu Ala Asp Val Asp Val Leu Met
     200     205

```

<210> 21

<211> 833

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7469966CD1

<400> 21

```

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

```

WO 02/064792

PCT/US02/00254

```

Ile Ala Leu Tyr Cys Leu Lys Glu Leu Glu Arg Ser Tyr His Leu
245 250 255
Asp Lys Cys Gln Ile Ala Leu Asn Ile Leu Glu Glu Asn Leu Phe
260 265 270
Tyr Asp Ser Gly Ile Gly Arg Asp Lys Phe Gln His Asp Val Gln
275 280 285
Thr Leu Leu Leu Thr Arg Gln Arg Lys Val Tyr Gly Asp Glu Thr
290 295 300
Asp Thr Leu Phe Ser Pro Leu Met Glu Ala Leu Gln Asn Lys Asp
305 310 315
Ile Glu Lys Val Leu Ser Ala Gly Ser Arg Arg Phe Pro Gln Asn
320 325 330
Ala Phe Ile Cys Gln Ala Leu Ala Arg His Phe Tyr Ile Lys Glu
335 340 345
Lys Asp Phe Asn Thr Ala Leu Asp Trp Ala Arg Gln Ala Lys Met
350 355 360
Lys Ala Pro Lys Asn Ser Tyr Ile Ser Asp Thr Leu Gly Gln Val
365 370 375
Tyr Lys Ser Glu Ile Lys Trp Trp Leu Asp Gly Asn Lys Asn Cys
380 385 390
Arg Ser Ile Thr Val Asn Asp Leu Thr His Leu Leu Glu Ala Ala
395 400 405
Glu Lys Ala Ser Arg Ala Phe Lys Glu Ser Gln Arg Gln Thr Asp
410 415 420
Ser Lys Asn Tyr Glu Thr Glu Asn Trp Ser Pro Gln Lys Ser Gln
425 430 435
Arg Arg Tyr Asp Met Tyr Asn Thr Ala Cys Phe Leu Gly Glu Ile
440 445 450
Glu Val Gly Leu Tyr Thr Ile Gln Ile Leu Gln Leu Thr Pro Phe
455 460 465
Phe His Lys Glu Asn Glu Leu Ser Lys Lys His Met Val Gln Phe
470 475 480
Leu Ser Gly Lys Trp Thr Ile Pro Pro Asp Pro Arg Asn Glu Cys
485 490 495
Tyr Leu Ala Leu Ser Lys Phe Thr Ser His Leu Lys Asn Leu Gln
500 505 510
Ser Asp Leu Lys Arg Cys Phe Asp Phe Phe Ile Asp Tyr Met Val
515 520 525
Leu Leu Lys Met Arg Tyr Thr Gln Lys Glu Ile Ala Glu Ile Met
530 535 540
Leu Ser Lys Lys Val Ser Arg Cys Phe Arg Lys Tyr Thr Glu Leu
545 550 555
Phe Cys His Leu Asp Pro Cys Leu Leu Gln Ser Lys Glu Ser Gln
560 565 570
Leu Leu Gln Glu Glu Asn Cys Arg Lys Lys Leu Glu Ala Leu Arg
575 580 585
Ala Asp Arg Phe Ala Gly Leu Leu Glu Tyr Leu Asn Pro Asn Tyr
590 595 600
Lys Asp Ala Thr Thr Met Glu Ser Ile Val Asn Glu Tyr Ala Phe
605 610 615
Leu Leu Gln Gln Asn Ser Lys Lys Pro Met Thr Asn Glu Lys Gln
620 625 630
Asn Ser Ile Leu Ala Asn Ile Ile Leu Ser Cys Leu Lys Pro Asn
635 640 645
Ser Lys Leu Ile Gln Pro Leu Thr Thr Leu Lys Lys Gln Leu Arg
650 655 660
Glu Val Leu Gln Phe Val Gly Leu Ser His Gln Tyr Pro Gly Pro
665 670 675
Tyr Phe Leu Ala Cys Leu Leu Phe Trp Pro Glu Asn Gln Glu Leu
680 685 690
Asp Gln Asp Ser Lys Leu Ile Glu Lys Tyr Val Ser Ser Leu Asn
695 700 705

```

WO 02/064792

PCT/US02/00254

```

Arg Ser Phe Arg Gly Gln Tyr Lys Arg Met Cys Arg Ser Lys Gln
710 715 720
Ala Ser Thr Leu Phe Tyr Leu Gly Lys Arg Lys Gly Leu Asn Ser
725 730 735
Ile Val His Lys Ala Lys Ile Glu Gln Tyr Phe Asp Lys Ala Gln
740 745 750
Asn Thr Asn Ser Leu Trp His Ser Gly Asp Val Trp Lys Lys Asn
755 760 765
Glu Val Lys Asp Leu Leu Arg Arg Leu Thr Gly Gln Ala Glu Gly
770 775 780
Lys Leu Ile Ser Val Glu Tyr Gly Thr Glu Glu Lys Ile Lys Ile
785 790 795
Pro Val Ile Ser Val Tyr Ser Gly Pro Leu Arg Ser Gly Arg Asn
800 805 810
Ile Glu Arg Val Ser Phe Tyr Leu Gly Phe Ser Ile Glu Gly Pro
815 820 825
Leu Ala Tyr Asp Ile Glu Val Ile
830

```

<210> 22
 <211> 655
 <212> FRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 6920129CD1

```

<400> 22
Met Ser Gln Arg Asp Gly Val Cys Gly Ser His Glu Val Ala Gly
1 5 10 15
Ala Ala Ser Pro Gly Ala Asp Gly Gly Leu Ser Leu Ala Ala Tyr
20 25 30
Cys Lys Asn Ser Val Asp Gly Leu Trp Tyr Cys Phe Asp Asp Ser
35 40 45
Asp Val Gln Gln Leu Ser Glu Asp Glu Val Cys Thr Gln Thr Ala
50 55 60
Tyr Ile Leu Phe Tyr Gln Arg Arg Thr Ala Ile Pro Ser Trp Ser
65 70 75
Ala Asn Ser Ser Val Ala Gly Ser Thr Ser Ser Ser Leu Cys Glu
80 85 90
His Trp Val Ser Arg Leu Pro Gly Ser Lys Pro Ala Ser Val Thr
95 100 105
Ser Ala Ala Ser Ser Arg Arg Thr Ser Leu Ala Ser Leu Ser Glu
110 115 120
Ser Val Glu Met Thr Gly Glu Arg Ser Glu Asp Asp Gly Gly Phe
125 130 135
Ser Thr Arg Pro Phe Val Arg Ser Val Gln Arg Gln Ser Leu Ser
140 145 150
Ser Arg Ser Ser Val Thr Ser Pro Leu Ala Val Asn Glu Asn Cys
155 160 165
Met Arg Pro Ser Trp Ser Leu Ser Ala Lys Leu Gln Met Arg Ser
170 175 180
Asn Ser Pro Ser Arg Phe Ser Gly Asp Ser Pro Ile His Ser Ser
185 190 195
Ala Ser Thr Leu Glu Lys Ile Gly Glu Ala Ala Asp Asp Lys Val
200 205 210
Ser Ile Ser Cys Phe Gly Ser Leu Arg Asn Leu Ser Ser Ser Tyr
215 220 225
Gln Glu Pro Ser Asp Ser His Ser Leu Arg Glu His Lys Ala Val
230 235 240
Gly Arg Ala Pro Leu Ala Val Met Glu Gly Val Phe Lys Asp Glu

```

WO 02/064792

PCT/US02/00254

```

245          250          255
Ser Asp Thr Arg Arg Leu Asn Ser Ser Val Val Asp Thr Gln Ser
260          265          270
Lys His Ser Ala Gln Gly Asp Arg Leu Pro Pro Leu Ser Gly Pro
275          280          285
Phe Asp Asn Asn Asn Gln Ile Ala Tyr Val Asp Gln Ser Asp Ser
290          295          300
Val Asp Ser Ser Pro Val Lys Glu Val Lys Ala Pro Ser His Pro
305          310          315
Gly Ser Leu Ala Lys Lys Pro Glu Ser Thr Thr Lys Arg Ser Pro
320          325          330
Ser Ser Lys Gly Thr Ser Glu Pro Glu Lys Ser Leu Arg Lys Gly
335          340          345
Arg Pro Ala Leu Ala Ser Gln Glu Ser Ser Leu Ser Ser Thr Ser
350          355          360
Pro Ser Ser Pro Leu Pro Val Lys Val Ser Leu Lys Pro Ser Arg
365          370          375
Ser Arg Ser Lys Ala Asp Ser Ser Ser Arg Gly Ser Gly Arg His
380          385          390
Ser Ser Pro Ala Pro Ala Gln Pro Lys Lys Glu Ser Ser Pro Lys
395          400          405
Ser Gln Asp Ser Val Ser Ser Pro Ser Pro Gln Lys Gln Lys Ser
410          415          420
Ala Ser Ala Leu Thr Tyr Thr Ala Ser Ser Thr Ser Ala Lys Lys
425          430          435
Ala Ser Gly Pro Ala Thr Arg Ser Pro Phe Pro Pro Gly Lys Ser
440          445          450
Arg Thr Ser Asp His Ser Leu Ser Arg Glu Gly Ser Arg Gln Ser
455          460          465
Leu Gly Ser Asp Arg Ala Ser Ala Thr Ser Thr Ser Lys Pro Asn
470          475          480
Ser Pro Arg Val Ser Gln Ala Arg Ala Gly Glu Gly Arg Gly Ala
485          490          495
Gly Lys His Val Arg Ser Ser Ser Met Ala Ser Leu Arg Ser Pro
500          505          510
Ser Thr Ser Ile Lys Ser Gly Leu Lys Arg Asp Ser Lys Ser Glu
515          520          525
Asp Lys Gly Leu Ser Phe Phe Lys Ser Ala Leu Arg Gln Lys Glu
530          535          540
Thr Arg Arg Ser Thr Asp Leu Gly Lys Thr Ala Leu Leu Ser Lys
545          550          555
Lys Ala Gly Gly Ser Ser Val Lys Ser Val Cys Lys Asn Thr Gly
560          565          570
Asp Asp Glu Ala Glu Arg Gly His Gln Pro Pro Ala Ser Gln Gln
575          580          585
Pro Asn Ala Asn Thr Thr Gly Lys Glu Gln Leu Val Thr Lys Asp
590          595          600
Pro Ala Ser Ala Lys His Ser Leu Leu Ser Ala Arg Lys Ser Lys
605          610          615
Ser Ser Gln Leu Asp Ser Gly Val Pro Ser Pro Gly Gly Arg
620          625          630
Gln Ser Ala Glu Lys Ser Ser Lys Lys Leu Ser Ser Ser Met Gln
635          640          645
Thr Ser Ala Arg Pro Ser Gln Lys Pro Gln
650          655

```

```

<210> 23
<211> 330
<212> PRT
<213> Homo sapiens
<220>

```

WO 02/064792

PCT/US02/00254

```

<221> misc_feature
<223> Incyte ID No: 71514704CD1

<400> 23
Met Gly Ser Arg Lys Lys Glu Ile Ala Leu Gln Val Asn Ile Ser
1      5      10      15
Thr Gln Glu Leu Trp Glu Glu Met Leu Ser Ser Lys Gly Leu Thr
20     25     30
Val Val Asp Val Tyr Gln Gly Trp Cys Gly Pro Cys Lys Pro Val
35     40     45
Val Ser Leu Phe Gln Lys Met Arg Ile Glu Val Gly Leu Asp Leu
50     55     60
Leu His Phe Ala Leu Ala Glu Ala Asp Arg Leu Asp Val Leu Glu
65     70     75
Lys Tyr Arg Gly Lys Cys Glu Pro Thr Phe Leu Phe Tyr Ala Gly
80     85     90
Gly Glu Leu Val Ala Val Val Arg Gly Ala Asn Ala Pro Leu Leu
95     100    105
Gln Lys Thr Ile Leu Asp Gln Leu Glu Ala Glu Lys Lys Val Leu
110    115    120
Ala Glu Gly Arg Glu Arg Lys Val Ile Lys Asp Glu Ala Leu Ser
125    130    135
Asp Glu Asp Glu Cys Val Ser His Gly Lys Asn Asn Gly Glu Asp
140    145    150
Glu Asp Met Val Ser Ser Glu Arg Thr Cys Thr Leu Ala Ile Ile
155    160    165
Lys Pro Asp Ala Val Ala His Gly Lys Thr Asp Glu Ile Ile Met
170    175    180
Lys Ile Gln Glu Ala Gly Phe Glu Ile Leu Thr Asn Glu Glu Arg
185    190    195
Thr Met Thr Glu Ala Glu Val Arg Leu Phe Tyr Gln His Lys Ala
200    205    210
Gly Glu Glu Ala Phe Glu Lys Leu Val His His Met Cys Ser Gly
215    220    225
Pro Ser His Leu Leu Ile Leu Thr Arg Thr Glu Gly Phe Glu Asp
230    235    240
Val Val Thr Thr Trp Arg Thr Val Met Gly Pro Arg Asp Pro Asn
245    250    255
Val Ala Arg Arg Glu Gln Pro Glu Ser Leu Arg Ala Gln Tyr Gly
260    265    270
Thr Glu Met Pro Phe Asn Ala Val His Gly Ser Arg Asp Arg Glu
275    280    285
Asp Ala Asp Arg Glu Leu Ala Leu Leu Phe Pro Ser Leu Lys Phe
290    295    300
Ser Asp Lys Asp Thr Glu Ala Pro Gln Gly Gly Glu Ala Glu Ala
305    310    315
Thr Ala Gly Pro Thr Glu Ala Leu Cys Phe Pro Glu Asp Val Asp
320    325    330

<210> 24
<211> 276
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7715945CD1

<400> 24
Met Cys Ile Ile Phe Phe Lys Phe Asp Pro Arg Pro Val Ser Lys
1      5      10      15

```

WO 02/064792

PCT/US02/00254

```

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

```

<210> 25

<211> 232

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7025368CD1

<400> 25

```

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

```

WO 02/064792

PCT/US02/00254

```

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

```

```

<210> 26
<211> 120
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7500954CD1

```

```

<400> 26
Met Lys Asn Asp Thr Asp Thr Gly Asn Met Gln Lys Glu Val Met
1      5      10      15
Ser Val Thr Glu Gln Val Glu Lys Lys Lys Asn Asp Ile Glu Lys
20      25      30
Asp Asp Thr Gly Arg Lys Arg Lys Pro Asp Ile Ser Leu Leu Glu
35      40      45
Val Ile Val Asp Val Ala Met Lys Val Lys Lys Glu Ile Val Thr
50      55      60
Gly Asp Thr Asn Thr Lys Asn Leu Lys Glu Ala Lys Lys Glu Lys
65      70      75
Lys Arg Ala Val Ser Leu Pro Leu Asn Arg Arg Ala Pro Lys Leu
80      85      90
His Leu Gln Asn Arg His Gly Phe Gly Leu Leu Cys Ile Leu Val
95      100      105
Pro Glu Val Asp Thr Ile Asn Leu Val Ile Phe Leu Asp Asn Val
110      115      120

```

```

<210> 27
<211> 675
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 6242606CB1

```

```

<400> 27
attctgggca ttttoagagt tttctcttcc atgtgaattg ctccagctaa cttaatatta 60
gtttctcaag ggtttgtgta gacctggctt tgtctcttcc catctgttcc tctctttccc 120
tttaacgtca cttgctgag gtctataagg ccctttaagg ttgccagccc agagccagaa 180
gagatttggtg tccctgaagt taagaaagcc acctttgcct acagatgggtg ttcacagtgt 240
tggtgtggag tcacagcttg acctttgggg gccccaggag atgttgaccc agcagggcat 300
ggcgtgcag aactacgaca acaagctggg caaatgcata gaggagctat gccagaagca 360

```

WO 02/064792

PCT/US02/00254

```

ggaggagctg tgcctggcaga tccagcagga ggaggacaag aaacagcggc tgcagaatga 420
ggtgaggcag ctgacagaga agctggcctg cgtcaacagag aagctggccc gctcaacga 480
gaacctggca cgaagattg cctcttgcag taagtctctac cagaccatcg cggagacgga 540
ggccacctac ctcaagatgc tggagagctc ccagactttg ctgagtgctc tgaagaggga 600
agctgggaac ctgaccaagg ctacagcctc agaccagaaa agtagtggtg gcaggggacag 660
ctgaccagac catgg                                     675

<210> 28
<211> 1387
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 997105CB1

<400> 28
gcaacattct attttgcgta ggtgtcggga gagctttctt ttgcgatttg ttctctgac 60
cggaagaaaa gatagcaaat atcaacttgc attggattgc atcataaggg ttgggcttta 120
gcgagccagt cactctggca accggacgtg acgaccccg aaagcggccc accgagcttt 180
gctccagcta teggtctcag tgggcgggcg ggccagagcca gactgaagca tcttcagcac 240
tgaccaggtc ctgttggggg tgggtgcacc aaggccacgc cccctggccc tgctccggcg 300
gagctctggg agttgtcgtg ggtctcgatg ggggagactc acgcccgttc tctctcagcg 360
tagggctgtg agggcgaggt ccacggccag gagctctcgg gtttctggcg ggtgcgcgac 420
ctgtcccgga agctctggga tacaacctct cgagagcccg ctgtccctct ccgttaaggt 480
gcaacccctc acagttgctg tgggcaactc cagcccaaca ttcctctcgt ctggttctcg 540
cccattggg aaactcggcc ccacgcttcc cactttctgt gatgaggtgt cccctttctc 600
cccactaaaa tgtcaataaa cctacggagg gtcttcttga aaocgcaga ggaaaaattc 660
ggcaacgcct cgcgttgtgt ttoaggctgc atgtaccaag tagttcagac gattggctcg 720
gatggaaaaa atcttctgca attacttcca atctctaagt cttctggaaa tcttatacca 780
ctagttcaat cttcagtcac gtctgatgct ttgaaaggga atacaggaaa accagttcaa 840
gttacttttc agactcagat ttccagctct tccacaagt catcagttca attgccatt 900
tttcagccag ccagttcttc aaactatttt cttacaagaa cagtagatac atcagaaaaa 960
ggtagagtta cttctgtggg aactggaat ttttcttcat cagtttctaa agttcagagt 1020
catgggtgta aaattgatgg actcaccatg caaacatttg ctgttctctc ctaaacacaa 1080
aaagactcat cttttattgt agttaatacc cagagtcttc cagtgaactg gaagtctcca 1140
gttttgcctt ctgggcatca tttaacagat ccagcccatg ctgaagtga atctgtacca 1200
gcgtcatcat tgcctccttc agtgcagcaa aagatacttg caactggcac ccaagttacc 1260
tcaggatgg ttgaggctc ccaaatgcca accgttatt atgtatctc tgaatcctt 1320
acaacttggg agtctgagga aggaggtaca cttgaggcca ggagttcgga tggcaatgg 1380
caataga                                     1387

<210> 29
<211> 1534
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1482843CB1

<400> 29
ggttttgagg attgggaact ctagggtatt tttttttttg ccggcccctc caggtttatt 60
aggttgcat caggaccttg aatagttcca cagggtcaca taagtgaat tagaatcgt 120
tcoccaaaaa gctatttaca ttaatccagg gtttcattaa aaaaaactag caaacatcac 180
atctagaaa ccaacaatg cacaatgaac cgcctggggg tctgcagaaa ctacagtgtg 240
gctctcttgg cccacagggg cagattccag aggtgcocag caagcccagc tctcagttcc 300
ctacaggcca caccocccgc gacgcagcga caggtctcgg ggagggggcg gctcgccct 360
cggaagaagt ggcctggtca ttttcatgag gaattcaatc ccggagggaag gcgagaggca 420
ctcaggtctt tcatctggg aatgaggaca gtgtggcaag gctgcagta cctgcagag 480
acaccccag ctctcaccga gaggcccctc tctgcccaca gctctggccg aggcactgaa 540
ccaggacact ggcttgcatg cccagagaca gtggccnagg ggaccaacat ggaagaggaa 600

```

WO 02/064792

PCT/US02/00254

```

ggcagcagac cagattgcag ggcgggactg gccacacagt ggcctgggtg aggggtgggt 660
ctgtctacgt aatagtggac agtggccacc aggcctgagg ggtgcctcgg tggtagctgg 720
tgccataacc tttagactg atggggctga gccctagctc taccaagggg tgccctcgt 780
cctgtgaggc cagtcacagg ccaggttctc ggggggtgtg cacacctcac acacacatcc 840
ccgcctctta ccgcactggc tgtcagatgg ctttggtttc tagttgtcct tggtaggaca 900
gagtgctcct ccagtcaccc tcagaacacgc tctgtgaggt aggcaggaaa aatgcgacca 960
tccctgtata ccagtgggaa aaccagggca cagagggctg gcattggcca ggtcttgtct 1020
gcccaacctc cctggtgctg ggcattggtg gggggatgga ccgtggggct cctgggtgca 1080
gggcagaccc acagcccgag agcagcaggg gcacagggag tgccaccaga accccagct 1140
gctggatgga gactctgcc atgtctctcc cttgcgtggt taaggatgga gctgtgcagc 1200
cagatgggag gtaagggtga ggcaggaggg ctccaggaaa ggggctgcc acagccagga 1260
tgacctaggc agcctgtggc ccttctctct aaaaacctgt ctgaggccag aaagccag 1320
caaggatggt tccaccttca totcctact gagtgtgacc cttgggaaag aaacaggaa 1380
agcaggcagg acgggtgac tcaagcctgt aatccacgca ctttgggagg ccgaggcggg 1440
tggtacactt gaggtcagga gttcaagac agcctggcca acatggcaaa acccattctc 1500
tactaaaaat acaaaaaaaa tagccagggt tggt 1534

```

<210> 30

<211> 4203

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Ineye ID No: 3218219CB1

<400> 30

```

ccgggctccg tccgtgttaa ccgcagctct cagaagccta taacaggcta ccatttttat 60
gaagctcaaa aacgaaaatg aatttagtga tacagtgttt agacacatac tttgttaaga 120
aaacttaaaa acttggggga accataaaca taaaattgaa gacagtgttt accottgggg 180
tgagggaagc tatgcgggat agaccgggct agaattctga gtgcgggaga gtgctggacc 240
ctgagtgatg gggccggcgc cagctggaga gcagcttccg ggagcggact gagagccaga 300
gggtgagtaa ccagccctgg aaggcacagg caaagaaggg aagaaagcat cctccaggaa 360
gcgtacattg gctgaacctc cagcgaaggg cctcctgcag ccagtggaag tcagcagggc 420
agaaactgtac aaggagoccta ccaatgagga gcttaatcgc cttcgggaga ctgagatctt 480
gttccactcc agcttgtctc gtttacaggt agaggagcta ctaagggaag taaggctgtc 540
agagaagaag aaggatcgga ttgatgcctt cctacgggag gtcaaccagc gggttgtgag 600
ggtgcoccta gtccttgaga cagagctcac tgaccaggca tggctccagc ctgggggttcg 660
agtgcccttc caccagtgcc cctatgcctg gaagggctgt tcccgcttcc tgcccccagc 720
ccaggttaact gttgtgggca gctaccttct gggcaacctg atccgaccag acatcaatgt 780
ggatgtggca ctgacctatg ccagggaat cctacaggac aaggacgggc tgaaccagcg 840
ctacttccgc aagcgtgcc cctacctggc ccaacttggt caccacctgg ccagggaacc 900
cctcttggcg agtgtttgtc totactacac aaatggctgc cacttgaacc cctcactgtt 960
gctgcggccg cgtgaaagg atgagcgctt ggtcactgta cgtctgcac cgtgccctcc 1020
acctgaettc ttcggcccg ggcgcttgct gcccaaccag aacaatgtgc gctctgctg 1080
gtaccgaggg cagagtcttg caggggatgg tagccagag cctcctaccc ccgctataa 1140
cacatgggtc ctgcaagata cagttctcga gtccatttg cagctgtctg caaccattct 1200
gagttcagcc cagggcctga aggatggcgt ggcaactctg aaggtctggt tgccgcaagc 1260
ggagctggac aagggcaggg gtgggtttac tgggttccct gctccatgct tgggtgtctt 1320
ccttgtgtct acacgcaaga tccataccac catgaagtgc taccaggtcc tgagaagtgt 1380
cttgacgttt ctggccacta cagacctgac agtcaacggg atcagtttat gtctcagctc 1440
agatccctct ttgcggccc ttgctgactt ccaccaggcc ttctcgttg tottctctga 1500
ttcttcaggc catctcaacc tctgtctgta tgtcaactgc tctacttacc accaggtaca 1560
cagtaggga cggctgtcta tgatgtgtgt ggaacagaga gctgagagc ggttccact 1620
gcgtgtgatg atcccaaac caatgatccg ggttttgac caktctctgc atctccgtcc 1680
actgagtgc ctgcaggcag cgtgcacagc gctgaagctc tggccagagc tgcaggacaa 1740
tggtagggac tatgtctcag ctgctttggg cccctgacc accctcctgg agcagggtct 1800
gggggctccg ctgaacctgc tggctacact tcgaccccca gtcocagagt gggacatcag 1860
ccagatcca ccaagcaca aagactctgg gacactgacc ctgggactcc ttctccggcc 1920
tgagggactg accagcttcc ttgagctggg tcagaggcca gaccagctg aggtctgata 1980
attccgcag ttctggggat ccgcctcgga gcttcggcgt ttccaggagc gagccattcg 2040
ggaagctgtg gctcggggag cagcctctat gtccagaag cgccttatte cccaccaggt 2100

```

WO 02/064792

PCT/US02/00254

```

gggtaccacat ctcttggcac tccatgtga catccagaa acctgtgtcc actatgtggg 2160
ggggcccccgt gatgcactta tccaaaggcct gaagagagacc tccagcacag gtgaggaggc 2220
cctggtagcg ggggtacgtt gtaacgacga cctcagtcgc ctactgtggg ggtagaggg 2280
tctccacactg acggtgtctg ctgttcaggg agctcaccga gtgctgtgct acacagaggt 2340
gttccacacca actccagttc gtccagcctt ctctctctat gagactctgc gggagcggtc 2400
ctcactgtcg ccccggtctg ataagccctg tccggcctac gtggagacca tgacctgggt 2460
tgtcaacctg gaggggagtg gccagtgccc acaggagcgt gaggcgtgtc agcgggtccc 2520
agctgccttc cagctgcgcc tggcagagct gttgacacaa cagcatgttc tgcagtccc 2580
tgccactgcc acccacacgg atgtccttaa ggtatggatt gtgtttcgga ttcgctgtgc 2640
ctatcagcgg gagccccaga tccatgaagg ggtgcagago ccagagggga tgatctcgt 2700
gagggaacaca gctgcctccc tccgccttga gagagacaca aggcagttgc cactgtcac 2760
cagtgcctcg cagggactgc agcagcagca cccagccttc tctgggtgtg cagggtggc 2820
caagcgtgtg gtgcgtgccc agcttcttgg tgggggttcc gctgatgaga gctgggatc 2880
gggtggcgtc gcccttttcc tgcacctga gcccttcacc cctccagtt ccccccaggt 2940
tggtctcctt cgtatctctt tctgtgtatc aacgtttgat tggagaacaa acccctctt 3000
tgtcaacctc aataatgagc tcactgtgga ggaagcagtg gagatccgca gtggtctct 3060
ggcagctcgg gccacagctc cgttcattgt cattgttacc ccccaagacc gcaaaaaact 3120
tgtgtggaca caggatggac cctcagccca gatcctgcag cagcttgttg tctggcagc 3180
tgaagccctg cccatgttag agaagcagct catggatccc cggggaactg gggacatcag 3240
gacagtgttc cggccgcccct tggacattta cgaagtgtctg attcgcctat ctctccgca 3300
tatcccggg caccgcagg ctgtggactc gccagctgct tctctctgcc ggggcctgct 3360
cagccagcgg ggcctcctat cctgtatgcc cgtgctgggc tatgatcttc ctacgtcta 3420
tctgacgcag ctccaggagg ccttggggga tctggccctt tctctctatg accagcatgg 3480
tggaaggggtg attggtgtcc tctggaagcc caccagcttc cagccgcagc ccttcaaggc 3540
ctccagcaca aaggggcgca tgggtgatgc tccaggttgg gagctagtaa tgggtcccaa 3600
tgttgaagca atcctggagg actttgtctg cctgggtgaa ggcctgtgtc agactgtgga 3660
ggcccgaaat gagaggtgga ctgtgtgctc ccagctctgg agcaagctgt agacggaacg 3720
caggacatgt gacctctaga gcaagatgtc agttagatga cctccaccct ccttgacat 3780
gaatcctcca tggagggcct gctggctgaa catgctgaat catctccaac aaaaccacgc 3840
cccaacttcc tctctatgc tccagcattg gggcaggggc atggtgtccc atgtagtctc 3900
ctgggctcca ccatcccaag agaggagtgg gagccagctc agagaaggaa ctgaaccacg 3960
gagatccatc cactattag cctgggccc tggacctcct cgcatttccc actccttctc 4020
tagtctctct ccagaaacag agaaggggat gtgtgcttgg gagaggtctc gtctcctcc 4080
tgtgtccagg acctgtgctt gaccttagca tgccttccac tgcagtgtca ggcctttaga 4140
tgggacccag cgaataatgt gccctcttga gtcacatcac cgcactgag cagtggaaag 4200
ggg 4203

```

```

<210> 31
<211> 2082
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7371678CB1

```

```

<400> 31
cggttgctct ggagagggtt gtggcactcg gggctgggtg gctctcccta agcttgcctc 60
gacaaaagag ttttgagttg gtcttttggc tgagcctgcc caggaaaggca ggcctctgca 120
ggagctcttg aggttcggat gcgggcggcg atgaggatga ggcgtggcgg aagatgctt 180
tggtctctga gacactgcat cgggcagcag gggactctgg gagctgtgtg cagccagaag 240
gcattgctct tgacagcctt ctatgagaat ctctggaatt gtgcatgtcc ccccgccctc 300
cagccagcct gtgcagactt gctgcctgct gtgtcacccg gaacgcaag gctgggaaga 360
agggccctct caaatggac tgggtgttga ggttgagaag ctgccccctg acttcatgcc 420
aaagctgtgc aagaatctcc taggcagatg gctctgttgg gtctgccaga gttgccgaaa 480
gagcatggag gaagatgaaa ggcagacagg tccagacatc gcagtggaga tctccttgtc 540
acacacatcc tgcacatcac agtcttcttg agatgactct cattgtctct cgtcttcttc 600
ctcatcatcc tcatctctgt cctctcttcc ctgcccctgg aactcgggag actgggatcc 660
tagctgttcc ctgtcggcac ataagctctc gggcctctgg aattccccac attccagtgg 720
ggccatgcca ggcagctctc ttgggagttc tctaccatcc cccggtgagg ctltccccgt 780
ctcggagcac caccagcact cagacctcac tgcctcccc taccagccca ccggccacca 840
cccgcagcca gcatctctaa tcccgctcca cccagctcc tttgctctcc caccacca 900

```

WO 02/064792

PCT/US02/00254

```

acacctgctg cccaccaccc cggcagcacc tttccctgcc caggcttcag agtgccctgt 960
tgctgctgcc actgcccacc acactccagg gccatgtcag agtccccatc taccctccac 1020
cagcatggcg ctcctgaaga tgcccaccac attctcgggg tgacgccacc cctgcaagcg 1080
gcactgtggt gggcactgca gtgggcctct cctccacacc ccgagctctc agccactccc 1140
tagcactcac agggatcccg ggtgcaaggg gcacaagttt gcacacagtg gctggcttg 1200
ccagctgcc cagccctgcy aggcagatga ggggtgggt gaggaagagg atagcagctc 1260
tgagcgaagc tctgcacct catctccac cccacagaga gatgggagat tctgtgactg 1320
ctgctactgt gagtctctcg gccacaaatgc ggtgagtga cctgcccagg ctgagagagg 1380
ctggggggcga gccaaagtgga gcaaaagag caccctgtct tcccacaga ccccgctgct 1440
caggtgtgta attaccgtgc tggggccgtg agtctctgca cctgctctg ctgctgtgat 1500
cctccacaga catttgtgag gtgggggttt ttaaacattt tgctaccttt gctcttctca 1560
tggttttctc taagattgac ttttttctct taggcaaaag gaaaggaaat gccagagaga 1620
aagctatgat tctgatgagt atgtatacgt gtgtaatccc agagaagtga acgcttgga 1680
gtgatgaagg cagagtggaa gcaaaaaggc tctcagtcct caagtgatga cagccagccg 1740
agggacaggc cgtgagcaca gacggcgcca ggaaggaggc tcagatcaga gggcatgctg 1800
cgtctggcca gggggaggaa gcagtgcaga agtctcataa gccaccgct gccccaga 1860
gtcgggaacta taccgagatc cgggagaagc tccgtctgag gctgaccagg cggaaagagg 1920
agctgcccac gaaggggggc accctggggc ggtatccctg ggaagccccc gtggaaccac 1980
gagatgtgga tgagctgctg gaattcatca acagcacgga gcccaaaagt cccnacagcg 2040
ccaggggccg caagcggggc cggcacaagc tgaaaaaaa aa 2082

<210> 32
<211> 1459
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7473883CB1

<220>
<221> unsure
<222> 1438
<223> a, t, c, g, or other

<400> 32
gctacctagt gggctctggg gaccttcgaa atcgcccgcc ctctcacaat ggttgggtc 60
cagactgcgc caccagcctc cgggagaagt gggcctcgg aaacttttta gtgcggact 120
ccggggcgca gccagtcctc cggcagcagg atcacagctc ttgaggctag tgcgcctcta 180
ggcaagatgt ccttgcccat cgggatatac cgcggggcag tcagctatga tgataccctc 240
gaggacccct cggccatgac tctctctcca tcggacatgt gcagcgtccc ttggaagcca 300
gtgatctccag agcgcaagta tcagcacctc gccaaagtg aggaaggaga ggccagtcta 360
ccctcccctg ccatgacctc gtcatcagcc attgacagtg tggacaaggt cccagtgggt 420
aaggctaaag ctaccatgt catcatgaat tctctgatca caaaacagac ccaggaaaagc 480
attcagcatt ttgagcgaca ggcaggggctg agagatgctg gctacacacc ccacaaagggc 540
ctcaccaccg agggaccaca gtaccttcga gtggccgaag cactccacaa actaaagtta 600
cagagtggag aggttaacaaa agaagagagg cagcctgcac cagcccaagc caccccaagc 660
accactccgc actcttcacc taagcagagg cccaggggct ggttcacttc tggttcttcc 720
acagccttac ctggcccaaa tctagcacc atggactctg gaagtgggga taaggacaga 780
aaotttgtcag ataagtggag cctctttgga ccgagatccc ttcayaagta cgattctgga 840
agtttttgca cccaggctcta ccgaggagcc cagaagccct ctccattgga actgatacgt 900
gcccaaggca accgaatggc tgaagatcca gcagcctga agcccccaaa gatggacac 960
ccagtgtatg aaggaaagaa acagccaaca cgggcccata acctcaaac ccgtgacctg 1020
aatgtgctca caccactgg cttctagagc cctctttcca gggatctctg taaagtggt 1080
ttctgtctc ccactccctc tttacctgg ctttgacata ggaaggtat atttaaaac 1140
ttaatcagct gggcgtgggt gctcagccct gtaatcccaag caatttgga ggcacaagta 1200
ggttgatacc tgagtgacgg agttcaagac cagcctggcc aacatgtgta aaccccgctc 1260
ctactaaaaa taaaaaatt agctgggcgt ggtgtgggccc cctgtagtc ccagctactt 1320
gggagggctga gccaggagaa tgcctgaac ccagggaaga gatgtgtgac cagctgaga 1380
tcatgccatt aactccagc ctgggcgaca gaatcgagac ttccagcac actgcgcncg 1440
tacaagtgag ccgagctcg 1459

```

WO 02/064792

PCT/US02/00254

<210> 33
 <211> 1806
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Ineyte ID No: 7478662CB1

<400> 33
 acttctctcca tectgtccag caggatgggc tggacagtgg gacagcctgt gtgcacattt 60
 tgtggcaagt aggagtgaaca caccatccct gggaggaccc atgggtlctg ccaaacccaa 120
 ccccaagaact ctgtccctga ggtggtttta ccaaaaccca aaaccacaga ttgtgcttgt 180
 ggctcagggg tcagcacctg ctagtaccag gacactactg ggaggctggg acctgaccga 240
 agcccatggt gtctgtggcc taggacagg gtgtgttggg gccataagtc ccggccacca 300
 atggccatgt ggtctagggt cctcagcccc agtgtttggc cttccctggc tcctctgggt 360
 tcagtcccat tagggccctg gaggcccaaga cccagcacc aaggtccctt ccaggatatac 420
 tggcggtctg gcttctctta tcatgtttca tctgagagca aaaatgtcag atcggatgca 480
 cagaaaaatg gcccaaatgt ttaataaact agaagaaata taggagcagc aagagggtaa 540
 tatggagagg ggaggggctc catgaccggt gtctgcagag ccagggggtac aggcacccag 600
 tgtlgtggcc tggcaccacc ggcctctcag agcgtgggtg gccacgtgt ctttaaccgg 660
 aggaacagcag gccctgacac cagctttctt acctgtccct gtaagcaca cgttcttaga 720
 agaaaatctc atgcacagac ttgcacccat cctagcttgg gggttagggg ttgtctcttg 780
 gtgacctaaa tgaaaaaata ggtccagatc agagtctctg accgagagca ctacccact 840
 ctttgaatcg tgggagggga ggcctggttt tagttaaac taacctcttt gaggaaccac 900
 agagcccaag actggaagcc ttcagaatct tctggccccc aacctctctt ggggacccct 960
 gtggccctgtc tcaccagagc actcttctgt ctgtagatgt ctcagctgct ctacaaggga 1020
 gtcccatctc agtgtgtggg ctgggcctag tcaactctgc tggatgtcta gaaggtggaa 1080
 actaaggacc taggaaaatg ccagatacag cctttccacc ctcattccaga gcaggacaaa 1140
 caggcccggt ggtgtcagga gccacaggtct ccagctggag ggaacgtcaa cctgcagtg 1200
 ggagccagggt cccatcgcac atccctaggca cagatgctaa tgcaggcact gcaggtaagc 1260
 tgggcttgggt atccttccct ggcctcagaa agaagccaac aaggagcgtt ttgcagaaatg 1320
 aaacctttgt ttccagaagc actgctgact gtaagtgttt gccgttttgt gcagttagca 1380
 ttttgtccat tctgaggttg gattgttttc tctttttggc cttgcccctg cctacagacc 1440
 ataaaggaga acagcaagaa gcccccagca aacatccaca gatggccctg gacatcagcc 1500
 acattctgag gaacatgtca tgttctggga gggctaagc atcaagtaag gcctgtgggg 1560
 ctggaggatc ccaggcaagg tggggcaatc cagagccatg ggggcttccc atgggaattg 1620
 ggaggtccca aggcagagtc agaggttcca caggaggagt cagagagtca ccaagggttc 1680
 tccgtggcca gggagcagtc aacacatgg actgaacact tgcgtgggtc caaccttgg 1740
 gccagcgtgc ccatgtgggg ccaggaggca gctcagagtg ggagacagag agacaagtgt 1800
 gctcag 1806

<210> 34
 <211> 1698
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Ineyte ID No: 7650474CB1

<400> 34
 cactgttcca ggacgaattg ttgattccca ggaggggccc ctttccggtc tgggtccccc 60
 agaggactgc cttgctcacc tgtccctctg ggcggggccc gggagactcc cgaagagccc 120
 ccgggatcgc tggcctctccg aactccacag caatgagcaa gttgggcaag ttctttaaag 180
 gggggggctc ttctaaagag cgagccgctc ccagtcacca ggaggccctg gtccgacttc 240
 ggagactcga ggagatgctg ggcagaagaa aagagtacct ggaaaatoga atccagagag 300
 aaatgccctt gcccaagaag caccgcacgc agaataagcg agctgcatta caggcactaa 360
 agagaagaag gaggttctag aaacagctca ctcagattga tggcacactt tctaccattg 420
 agttccagag agaagccctg gagaactcac acaccaacac tgaggtgttg aggaacattg 480
 gctttgcagc aaaagcagtg aaatctgttc atgaaaacat ggatctgaac aaaatagatg 540
 atttgatgca agagatcaca gagcaacagg atatcgccca agaatctca gaagcatttt 600

WO 02/064792

PCT/US02/00254

```
ctcaacgggt tggcttttgt gatgactttg atgaggatga gttgatggca gaacttgaag 660
aattggaaca ggaggaaatta aataagaaga tgacaaatat cggccttcca aatgtgcctt 720
cctctttctc cccagcacag ccaaatagaa aaccaggcat gtgtccact gcagtcgat 780
cccagcagc atcttcccag agggcagaa aaggagatga tgatatcaaa caattggcag 840
cttgggtac ctaaaactaaa acacattttt gatacctaaa ttaatgagct atagataaaa 900
tataaaaaat gtttttacc agttcagaa gtacaaaga ctgtgttta taattatatt 960
gaatgaataa ttgtgttta agcctcttaa gtanaagtaa aaaaggagtc atgtgcatac 1020
atagaalccag tgaaggaggg caggcaggtt atctcatgoc tgaataccca gctacttggg 1080
aggctgagtc aggaagatcg cttgagccca ggagatggag gttgcagtga gccaaagatca 1140
tgccactgca ctccagactg ggcaacagag ggagactcgg tctcaaaaaa taanaaaaaa 1200
aaatacattt agtatagcgg cgggtggggg ggagaaataa tgttatttcc tatgcgaagt 1260
acgtgtatcc ctgtaccatc ggtaaatgta aatatactgt gtctcttttg ggagagcctt 1320
ttagtagagg agtcttatat gagtctctac ataagtagtt tcaactgagt tttgcagttt 1380
gaaatcttaa agagcttta attgacattt attataccaa ttaagcttgg aatggggcaa 1440
tgagtgcatc tcccaaaacg tgtgaagca ctaacagctt atattgctga atgagaatct 1500
cctgggtgta atttagccac ttagggaaact ggtgaacac tcccaggcca ttatgatgct 1560
gttacagctt cagtgtataa atgcataagt attctttctg ttctgttttg tgcctcttgg 1620
taacatttatt taacctttac agaattattc ttgtaataac ataaaaatat tggcaattaa 1680
aagtacatct tgaataaa 1698
```

<210> 35

<211> 1753

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 8092378CB1

<400> 35

```
tttcacataa ctagaagggt atagacagta tgacagtgya ataccttttg ttctcattta 60
ggacagtgtc tgaattgtag tgtgagagac tlaagtaga taatttagtg catttccatg 120
tagtgagagt tgttaattgt taacacttgg aaggaaatgg tatcataagg aattgacatc 180
ataagttttg gaaatctttt ttgtctagga tgcctttatg cctatctgaa ttgaaagt 240
tagttagatc taactgtttt tttagcatat tattctgttg tattccallt aggtatttgt 300
aattatccag aatgcagacc ccacgatata gaggtacaaa ttatatctct tccaaactaa 360
taccanaact atattatgtg cttttctttg taanaacgti tgttagggtt gtgctatttg 420
aacatatttt ctggtggata ttctcagctt tgaacaactg acatttactg gtaagatttt 480
atttataaat taccatattt aaggaaatct ctatacttct aanaactcga atattttttg 540
aaagaggtaa ttgctatata ttgcttttcc ttctaaaata ttctaaalat atctgttggc 600
taactgtttt tglatggaga agtgaaggca ttgagattaa ctttctcata cttacaaaaa 660
ggctgtgttg tgggtctgaa aataagaccg gtgtcctgac tcccagtttg ttttctgtgt 720
attagaccag actgcttact taaagtttta ttgcgtataa aatctgttca taagtttgag 780
ctccattttt ttctgtccat tattacaaca tagagaagca cattcatatc cccagagaaa 840
tttagtgata tgcagcaact ttgtttcact cttgacatta agtgtacatt gttaacattt 900
tttatcttat gattgaatgt ttcaggtttt cctcctccac caggcgtccc acctccatct 960
cttataccaa caatagaag tggacattcc ttgtgttatg atagtgttc tgcacgtgca 1020
tttccatagc gcaatgttgc ctttcccatc cttctgtgtt ctgtctcttc gtggcctagt 1080
cttgtggaca ccagcaagca gtgggactat tatgccagaa gagagaaga cccagataga 1140
gagagagaca gagacagaga gcgagaccgt gatcgggaca gagaagaga acgcaccaga 1200
ggagagagga gggagcgtga tcaagtcctt acaccaagtg ttctcaacag cagtgaaga 1260
cgatacagat caagggcaat tgcagaagga ggttatgagc gtcacagagc aagtgcagaa 1320
aaagaagaa cactatagaga aagacgacac agggagaaag aggaacccag acatagtct 1380
ttctgagnta atagtagacc tcccatgaa agtgaagaa gagatagtc caggagccac 1440
aaacacaaaa aatctaaaag angcaagaa ggaaaagaa cgggcagtga gcttgcctct 1500
gaacaggaga gcaccgaagc tacacctgca gaataggcat ggttttggcc ttttgtgtat 1560
attagtacca gaagtagata ctataaatct tgttattttt ctggataatg ttttaagaaat 1620
ttaccttaaa tctgtttctg ttgttagata tgaagaattt accttttttt ccaaaataaa 1680
agagtgaatt ttctatgta agttaaaaa ctttgtcttg taactattca aaaaataaaa 1740
gcagcgaatg att 1753
```

<210> 36

WO 02/064792

PCT/US02/00254

```
<211> 1489
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1263178CB1

<400> 36
cgccgggtggg cgcagatga agaggaggcg gtggcagtgg tggagaaga ggcggcggcg 60
gcgggggttag ggagcctgga aacgcgagcg gggatggcct ttggacaac tcttctgaaa 120
ggtggggaac ctagaagcc agttgatgcc acctgtatt ggagttgtgg gttggctggc 180
tcctctggct gatggcatgt tgggtacat gggccagcgg cagcaggggc atccaatcca 240
gaggggtcca ctctagaggc caggccacca gcaccatgg ccagtggtgc accaagtgtc 300
agaatccctc atcgaccctg ggcgcaaaa atggagaccg tgagccccag aacaagtcac 360
atagcaggcg ggggtcaggc caccgtgagg agcaggtaac acctgtggc aagccaggty 420
gagatctcct cgtcaacggg accaagaagg ccgaggctgc cactgaggcc tgcagctgc 480
caacgtcttc gggagatgct gggaggagg ccaagtcca tgcaggagg tcttctctgc 540
aaagattgga agaactgttc aggcgtatca aggatgagcg ggaagatgca attttggagg 600
aaggcatgga ggccttttgc aatgacctgt gtgttgacc caccagaatt cgagtgtctg 660
tcttggcttg gaagtccag cctgcaacca tgtgcaaat caccaggag gagttttttg 720
atggctgcaa agcaataagt gcagacagca ttgacggaat ctgtgcagg ttcctagcc 780
tcttaacaga agccaaacaa gaggataaat tcaaggatct ctaccggtt acatttcagt 840
ttggctctga ctctgaagaa gggcagcggc cactgcactg ggaatatagg attgcccgt 900
ggaactagt cttaaccag aacaatcttc cgttattgga ccaatggcta aacttctaa 960
cagagaaccc ctccgggac aagggcactc cccgggacac ttggacact ttccttaac 1020
tcactcaggt gattggccct gacctcagca actacagtga agatgaggc tggccaagtc 1080
tctttgacac ctttgtggag tgggaaatgg agcgaaggaa aagagaagg gaaggagag 1140
gtgcactcag ctccaggcct gagggcctgt gtcccagga gcagacttag tggctctgtc 1200
ccaggagcag cagcaaggat ctgcccagct cctgcagccc aactgaggaa ttggaccatt 1260
ttggaattaa ctgaagatcc ggatatttct tactttacac ctttctctgc cttgtatctg 1320
aaagggctct aaaaatgctg atcatgtttt aggcacttct ttcatttttt tggatttttt 1380
ggttatttcc tttttggggg gatctcccag aatatttgaa cctggttaca tgttgtgtat 1440
ctttttttga agccttcaga tagaataagc ctgccatttc ttgcacaaa 1489

<210> 37
<211> 1383
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3535417CB1

<400> 37
gggagggtcac aatcacattg agccaaaacg catccagtg tttctccagt tacaaaataa 60
acgaatatgc ccatgctgct accacatcct caccagcatt tccataaagg ccttttaaga 120
gcacctttcc gatgttacca ctctactctt cactcaagta ctactctcgg atcaggaaac 180
ccatgtgtgc agccgtttta ttctcttgga ctccattgta caaagtggat gctgctgtca 240
gatggcttaa agagaaaatt atgtgtacaa acaactttaa aggcacacac agaaggactt 300
tcgtataaag agcaagatt tgtggataaa cttatactg gtttaacca agggcaagg 360
gctgttttag cagaggccat aactcttgta gaatacaact acagcaggaa aaaggagtta 420
gccagggtgc ttcttcagaa agtattactt taccacagag aacaagaaca atcaaataaa 480
ggaaaaaccac tagcatttgc agtaggattg tctgggcccc ctgggtctgg aaaaatcaaa 540
tttatagaat attttgaaa aatgcttact gagagagggc acaattatc tgtgctagct 600
gtggaccctt cttctgtac tagtgttgga tcaactctag gtgataaac ccgaatgact 660
gagttatcaa gagatagaa tgcatacacc aggccatctc ctactagagg aactttagga 720
ggcgtgacaa ggaccacaaa tgaagctatt ctgttgtgtg aaggagcggg atatgacata 780
attcttatgt aaaccgttgg tgtgggtcag tcggagtttg ctgttgctga catggttgac 840
atgtttgttl tactactgcc accagcagga ggagatgag tgcaggggtat caaaaggggt 900
ataatcgaga tggcagatct ggtagctgta actaaatctg atggagactt gattgtgcca 960
gctcgaagga tacaagcgga atatgtgagt gcactgaact tactccgcaa acgtttacaa 1020
```

WO 02/064792

PCT/US02/00254

```
gtctggaac caaaggtaat tegtatttct gccgaagtg gagaggggat ctctgaaatg 1080
tgggataaaa tgaagatttt ccaggacctt atgcttgcca gtggggagct gactgccaat 1140
cagcggagac aacagaaggt ttggatgtgg atctctattc aggsaagtg gtttagaat 1200
ttcaggacac acccagagc ccgggaacag attccacttc tggaaacaaa ggttctcatt 1260
ggggccctgt cccagagact agcagcagac ttcttgtaa aagcttttaa aagcagagac 1320
taataaatt catctgtat aataatttta catatcattt cataaagtat ttaatataga 1380
aaa 1383
```

<210> 38

<211> 3519

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7473555CB1

<400> 38

```
aaatgtgat gctacccagg tccctgcca cttctgtctt atcctgcata cttttctga 60
atcactcaac taatttttag ttccaacaaat tgcttgata tcatgacccc atacctctcc 120
caccgagctc cagaactgtta agttacatta ctccctggat tgtgcatgtt ccaactctag 180
ctcaggccac tccccacac cttattcttc ttgtactctg tgagcagcat cactgttccc 240
ccaagacagc cccctggaag ccatcttctg accaactcct tgacagtgtt ggccccagta 300
gccccaggtg tcttacctgt actgagaaga cagtgtctga tcatgggaaa tgcattctctg 360
aatgcccctg cgggtactat gctgatgcca ctggcaggtg caaagtltgt cataactcat 420
gtgcagctgt ctctgggccc acacctcttc actgtacagc ctgcagcccc ccaagagctc 480
tgcgtcaagg cactgtctgt ccccgctgtg gagaggggtt ctactctgac cagcagtgct 540
gcaaaagagc cagaggaagg actgcaagtg gagaagctgt ctggcgtggg catccctctc 600
ggcagtgctc tagccagtg tagagcccat tttacttgg agagcactgg cctatgtgaa 660
gaacacccag aactttggaa attggtattt gccatgtctg aactgcagat atgggaagga 720
aaaactatcc tgaactgtgt tcccaactac atcccaactc ctacgtatca acgtagaccc 780
agacattctg ctttacttac ctccaaactg actgttccca tccagagttg tgttaagcct 840
cctacactgc tgttagtagg aaactatcaaa atttgggcaa ataaccaat tgcacatgc 900
attaattgtc atctatacac ttgtattaac tccatttttg actccaggaa aagtgtaatg 960
ttggttcagc ctgagaaagg aactctgatt ccgcttgcca ccagctctgt ttcatgtgtg 1020
cagggaagag ccaacataac tgcacagact gtgggccttc ccagtgtctg ttgatgggc 1080
agtgcctctc ccagtgcaca gatggctact ttccacagga aggtagtgtg acagagtgtc 1140
accacacctg caggcaatgt catggccctt tggagctcga ctgcactctc tttacacctc 1200
acatctctct taccaatgtt aactgcagga ccagctgcag ggaagagcag ttctccaccc 1260
tcgtgggata ctgtctgac tgcacacacc tgtgccagca ctgtgcagct gatctccaca 1320
acactgggag catctgcttc aggtgcagca atgcccacta cctgtctctc ggggaccact 1380
gtgttctcga ctgccttcca ggatactatg cagagagagg agcttgtaaa aaatgccact 1440
ctctctcgag aaactgcagc ggcagaggac ctltctctct cctctcatgt gacacacacc 1500
tcgtgtctgc ccacactggc acctgcagca ccacctgctt cccctgggac tatcttgatg 1560
acaatcatgt ttgcagcacc ccagagccaa tggagtattg aagttaggag ggatgacct 1620
ttcttagacc tacagcagaa aacaagtctg ttaagaagg tgattgaacc ccccaactca 1680
catccagatg gtgatggta cagcacatcc cagcttccct cagtgggttc tggctcactc 1740
tgcacttttc tcaactgct cctgcaactg ctgtgggggt cactgaactc agagaactgc 1800
ccctgtactc ccaggtgttg ccacaccaaag atgtctggt cagtlactaa tgcacacacc 1860
actgtggaag ctgtgatcca caggccagct gtacctctct ccagatccca aacaaagttc 1920
tgcctcttgg ggaatgtcaa tacgagagct gcgcccaaca gtactatctt gaacttccca 1980
ccaacacgtg caaagtgcac ccgtagggtt aaacggtgta gctgtctctc ctgctggagt 2040
gaacaggatc tcttcagttc agcagtcctc accattccct ttctatcaca caaggcctat 2100
tctgaagata cattacttct ctgggctgta ggtgtttgca ttacaaactt gaggattgta 2160
cttccaagct tgataagtag ttggaattct cggggcttac acactttgaa tggtagtgct 2220
cagcaacagc aaagcaaaagc tgcagacaga gtccctcataa atgaactgtc gggctctgag 2280
gtggacagaa aggaagataa tctaatgcag actttctttt tagagtgtga ttggagctgc 2340
agtgcactga gtgggccccg gaaaaacagc tgctgcagtg gcatggatgg ctatgttctc 2400
caggtatggg cctgcttgga gcaagtgttg tcatcatttt accaggactc gggcctctgc 2460
aagaactgtg acagctactg tctccagtgc caaggtcccc atgagtgtac ccgctgcata 2520
gggcccattc tctcttgga agcccagtg gtccaggaaat gtgggaaggg gtaacttgca 2580
gatcatgcaa agcacaatag cacagcctgc cctcaggggt gcttgacagt cagccacagc 2640
```

WO 02/064792

PCT/US02/00254

```

gaccgttgtc accctctgtga ccacgggttc tttctgaaga gtggcctctg tgtttacaac 2700
tgtgttcctg gcttttctgt ccacacctct aatgaacat gttctggcaa aatacacacc 2760
cctagttctc atgtgaatgg ttccctgato ctccaattg gttcaataaa gccactggat 2820
ttttccctcc tgaatgtcca agaccaggag ggttaggtca aagatctcct attcactgtt 2880
gtgagcactc ccaccaatgg tcagctagtg ctctcaagaa atggaaaaaagggttcagctg 2940
gacaaaggctg gccgttttag ctggaaagat gtgaacgaga agaaagtgcg ttttgtgcac 3000
agcaaaagaaa aactcaggaa aggttacctt ttcttgaaaa ttagtgaaca gcagttcttc 3060
tctgagccac agctgatcaa catcaagca ttctcaaac aggcctctca tgtctgaga 3120
aatgaagttc tccacattag cagagagag agggcaacca taaccaccca gatgctgac 3180
atccgagatg atgacaaccc acaggtatgt gtcatgaaa taatcgatcc tccactcat 3240
ggccaaattgc ttccagacac tcagtccccc gcaaccccta tctatcaatt ccagctggat 3300
gaactctcta gaggccttct ccactatgct catgatggtt cagacagcac atccgatgtt 3360
cgagtctctg aggcacatga tggacactcc ttccataata tactgttcca agtgaagacc 3420
gtgcctcagg gccagttcta tgatggtatg tcttctcttg atctcgacc atcaagcagg 3480
gtgtgagag cctttatatt cctcatgtta gttccctga 3519

```

<210> 39

<211> 2763

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Inocyte ID No: 7474985CB1

<400> 39

```

atggcagatt oacatcaco ttctttcttt cctgattttg ggetgctatt gtattttggg 60
gagctaaaca aagaggaatt aaatcacctc aagttattcc taaaggagac catggaacct 120
gagcatggcc tgacaccctg gaatgaagtg aagaaggcca ggcgggagga cctggccaat 180
ttgatgaaga aatattatcc aggagagaaa gccctggagt tgtctctcaa aatctttggc 240
aagatgaacc tgaaggatct gttgtagaga gcgaagaag agatcaactg gtccggccag 300
actataggac cagatgatgc caaggctgga gagacacaag aagatcagga gccagtgctg 360
gtcalagtta acacaggggg ccccaactcc tgggccaacg accctactg gtccggcgcc 420
cctcgggaat caggtcgcat agccaggagt gatggaacag aatacagaaa tagaataaag 480
gaaaaatttt gcataccttg ggacaagaag tctttggctg gaaagcctga agatttccat 540
catggaattg cagagaaaga tagaaaaactg ttggaacact tgttogatgt ggaigtcaaa 600
acgggtgcac agccacagat cgtggtgctt caggagagctg ctggagltgg gaaaacaacc 660
ttggcgagaa aggcaatgtt agatgtggca gaggcgagtc tctaccagca gaggtttaag 720
tatgtttttt atctcaatgg gagagaattt aaccagctga aagagagaag ctttgtctaa 780
ttgalatcaa aggaactggcc cagccagaaa ggcctcattg aagaaatcat gtaccagcaa 840
agtagcctct tgtttattat tgacagtttc gatgaactga actttgcctt tgaagaacct 900
gagtttgcac tgtgcgaaga ctggacccaa gaacacccag tgccttctct catgagtagt 960
ttgctgagga aagtgatgct cctcaggcca tccttatttg tgacacaag actcaacct 1020
tctaagagac taaagcagtt gttgaagaat caccattatg tagagctact aggaatgtct 1080
gaggatgcaa gagaggagta tattaccag ttttttgaa ataagaggtg gcccatgaaa 1140
gtattcagtt cactaaaaag caatgagatg ctgtttagca tgtgccaaat cccctagt 1200
tgctgggccc cttgtacttg tctgaagcag caaatggaga aggggtggtg tgtcacattg 1260
acctgccaaa caaccacagc tctgtttacc tgctatattt ctagcttggt caaccagta 1320
gatggaggct ctctagtcct acccaaccaa gccagctgtg gaagactgtg ccaagtcgct 1380
gccaaaggaa tatggactat gaactacgtg ttttacagag aanaatctcag aaggcttggg 1440
ttaactcaat ctgatgtctc tagttttatg gacagcaata ttattcagaa ggacgcagag 1500
tatgaaaaact gctatgtgtt caaccacctt catgttcagg agttttttgc agctatgttc 1560
tatatgttga aaggcagttg ggaagctggg aacctctctc gccagctctt tgaagatttg 1620
aagtoactac ttcaagcaco aagtataaaa gaacccactt tgacacagat gaagtgtctt 1680
ttgtttggcc ttttgaatga agatcgagta aaacaactgg agaggacttt taactgtaaa 1740
atgtcactga agataaaatc aaagttaact cagtgtatgg aagtatagg aaacagtgc 1800
tattctccat cacagctggg attctcggag ttgtttcaot gctgtatga gactcaagat 1860
aaagcgttta taagccaggc aatgagatgt ttcccaagg ttgccattaa tatttgtgag 1920
aaaatacatt tgcttgatc ttctttctgc ettaagcact gccggtgttt gcggaccatc 1980
aggtcgtctg taactgttgt atttgagaag aagatattaa aaacaagcct cccaactaac 2040
acttggttgg aatactgtgg tttagacatc ctctgctgtc aagatctctc ctctgctctt 2100
atctgcaaca aaagactgat aaaaatgaat ctgacacaga ataccttagg atatgaagga 2160

```

WO 02/064792

PCT/US02/00254

```

attgtgaagt tatataaagt cttgaagtc cctaagtgta aactacaagt tctaggctta 2220
gaaagctgtg gtctcacaga ggtcggctgt gagtatcttt ctttggctct catcagcaat 2280
aaaagactga cacatttggt cttggcagac aatgtcttgg gtgatgggtg agtaaaagctt 2340
atgagtgatg ccttcgaaca tgcacaatgt actctgaaga gccttgattt gatgggctgt 2400
gttctcacta atgcattgtt tctggatctg gcttctgtta ttttgaalaa cccaaacctg 2460
aggagctggg accttgggaa caacgatttg caggatgatg gattgaaaaa tctgtgtgat 2520
gctttgagat atccaaactg taacattcag aggtcgggtt tgcctattac aaacacaaac 2580
ctggaaaagt gcccaagact tcaactcagt gaccagccca gaaggctgag aaagaagagc 2640
atggaaatgg tctcttttca aggtgcttgc tggctaccac acactattga tctgattgtg 2700
gactctctct tggccgatga aagtgaaggc attagcagtc atactgttct ccacattgtg 2760
tga

```

<210> 40

<211> 1023

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7475137CB1

<400> 40

```

atggggagaa accagagcgg aaaagctgaa aattctaaaa atcagagcgc ctctctctct 60
ccaaggact gcaactcccc gccagcaatg gaacaaagct ggatagagaa tgactttgac 120
gagttgacag aagtaggctt cagaaggctg gtaataacaa actactctct cgagctaaag 180
gagcatgttc aaacccattg caagaagctt aaaaaccttg aaaaatgggt agacgaatgg 240
ctaaatagaa taacacgtgt agagaagacc ttaaatgacc tgaaggagct gaaaacctg 300
gcacagaac ttcgtgaggt atgcacaagc ttcaatagcc aattcaatca agtggaagaa 360
aggggtgtcag tgattgaaga tcaaattaat gaataaagc gagaagacat ggttagagaa 420
aaaagatata aaagaacaaa acaaaccttc caagaaatat ggtactatgt gaaaagacca 480
aatctacatt tgattagtgt acctgaaact gatggggcga acagaaacca gttggaacac 540
actctctcat atattatcca gaacttgccc aaaaatagcaa ggcaggccca cactcaaat 600
caggaataac agagaacacc acaaaagatac tctctgagaa tagcaacccc aagcatgtta 660
gtcagattca ccaaggttaa aatgaaggaa aaaatgttaa cggcagcccg agagaaggt 720
cagggtaccc acaaggggaa gcccatcaga ctaacagcag atctcttggc agaaacctta 780
caagccagaa gacagtgggg gccaatattc aacattctta aagagaagaa ttttcaatcc 840
agaattccat atccagccaa actaagtttc ataagtgaag gagaataaaa atcttttaca 900
gaanaacaaa tgcagagaga ttttgcacac accagggctg ccttacaaga gtccttgag 960
gaagcactaa atatggaag gaacaaccag taaccagcac taaaaaaa tgccaaattg 1020
taa

```

<210> 41

<211> 2046

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5036986CB1

<400> 41

```

gtgtctctct ccaaggctgt agggattctg gagctgctgg ctggagagga ggggtggaca 60
agctctctcc agaaagacat cctgagagga cttggcaggg ctgaacatgc attggtctgc 120
aaaagttcag ggacttttga cctgtgtggg tactcagatg tccagccgca ctctctacat 180
taattgtagg caactgttgt cctgcagtg gggccaccag gaagtgcagg ccaagtttaa 240
ctttgtctgt gatgtgttgg atcactgggc tgacatggag aaggctggca agcgaactcc 300
aagccacagc ctgtgttggg tgaatgggaa ggggaaggaa ttaattgga atttcagaga 360
actgagtgaa aacagccagc aggcagccaa cgtctctctg ggagcctgtg gctgcagcag 420
tggggatcgt gtggcagtga tgcctgccc agtgccctgag tgggtggctg tgatcctggg 480
ctgcattcga gcaggtctca tctttatgcc tggaaaccac cagatgnaat ccactgacat 540
actgtatagg ttgcagatgt ctaaggccaa ggctattgtt gctgggggat aagtcattca 600
agaagtggac acagtggcat ctgaatgtcc ttctctgaga attaagctac tgggtgtctg 660

```

WO 02/064792

PCT/US02/00254

```

gaaaagctgc gatgggtggc tgaacttcaa gaaactacta aatgaggcat ccaccactca 720
tcactgtgtg gagactggaa gccagggaagc atctgcaatc tacttcaacta gtgggaccag 780
tggtcttccc aagatggcag aacattcccta ctctgagctg ggcctcaagc ccaagatgga 840
tgcctggttg acaggctgc aagcctctga tataatgttg accatctcag acacaggttg 900
gatactgaac atcttgggct cactttttga atottgggaa ttaggagcat gcacatttgt 960
tcactctctg ccaaaagtgt acccaactgt tattctaaag acactctcca gttatccaah 1020
caagagtatg atgggtgccc ctattgttta ccggtgtttg ctacagcagg atctttccag 1080
ttacaagtcc ccccatctac agaactgcct cgttggaggg gagtcccttc ttccagaaac 1140
tcctggagaac tggagggtccc agacaggact ggacatccga gaattctatg gccagacaga 1200
aacgggatta acttgcatgg ttcccaagac aatgaaaato aaaccaggat acatgggaac 1260
ggctgcttcc tgttatgatg tacagggttat agatgataag ggcaacgtcc tgcctccggc 1320
cacagaagga gacattggca tcagggtcaa acccatcagg cctataggca tcttctctgg 1380
ctatgtggaa aatccccaca agacagcagc caacattoga ggagactttt ggcctcttgg 1440
agaccgggga atcaaaagat aagatgggta ttccagttt atgggacggg cagatgatat 1500
catttaactcc agcgggtacc ggattggacc ctccgaggta gagaatgcac tgatgaagca 1560
ccctgctgtg gttgagacgg ctgtgatcag cagccagcac cccgtccgag gagaggtggt 1620
gaaggcattt gtgatactgg cctgcagttt cctatcccat gaccagaac agctcaccaa 1680
ggagctgcag cagcatgtga agtcagtac agccccatcc aagtacccaa gaagataga 1740
gtttgtcttg aacctgcaca agactgtcac agggaaaatt caacgaacca aacttcgaga 1800
caaggagtgg aagatgtccg gaaaagcccg tgcgcagtga ggcgtctagg agacattcat 1860
ttggattccc ctctctcttc tctttctttt cctttggggt ccttggcctt actatgatga 1920
tatgagattc ttatgaaag aacatgaatg taagttttgt cttgccctgg ttattagcac 1980
aaaaacattc tatgttagat attgaataa ggaagaaaag aaagaggaga tgaagggggg 2046
agaaaaa 2046

```

<210> 42

<211> 857

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1375644CB1

<400> 42

```

ggccccgcgc gccacggcga ggcactgggc tcctctgtct ccccgtaggc cgtccccgc 60
gtggggccac tgccccgcgc ccccgcatg gtccggattt caaagcccaa gacgtttcag 120
gcctacttgg atgattgtca ccggaggtat agctgtgccc actgcgcgcg tcacttgccc 180
aacccagcag acctcatctc caagtcttcc cagggcagtc aggggcgtgc ctacctcttc 240
aactcagtgg tgaactgtgg ctgggggcca gccagggagc ggtgtctgct gaccggcttc 300
catgctgtgg ccgacatcca ctgcgagaac tgcaagacca ctttgggctg gaaatatgaa 360
caggcctttg agagcagcca gaagtacaaa gaggggaagt acatcattga actcaaccac 420
atgatcaaa acacggctgt ggaactgacc ccgtccccc gagcactgtg gctccagccc 480
ggcctggccg ccaggagcgc ccaactggctt cccgcccccc gaaggagctc ctggaccctc 540
agagccctgt cagaggacgg atctagctcc tgtatatata tttttattga tgcactgtga 600
ccttgggggg agaacagaag ggggacgacg cccccgcacc tccctgcgac tggctggctt 660
ggatctogtt tttaaccogt tcttgcccca cctgcacctat agttatggcc tgggtctctg 720
tcctctgggc ccagtgcacc gtctgtctgt tgaactcttc ccaccaggcc cttcttcccc 780
cacgcgtgtc tgtccccctc gttctgtagc gtttgtacat aataaaacaa tggagtgagg 840
acaaaaaaa aaaaaaa 857

```

<210> 43

<211> 1485

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1356261CB1

<400> 43

```

ggggcctggc cttcaggcca ctggctacgg aacccggggg ctcttcacca gtccagctcg 60

```

WO 02/064792

PCT/US02/00254

tttccagcac catgtcggtg cggacgctac cgtcgtcttt cttgaacttg gggggggaga 120
tgctttacac cctcgaccaa cggctcggtg cccagaaact cccggggagac aaggcccgca 180
aagatgag g gacagaggtg gacagaaaaa gatttctgaa tgacatcctc tccacctgt 240
tcaatgagaa gtttatggag gaattattca agcctcaaga gctctactcc aagaaggccc 300
tgagggtgtg ctatgagcgc ctggctcatg cctccattat gaaactgaac caggccaaga 360
tggataagct ctatgacctg atgacctatg ctttcaata tcaagtattg ctgtgtccc 420
gacccaagga tgtgtgtgtg gtccacttca atcacttgga taccatcaag ggattctacc 480
gagactcccc agccactctg cagcaagtgg acgagacttt cgggcagctg acagaaaata 540
atggtgtgtc ctctgcaggg gagtccagc tgaatccgca gacactctc atcttcttcc 600
aagacctgca catccgagta tccatgtttc taaaggacaa agttcagaat aataacggtc 660
gctttgtgtt gccgtgttcc gggcctgttc cttggggaac tgaagttcca ggactcatca 720
gaatgttcaa caacaaaggt gaagaagtga agaggataga attcaagcat ggtggaaact 780
atgtccctgc acccaagaa ggttcttttg aactttatgg agaccgagtc ctgaaactgg 840
gaactaacat gtacagcgtg aatcagcctg tggaaactca tgtgtctgga tcatcaaaaga 900
acttagcctc atggaccag gaaagcattg ctccaaaccc tctgtctaaa gaagagctga 960
atttcttgcc caggctgatg ggaaggatgg agattaagaa acccagtgcc cctgagccca 1020
gattccggtt gaattctctt accaccgatg aagaagagga accagcagcg ctaaccagcg 1080
cagaagagtt atcctatgaa gttatcaaca tacaagccac ccaggaccag caacggagcg 1140
aggagctggc tgaatcatg ggggagtttg agatcacgga gcagccaagg ctgagcacca 1200
gcaaaaggga cgatttctc gccatgatgg atgagttata gctgtcttga ccaggcgtcc 1260
ctgtccccc gggagaggct gctggatggt gacccctggg gaatgcccca tggccagaa 1320
tgatgtgtgt agttctctac tgagtgaagc cattaagctc atttctatt tatgttgtaa 1380
ggaatgtgtg gactctccct tgaggagcac tcaactctga aggcacacac atacacatat 1440
tttcagtga atatatctg acttttaaac ttgaaaaaaa aaaaa 1485

<210> 44

<211> 989

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1419379CB1

<400> 44

agaaaaatggc ggccccccaga ccgcactgcy ggcgtcgcg caggtgaagc ggtggtgga 60
gaggcgcctg aggttacttg ccgggcccgc ggtcctgag tctccagagc tgggtcttta 120
catcccccgy tccccaggtt tagcctgagc aggtttgttg aaggccggga ccaattgaca 180
gcacgatggt gacggagcag gagggtgatg ccatcgggca gacgtggtg gaccccagc 240
agccctctga gggccgcttc cgggcgctgt tcaagctgcy tgggtcggc ggcccaggcg 300
coattgcatg gatcagccag gccctgatg acgatccgc cctgtcaag caccagctgc 360
cctaactgct gggccagatg caggatgccc gcgccatccc catgtggtg gacgtgctgc 420
aagacacccc tcaggagccc atgggtgcgc atgaggcagg ggaaggcctg ggggccatcg 480
gggacccgga agttctggag atcctgaagc agtattcctc ggaccccgtc atcgaggttg 540
ccgagacctg ccagctggcc gtgcgcaggc tggagtggct gcagcagcag ggcggggagc 600
cgccggcggg accctacctc tccgtggacc ctgcccgcgc ggcgtaggag cgtgaagtgg 660
ggcgccctgcy ggaagcgctg ctggatgagt cccggccgct cctcgagcga taccgcgcca 720
tgttgcctc gcgcaacgcy ggaaggcagg agggccgctc ggcgtggcc gagggtctgc 780
actgtgggag cggcctcttc cgcacagagg tcggctacgt cctgggacag ctgcagcagc 840
aggcagagct ggggtctctg gaagaggctg acgttatggt ggcttcagct tcaataggaa 900
tgggacacag ggtctggggg cctctgactc cccaccccg aggcctgggt agggacaggg 960
tggtggctcc tgagtgggtc aggtagggtc 989

<210> 45

<211> 3886

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3509272CB1

WO 02/064792

PCT/US02/00254

```

<400> 45
gggggcagaa aagctttgtt aacctcttt tacagatgag gaaaaacaag atcagaggtg 60
ctaagtgtcg tagcctagtg coaggtcttc tgggcccact tctgggttct cccaagccc 120
atgtttcttc cctttctca caactttac ttcttctct gacctcacc accaccaaa 180
gtacttttaa ttctagaaaa gaaacccagg tgcacactgg cacacctgac ctctatgca 240
tcagaagctt tggatgattc cccatccaa atattagaga tgaatgaaa gcaagtagg 300
catctgcaa aagttgcttt ttccctctgt cattttagga cctcaagtaa gttttacca 360
gaaactgcta tcataccagg gattcattgt gtattlaaca acataggaat gcaatctggc 420
aaatttgaaa aactcttaac atacacccca aatccctgcc caaatttaag aactagagtg 480
gacacagtgc gtttttccat gtgcctctt ctgtgatggg gtaacgatac gtgggagcag 540
agaatgggga ggggtggagcy catgccagat gaggatctat cagcaatggg accggggcctc 600
cacttttagc tctccacctc gctctctca gaggaccgcc ttctattgca ttcagctgtg 660
atggtagcac gaacacaggt gcaccgagga cgaggagaga aggagccttg tctctctct 720
gcactgtagg caggacagca cagggtacgg agcagctctg agagaggcca gctcatcagg 780
gaagcacttg tcttcacctc tgggctttga ctgagcactg ggcgaattgg cctctgggat 840
caagcaaatc atcctaaaca gagttactct atgtcacact atggaaatgt ccaagtaggt 900
ggcctgtgtt tcaaaagatg tatttctccc ttttgttgtt gccatttcat aggttttaga 960
ttgggtgtgt gtttctcttc tctgaatggc actcgaatgt ttgctgactc ctactctgtg 1020
tgactggggt gtacagctat ggactgtgc atccatccc atcatcttc atgatacaag 1080
cagctctctc ttttttgaca gctgaagaag catcggtagg gaatccagaa ggagcgttca 1140
tgaaggtgtt acaagccggc aagcaactca caagcactga gctgattggt gagccagagg 1200
agcctctaga cagcagtggc atcaacttgt caggttttgg aggtgagcag ctgagacaca 1260
atgacagagag tgattttatc agtacactaa gttacatctt gcttatttcc tcaagcgtaa 1320
acctagatgt gaaatcactg ttaactccgt taattaaact gccaccaca ggaacacagc 1380
tggcaaaagt tcaaacgtta ggccaaaacc ggcagagagt gaagagagtc ctcatgggcc 1440
caagtagcat ccagaaaaag cacttcaaaag aggtaggaaag gcagagcacc aggggggaac 1500
aggttgccca ggcactctgt gagaacctgt ccgaagaaaa aaggctcggg agtccagccc 1560
caagggaggt ggaacagccc cacacacagc aggggcttga gaagttagng ggaacgccg 1620
tctacacaaa gcttctcttc acccaagagc ataaggcagc agtctctgtg ctgaaacct 1680
tctccaaagg cgcctctctc acctccagcc ctgcaaaagc cctaccacag gtgagagaca 1740
gatggaaagc cttaaccacac gctatttcca ttttagaang tgcaaaaggct agagttacaa 1800
atacgaagac gcttaaacca atcgtacatg ccagaaaaaa ataccgcttt cacaacactc 1860
gctcccagct gaacccacaga acacccaaag tcaaaaagag tccaaaggtc agaaagaaaa 1920
gttatctgag tagactgatg ctgcacaaac ggcttccatt ctctgcagcg aagagctcca 1980
taaatcccc ttcaaaaggg gctttttcat ccttaggaga cctgagtcct caagaaaccc 2040
ctttctggga agtatctgct ccttcagaa atttataga aaagaataat acaaacaca 2100
caactgcaag aatgtccttt gaagaaaatg attttatgga aaacactaac atgcccagaag 2160
gaacactctc tgaaaacaca aactacaalc atcctctgga ggcagattcc gctgggaactg 2220
cattcaactt agggccaact gttaaacaaa ctgagacaaa atgggaatac aacaacgttg 2280
gcaetgacct gtccccgag cccaaaagct tcaattaccc attgctctg tccccaggtg 2340
atcagtttga aattcagcta acccagcagc tacagtcctt tatcccaac aacaatgtga 2400
gaaggtctat tgctcatggt atccggacct tgaagatgga ctgctctggg gccatgtgc 2460
aagtgcactg tgcacagctc atctccagga caggccacct gatgaagctt ctcaagtggc 2520
agcaggaagt aaaggcatcc aagatagaat gggatacga ccaatggaag attgagaact 2580
acctaatga gagcacagaa gccacagatg aacagaagaa gaagtgcctt gagctcaaaa 2640
aagaagttcc aggatattggc tatactgaca aactcatctt ggcatttaatt gttactggaa 2700
tactaacgat ttgtattata cttttctgcc tcaattgtat atgttgtcac cgaaggtcat 2760
tacaagaaga tgaagaagga ttctcaaggg gcaatttccag atttctgcca cggaggggat 2820
gctctctgcg aagggaaggt caggatggac ttctctcatt tggacacggc ctctgggtta 2880
aagatattga caaacctctc agtgcacaaa gaataaataa tcaatgcatg aagctgcaca 2940
agaagtcac taatgaggac aagatactca acagggaacc tggggacagc gaagcccaa 3000
cggagagga ggaaggtgaa gccctgcoat aggagagaaa ccaagcccaac ctcaagcctc 3060
ctgcaaaaat acatagaata aacaacaaca gttactaaat gaatgaaaat tgtattccg 3120
atgaagcctg ccagagaaaa aaagcatttt ttaaaaagag aaataaggtg atatctgatt 3180
agggccaaaca tgatgcagac aagaaatgca ccggttcaga ggagggaagg tcaggccgcc 3240
tggggagagt ccatgaaaaa gatggaaagt gccagatgct gtacctggtg ctgggaaaaa 3300
gttgactagg ccagcatccc ttctccaaa gggggggctc ctgactcggg gggagggctg 3360
gaacatctga tacatctga ggagacagtg tgggacagca tgggtggcag ggaaccagcc 3420
gtggtctcgt tcttggtcgg ctggaagga gtagatgtaa gggatggttt agaagaagg 3480
aagtggaaag aaagttttct gagctgacaa gagggaagaa aggccgccta gaaggacac 3540
aaaaaggaag gagaagccct aagcagagtg agcaccagac tccacaggtt aagggtctag 3600
tcacacagga ccatccgcat gtccagcccc aggtgcaggg ccaagcatca cctatgcac 3660

```

WO 02/064792

PCT/US02/00254

```

tgaccaactg gctgtaaatt ggaggtcccc acaactccct cctcaggttt gaacatttgc 3720
tagaacagct catggaaccc aggaanaacag tttcttact agtgctgatt tattacaaag 3780
gatattttaa aggaacacaa tgatgaagcc agttgaagag atacacaggg tgagggttga 3840
aggctctgtg gagtggggtg accatctctg gacatggttg tcgcac 3886

```

```

<210> 46
<211> 2677
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Inocyte ID No: 708425CB1

```

```

<400> 46
ttgatttctt tcgtgttgtg tttttgaagg tggaaacctaa cttctgtgtg aacaagtgtt 60
gccagctcag aaggcagtgga ggagctgttt tcatctgtgt ctgttgagga tcaagatgat 120
tgctattccc tgtagatga tcaggacttc actctctttg atttatttcc tgaggggagt 180
gtctgcagtg atgtctcttc ttctatttagc acttactggg atgtgtcaga tagcgagttt 240
gaatggcagt taccaggcag tgacattggc agtggggagt atgtacttcc tgatgtcata 300
ccagttatcc caagtccacc ttgcctgctt cctaaaaaga aaaaacaaga cgggaattta 360
gatgaactcc ctggagtgcc aatgacaaat gatgagcagg tgggaatatat tgagtatctg 420
agtgggaag tgagtaactga gatgggtctt cggggagcaac ttgatactat taagatcatt 480
gatctctctg ctcaaatctc cctacagagc agtgagttta ttattgaact taactgtctc 540
acagatgaaa aactgaagca ggtcagaaac tatacaagg aacatagccc tcgccaacgg 600
ctctcaagag aggcctggaa gagaagcaac tttagtgttg caagcaccag tggagtggag 660
gggtgcagtg ccagcgccag cagcagcagt gccagcatgg tcagttctgc aagcagcagt 720
gggtccagtg ttggaacctc tgcttcaaac tccagtgcca acatgagtcg agcacacagt 780
gacagcaacc tgtctgaag tgccagcagag cggattcggg attcaaaaaa gcgatccaag 840
caggggaagt tacagcagaa ggccttcctg aagaggcagc tgaaggagca gaggcagccc 900
cggaaggaga ggcctcagtg gctctctctt aacgaagagg tgctgtctct gaaagtgaat 960
gaggaagacc atgaagcaga tgttgatgtt ttgatgtaat aaggggtgat ttatcaactg 1020
tcctttgtgag cattaataata ctccatctct atgggtttac atgcatcttg accaaaattt 1080
tgglttagggc tgtgtgatg taccattaat aatlttaagcc agtggttctt ggcaggtgat 1140
cccaagacct gcagcttcag catcaactga gaagtgttta ggaatgcata ctagtgggcc 1200
cngccccccg acatagttaa tcagaaacca acaggggagc gccatgcat gtttttttta 1260
caagtgtctg gttattctga tgcacagctc agtttaagaa coactacttt gggtaaacgt 1320
tttgactggt taaagtttat ggcggtgaag tgggctctt caaagactag tacttaaca 1380
gttttagaaga ttccaagta ctgctgacag tagtttata tgtcagcata catacatgtg 1440
tagagatcat aatttagttc cctcttaaat gttacaattt cttagtttac ttttctctaa 1500
gggcccatac ataattcttg atcctgggtg gaaattcttt ctgaggtgtg ggggtgggca 1560
aggtgtggat tgcgttttac gatagtgcct tcattagtgt tagttctgtc tgttttcaat 1620
cattattgac tcaaggttat tagaacagcc ccttatcttt ttctatttag atttattttt 1680
gttttttact ttatgtaagt tcagatctct tttttaagtg atgactactg atgaataaat 1740
gttaactagta gctgaatttt agacttgatg ctatgttgat taatatttaa atggtgaaag 1800
taattaggca aaataagcaa ttctgccttt ctgttgcttt cagaattttt aggcgtgtaa 1860
attgtctcag gcaggaagtc agtctatata tctctcaagt tcttattgta ggtcattttt 1920
ttttcagcac agctctaatt tttaaatatt tgggtataca gottctggca tgtgacctag 1980
ttaalggtga aaattaaact atgaaaaccc ttgtagatct gctaaacttc tctataaatt 2040
gctagaataa tacagagcaa tgagatatag aaaaacacct agcttgttgg ttattaatag 2100
agctctaat ttaactgtga aatcatggag tttctttaa tgttttctca gagtgtttgt 2160
tatcatgaac attagaataa gcattattcc ttttcttcta ccacatgtct tgaacattca 2220
catgggttaa agaagatgga aactgatact actgaacatc aagctgcttt ctgaatttca 2280
taattttctt tatattgttg taagcatttt gaaggtagcc gtgattctgt aggaagagct 2340
gtaatcaata tttttcaaac tctaagagta cagaggttag taagatttct ttttttttta 2400
ctttttgtta cgtcaagtag ctattataga aaaaagttca tgacttaaat tcttaattct 2460
cagttgatac aaaggcttat aatcatgtgt tttaaaattt tgatttoatt ttctgtctct 2520
gaaggtttaa gaatgaaac tggggataaa aattataata ctgcagcttt gccaaagctc 2580
gtgatctgtg tgaaccttt taaagactac taaagtata atttgaattt tttactaaaa 2640
taaacatatg gtgtctatgg ttttcaaaaa aaaaaaaa 2677

```

<210> 47

WO 02/064792

PCT/US02/00254

<211> 3031
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Inocyte ID No: 7469966CB1

<400> 47
 agaatacaaa agaagaacac totttatcga ggtggcaaa tatctcgttg gaactcttat 60
 ttttctctcg aaactatttc ttccagatttg ttaaaaggga cagttatgaa aagcttaaaag 120
 atttaataca ctgctgggca gagtctccta aaccaatatt tgcaaaaatc atcaatcttt 180
 atcatcatcc aggtctgtga ggtaccacac tggctatgca tgttctctgg gacttaaaaga 240
 aaaacttcag atgtgctgtg ttaaaaaaca agacaactga ttttgcaaaa attgcagagc 300
 aagtgatcaa tctgttcacc tatagggcaa agagccatca ggtattacatt cctgtgtctc 360
 tccitgttga tgattttgaa gaacaagaaa atgtctactt tctacaaaat gccatccatt 420
 ccgtttttag agaaaaggat ttgcgatatg aaaaacatt ggttaatttc ttaaatgcga 480
 tgagatcccg gaatccagat gaaagtgcga aattggcaga cagtattgca cttaattacc 540
 aactttcttc caaggaacaa agagcctttg gtgcacaaat gaaggaaatt gaaaagcagc 600
 ccaagaaactg tgaaaacttt tattccttca tgatcatgaa aagcaatttt gatgaacat 660
 atatagaaaa ttagtcagg aatatcctaa aaggacagga tgttgacagc aaggaaagc 720
 aactcatttc cttcttggct ttactcagct cttatgttac tgactctaca atttcagttt 780
 cccagtgtga aatatctttg ggaatcatat acactagtag accctgggaa cctgaagct 840
 tagaagaaca gatgggaact tattctacac ttctaataaa aacagaagtt gcagaatatg 900
 ggaatacac aggtgtgcgt atcattcacc ctctgattgc cctgtactgt ctaaaagaac 960
 tggaaagagc ctatcacttg gataaatgtc aaattgcaat gaatatatta gaagagaatt 1020
 tattctatga ttctggaata ggaagagaca aatttcaaca tgatgttcaa actctctctg 1080
 ttacaaagca gcgcaagggt tatggagatg aaacagacac tctgttttcc ccattaatgg 1140
 aagcttttaca gaataaagac attgaaaagg tcttgagtcg aggaagtaga cgattccac 1200
 aaaaatgattt catttgtcaa gccctagcaa gacatttcta cattaaagag aaggacttta 1260
 acacagctct ggactgggca cgtcaggcca aaatgaaagc accataaaat tctatattt 1320
 cagatacact aggtcaagtc tacaanaagtg aaatcaaat gtggttggtt ggaacaaaa 1380
 actgtaggag cattactgtt aatgacctaa cacatctcct agaagctgcg gaaaagacct 1440
 caagaccttt caaagaatcc caaaggcaaa ctgtagttaa aaactatgaa accgagaact 1500
 ggtcaccaca gaagtccag agacgatatg acatgtataa cacagcttgt ttctgggtg 1560
 aatagaagt tggctcttac actatccaga ttctcagct cactcccttt ttccacaaag 1620
 aaaaatgaatt atccaaaaaa catatggtgc aatttttacc aggaagtggt accatctctc 1680
 ctgattccag aaatgaatgt tatttggctc ttatgaagtt cacatccac ctaaaaatt 1740
 tacaatcaga tctgaagggt tgccttgact ttittattga ttatatggtt cttctgaaaa 1800
 tgggtatcac ccaaaaagaa attgcagaaa tcatgttaag caagaaagtc agtcgttgtt 1860
 tccggaataa ccaagaactt ttctgtcatt tggatccatg tctattacaa agtaaaagaga 1920
 gtcaattact ccaggaggag aattgcaaga aaagctaga agctctgaga gcagataggt 1980
 ttgtcggact ctgtgaatat cttaatccaa actacaaga tgctaccac atggaagta 2040
 tagtgaatga atatgccttc ctactgcagc aaactcaaa aaagcccatg acaaatgaga 2100
 aacaaaatcc cattttggcc aacattattc tgagtgtgtc aaagcccaa tccagttta 2160
 ttcaaccact taccacgcta aaaaaaacac tccgagaggt cttgcaattt gtaggactaa 2220
 gtcatcaata tccaggtcct tatttcttgg cctgcctcct gttctggcca gaaaatcaag 2280
 agctagatca agattccaaa ctaatagaaa agtatgttcc atccttaaat agatccttca 2340
 ggggacagta caagcgcatg tgcaggtcca agcaggcaag cacacttttc tatctgggca 2400
 aaaggaagggt tctaaacagt attgttcaca aggcacaaat agagcagtag ttgtataag 2460
 cacaataaac aaattccctc tggcacagtg gggatgtgtg gaaaaaaat gaagtcaaa 2520
 acctcctgcy tcytctaact ggtcaggctg aaggcaagct aatctctgta gaatatgga 2580
 cagaggaana aataaaata ccagtaatat ctgtttatct aggtccactc agaagtggca 2640
 ggaacataga aagatgtctt ttctacatag gattttccat tgaaggccct ctggcatatg 2700
 atataaagt aatttaagac aatcacatc ctgtagtcca aatatgttta ttatatctt 2760
 tatgatttta ttctctctct ctattctcat ggcacttcca taacattatg gctaacctct 2820
 aatcacagat ttgtcttttg cctccotgaa tgaattacaa gcctttttta gatatgaa 2880
 atgcctaccc ccagagcttg gcacaaagtg gagtcaatct tttatgttt taatatgca 2940
 ttctcagact caataatta agaagtttca ttgatatcca ctggtcacat cataactgtc 3000
 tatagggcaa taaaatctgt gttaaaactca a 3031

<210> 48

WO 02/064792

PCT/US02/00254

```
<211> 2415
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6920129CB1

<400> 48
cctggcacca ttttggctgg acagttttgt cttgcagggg gctgtcctgt gcattataga 60
atgttttaga gcattcgtgg cctctaccca ctatagtgcca gtagcacctc tcccttgagt 120
tgtgacaatc aaaaacatct ccattcattg ccaaatccca cccccccg cacagacaca 180
gttcccttgt tgagaccatc tggtttaaat aagtatgtgt ttcttaagat gaactggaac 240
tgcatctact tggaaatggt tggaaattct caagatattt tgcctcagtg tgatacagaa 300
tttagaattt ttttttaatc tctttctgtg ttgctatacg cagccttaaa acgttcttga 360
gttaattaga tgagccaaag agatgggtgtc tgggggtgc atgaagtggc tgggtcagcc 420
tcccttggtg ctgattggcg gctctctttg gcagcgtact gtaagaactc tggggacggc 480
ctctgttact gcttcgatga cagcgtatgt cagcagctgt cagaagatga ggtctgcagg 540
cagacagcat acatcctctt ctaccagagg cgacacgcca tcccgctatg gtacagcaac 600
agctcgttgg caggtctcac aagtctcttc ctgtgtgaac actgggtgag ccggctcccg 660
ggcagcaagc cagccagcgt gacctctgca gcttctcca gacgcacctc cctggcgtcg 720
ctctctgagt ccgtggagat gactggagaa aggagtgaa atgatggagg cttttcaact 780
cgaccatttg tgagaagtgt ccagcgtcag agtttgtcat ccagatcttc tgtccaccgc 840
cccttggcgc tcaatgaaaa ttgcattgaga ccttcatggt cctgtctgc taagctgcag 900
atgcgctcca attctccatc ccgattttca ggggattgc caattcacg ctctgtctcc 960
acccttgaga agattgggga ggcagcagat gacaaagtct ccatctcttg ctttggtagc 1020
ttgcgggaac ttctagcag ttaccaggaa ccaagcgaca gtcatagtc ccttgagcac 1080
aagcgtgtgg gcggggcccc tctgctgtc atggaaggcg tgttcaaga cgaatcgagc 1140
accgcagatc tgaactccag tgtctagat acacagagca aacattcagc acaaggggac 1200
cgctgcctcc cgtctctgtg tccatttgat aacaataatc agatcgcta tgtggatcag 1260
agcgaactcg tagacagctc tccagtcaaa gaggtyaaa cccccagcca cccaggctca 1320
ctgcacaaga aaccagagag cacaactaag agatcccca gttccaaagg caactctgag 1380
ccagagaaaa gcttgaggaa ggggagacca gcttggcaa gccaggagt atcccttcca 1440
agtacatccc ctctctctcc tcttctgtga aaagtctctc taaagccctc ccgtccccc 1500
agcacaagcag attctctctc caggggcagt ggcagcgcat cctccctgc cctgcccaca 1560
cccacaaagg agtccatccc gaaatctcag gaactcctgt cgtctccttc gccacagaag 1620
cagaagtcag cctcggccct cactacact gcttctcca catctgcaa aaaggcctcg 1680
ggccttgcca caaggagccc ttctccact gggaaagaca ggacttcaga ccacagcttg 1740
agttagagagg gctccagaca aagcttgggt tctgacagag ccagcgccc cccaccctc 1800
aaaccacaat cctcctgggt gagccaggcc cgagcagggg agggcagggg ggccgggaag 1860
cacgtgcgga gctctccat ggccagcctg cgtcccccga gcacacagat caagcttggt 1920
ttgaagaggg acagcaagtc tgaggacaag gggctgtctc tcttcaaatc agccttgaga 1980
cagaaggaaa ccggcgctc gacggatctt ggcaagacag ccttgcctc taaaaaggct 2040
ggtgggagct ctgttaagtc tgtctgtaag aacacgggg acgacgaggg agagagaggg 2100
caccagcttc cagcttccca gcagccaaat gcaaatacaa cgggaaaaa gcagcttgtc 2160
accaaggacc ctgcttctgc caaacattcc ctgctgtccg ctgcacaatc caagtcttcc 2220
caactagact ctggagttcc ctgctctccg ggtggcaggg agcttcgaga gaaatctcca 2280
aaaaagttaa cttctagcat gcaaacctct ctcaaaaac tcaagtatat 2340
ttctgcaatc gaagtgtttt atctgtaaag atgtttattt atttagaacc cctgcctccc 2400
caaaaaaaa aaaaaa 2415

<210> 49
<211> 1843
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 71514704CB1

<400> 49
gtgaacaca ttcattctca tcagaaagag cctcatcttt aattgcataa aacagaaagg 60
```



WO 02/064792

PCT/US02/00254

```

ttggtctcgca cttccctcgg ttggtgactc cgggcgcgtc gaagatcttg tacctgtcgt 120
tcgcagtggc tcactttccc aagccatggg cagcaggaag aaggaatltg cctgcaggt 180
caacatcagc acccaagagc ttggggagga aatgctcagt tccaaaggac taactgttgt 240
tgatgtctat caaggctggg gtggccctgt caaacctgtg gtgagcctct tccagaagat 300
gaggatcagag gtccggcctgg accttctgca ctttgcatta gcagaggcag atcgtcttga 360
tgtctcggaa aagtacagag ggaagtgcga gccaaccttt ctgttttatg caggaggaga 420
actgggtggct gtggttagag gagcaaatgc cccactgtct cagaaaacca tcttagacca 480
gctggaggcc gaaaagaaag tgcgtgctga agcagagaa cggaaagtga ttaaagatga 540
ggctctttct gatgaagatg aatgtgtttc ccataggaag aataatggtg aagatgagga 600
cahggtttca tcagagagga cctgtacctt ggccatcatt aaaccagatg cagtggccca 660
tggaagact gatgagatta tcatgaagat tcaggaaagt gggtttgaaa ttctaacaaa 720
tgaagagaga accatgacag aggcagaagt gcgacttttc taccacaca aagctggaga 780
ggaggcattt gagaagctgg tacatcacat gtgcagtga ccaagccacc tctgtatcct 840
caccaggact gagggcttgg aggcagtggt cactacctgg cgaaccgtca tgggcccccg 900
tgaccccaat gtggccagga gggagcagcc agaagtctc ctagctcagt accgacaga 960
aatgccttc aatgccttcc atggaagcgg ggacagagaa gatctgaca gagaactggc 1020
attgctcttc cccagtttga aattttcaga caaagatara gaagccctc agggcggtga 1080
ggctgaagca acagcggggc ccactgaggg gctttgcttt cctgaggatg tggattgaga 1140
tggtgtgtct gctcttcag agcacgtgac caggggctta ctgcacaaa cagacctccg 1200
gaatcgggac ttaacttttt gagtaccaat actttttggg ttagctgtct tttgtagaca 1260
atgaaaataa tatggctggg cgcagtgaat cacacctgta atcccagac tttgggaggc 1320
cagggcaggt ggaacacctg agataggag tttagacca gcttagccaa catggtgaaa 1380
cctgtctct actaaaata caaaaaatta gccggcggtg gtggtgtgga cctgtctcc 1440
agctactcgg gtggttagg catgagaatc gcttgaacc gggaggtgaa ggttgacgta 1500
agccagagatc atgccactgc actccagcct ggtgacaga gtgagactcc atettaata 1560
taataataac aaaaaaagt aacagaaaca acaaaacaa aaggggcgcg cccgcgaatt 1620
agttagacct tcgttagacc cgggaaatta atcccggaac cgtttccctt gcagggggct 1680
ctgcagaagt ttgcaatata aagctttatt ggaatcccg tcaaacctcg aaggggaggg 1740
ccgggttacc caaattcgcc cttatagtga gtgatattac gcagcgcata agtggccggc 1800
gttttacaag gtgtgcttgg aaaccccggg tttccaatta tgc 1843

```

<210> 50
 <211> 1827
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Inocyte ID No: 7715945CB1

```

<400> 50
gagtccggcgg cgggtggcga ggcgggtgagt gcgcggctcc ggggctggcc gactcccgta 60
gtggcccgcc cggcctgtgc tcgggggctc cgggctctgg gctctgggtg cggggacggg 120
gccaggtctc ttgaagacct cgcgacctgt gtcagcagag ccgcctctga ccacatgtg 180
catcatcttc tttaaagttt atccctgcgc tgtttccaaa aacgcgtaca ggtcatctt 240
ggcagccaac agggatgaat tctacagccg accctccaag ttagctgact tctgggggaa 300
caacaacagag atcctcagtg ggtgggacat ggagggaagg aagggaaggag gcacatggct 360
gggcatacag acacgtggca agctggcagc actcaacca taacctgcag cgcagctgga 420
ctggcaggcc cagggcgag gtgaacttgt caccacttt ctgacctgt acgtggacag 480
cttgtctac ctgaagaagg tctctatgga gggccatctg tacaatggct tcaacctcat 540
agcagccgac ctgagacag caaagggaga cgtcatttgc tactatggga accgagggga 600
gcttgatcct atcgttttga cgcagggcac ctacgggctg agcaacgcgc tgcgtggagac 660
tccctggagg aagctgtgct ttgggaagca gctctcctg gaggctgtgg aacggagcca 720
ggcgtgccc aaggatgtgc tcatgcccag cctcctggat gtgtcaca atgaaggagg 780
gcagctgcca gaccgggcca tcagagacca ggttggggag tacgtgcagc ccatgtgtg 840
caagtacgcg gctgtgtgcg tgcctgtccc tggctacgca accagaacca acactatcat 900
ctcgttagat gcggagggcc acgtgacctt aactgagcgt agcatgatgg acaaggacct 960
ctccaccctg gagaccagaa cctatgagtt cacactgcag agctaacccc acctctggcg 1020
ttggccagtg ggctcctggg gggccctgcc ttgaggggca ctgtggacag gaaaccttcc 1080
tttgccatac tgcattgcac tgcccgtggt ttggccagca tcccccgat caggggccctg 1140
tggttttgcgt gttacccatc tgtgtcccca tgcccagttc aggggtctgc tttatgccc 1200
tgaggagcag cagagctctg tactaggtct aggcacggcc gaggtatacc atgaacatgt 1260

```

WO 02/064792

PCT/US02/00254

```

gggtacacct gagccactc ttgcacatgt acacaggcac tcacatggca cacacataca 1320
ctcctgctgtg tgcacaagca cacacatgca aggcataatac atggacaccg acacaggcac 1380
atgtacatgc acaggtgtgc tacacatgtg cacacatgca cagttgcaca gacacacaca 1440
cacaggtgca cacacacgat gccgaacaag gcagaaggcg gactctcacc tctcatgtgc 1500
ttctggccag taggtctttg ttctggtcca acgacaggag taggcttgta tttaaaagcg 1560
gccccctctc tctgtgtgcc acagaacaca ggcgtgcttg gactcttgac aagcagacct 1620
gctcctgcag agggacacgc cacatttgga atggggcacc gagaagacct gagaaaaacc 1680
cactctctct tttttttttt ttttgagacg gagtcttgct ctgtcaccca ggttggagtg 1740
cagtggcagc actctggctc actgcaacct ccgctctcca agttcaagca attctctctg 1800
cccagctctc tgattagctg ggaattac 1827

```

```

<210> 51
<211> 794
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7025368CB1

```

```

<400> 51
gcccaggaga gotcgaccca cccaggcaca ccatagcccc agagatggct ggggacacag 60
aagtgtgaaa gcaaatgttt caggagttaa tgggggaggt gaagccatgg cacagggtga 120
cctggaacc agacaaggcg cttcttccca acgtctgaa gccaggctgg atgcaatacc 180
agcagtgagc cttgcgcagg ttccagtgct cctcctgctc tctgaactgg gctctgccc 240
aagtcttggt ccttttccac atgaactgga gtgaggagaa gtccaggggc caggtgaaga 300
tgagggtggt tacccagaga tgtaagaagt gcccccaacc tctgtttgag gacctgagt 360
tcacacaaga gaacattctca aggatcctga aaaaacctggt gtccgaatt ctgaagaagt 420
gtatagagag aagatttctag ttgatagagg aggttctctat gatcaaggac atctctcttg 480
aagggccaca caatagttag aactgtgagg catgtctgca gggctctctg gctggggcca 540
tacagggttac aagcctccgc ccatctcaga ccccaagagt acactccatt tacaagggtg 600
aggaggtagt taagccctgg gccctcaggag agaattgtct tctctacgca tgcacaaacc 660
acatctgtag gaacttaagc attttctgct gttgtgtcat tctcattggt atcgtgggtg 720
tgttgttaaa aactgctata tgagccttgg aaacatgacg ctaagtcatg aataaaatca 780
gattcaaaaa gtca 794

```

```

<210> 52
<211> 1539
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7500954CB1

```

```

<400> 52
tttcacataa ctagaagggt atagacagta tgacagtgga atacctcttg ttctcattta 60
ggacagtgtc tgaattgtag tatgagagac ttaagttaga taatttagtg catttcttag 120
tagtgagagt tgtaattgt taacacttgg aaggaattgg tatcataagg aattgacatc 180
ataagttttg gaaatctttt tctgctagga tgctttatag cctatctgaa ttgaaagtg 240
tagttagatc taactgtttt tttagccatat tattctgttg tattccattt aggtattgtg 300
aattattcag aatgcagacc ccaagatala gaggtacaaa ttatatctct tccaaactaa 360
tacacaaact atattatgtg cttttctttg taaaacagtt tgttaggggt gtgctatttg 420
aacatatttt ctgggtggata ttctagctt tgaacactgt acatttaact gtaagatttt 480
atttataaat tacacatttt aaggaaattt ctatactctc aaaaactgaa atattttttg 540
aaagaggtaa ttgctatata tctgctttcc ttctaaaata ttccaatat atctgggtgg 600
tactgtttta tgtatggaga aggttaaggca ttgagattaa cttctcataa cttacaaaat 660
ggcctgtgtg tgtgtgtgaa aataagaccg gtgtcctgac tccocagttg tttctgtgtg 720
atlagaccag actgcttact taaagtttta ttgcgtttaa aatctgttca taagtgttag 780
ctccattttt tctgtccat tattacaaca tagagaagca cattcatatc cccagagaaa 840
tttagtgata tgcagcaact ttgtttcact cttgacatta agtgtacatt gttacattt 900
tttatcttat gattgaatgt ttcaggtttt cctcctccac caggcgtctc acctccattt 960

```

WO 02/064792

PCT/US02/00254

```
cttataccaa caatagaaag tggacattcc tctgggttatg atagtcgttc tgcacgtgca 1020
tttccatgatg gcaatgcat gaagaacgat acagatacag ggaatatgca gaaagagggtt 1080
atgagcgtca cagagcaagt cagagaaaag aagaacgaca tagagaaaga cagacacagg 1140
agaagaggga aaccagacat aagtcctctc gaagtaatag tagacgtcgc catgaaagt 1200
aagaaggaga tagtcacagg agacacaaac acaaaaaatc taaaagaagc aaagaaggaa 1260
aagaaggagg cagtgcacct gccctgaac aggagagcac cgaagctaca cctgcagaa 1320
aggcatgggt ttggcctttt gtgtatatta gtaccagaag tagatactat aaatcttggt 1380
attttctggg ataatgttta agaaatttac cttaaatctt gttctgtttg ttagtatgaa 1440
aagttaactt tttttccaaa ataaaagagt gaatttttca tgttaagtta aaaatctttg 1500
tcttgacta tttcaaaaat aaaaagacag caatgactt 1539
```

【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
22 August 2002 (22.08.2002)

PCT

(10) International Publication Number
WO 02/064792 A3(51) International Patent Classification: C12N 15/12,
C07K 14/47, C12N 5/10, C07K 16/18, C12Q 1/68, G01N
33/50, 33/53, A61K 38/00, 39/00

(21) International Application Number: PCT/US02/00254

(22) International Filing Date: 4 January 2002 (04.01.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/260,168 5 January 2001 (05.01.2001) US
60/262,857 19 January 2001 (19.01.2001) US
60/262,736 19 January 2001 (19.01.2001) US(71) Applicant (for all designated States except US): INCYTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): BAUGHN, Mariah,
R. [US/US]; 14244 Santiago Road, San Leandro, CA
94577 (US). WARREN, Bridget, A. [US/US]; 10130 Par-
wood Drive #2, Cupertino, CA 95014 (US). HONCHELL,
Cynthia, D. [US/US]; 400 Laurel Street #203, San Carlos,
CA 94070 (US). XU, Yuning [US/US]; 1739 Walnut
Drive, Mountain View, CA 94040 (US). CHAWLA,
Narinder, K. [US/US]; 33 Union Square, #712, Union
City, CA 94587 (US). RAMKUMAR, Jayalaxmi
[IN/US]; 34350 Maybird Circle, Fremont, CA 94555
(US). YAO, Monique, G. [US/US]; 1189 Woodgate
Drive, Carmel, IN 46033 (US). LI, Yan [CN/US]; 3885
Corrina Way, Palo Alto, CA 94303 (US). YUE, Henry
[US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US).
THANGAVELU, Kavitha [IN/US]; 1950 Montecito
Avenue 23, Mountain View, CA 94043 (US). TANG,
Tom, Y. [CN/US]; 4230 Ranwick Court, San Jose, CA
95118 (US). DING, Li [CN/US]; 3353 Alma Street,
#146, Palo Alto, CA 94306 (US). BOROWSKY, Mark,
L. [US/US]; 122 Orchard Avenue, Redwood City, CA
94061 (US). HAEFALIA, April, J., A. [US/US]; 2227
Calle de Primavera, Santa Clara, CA 95054 (US). LI,
Dyung, Aina, M. [US/US]; 233 Coy Drive, San Jose, CA95123 (US). AZIMZAI, Valda [US/US]; 5518 Boulder
Canyon Drive, Castro Valley, CA 94552 (US). TRAN,
Bao [US/US]; 750 Selberg Avenue, Santa Clara, CA
95051 (US). NGUYEN, Dannel, B. [US/US]; 1403
Ridgewood Drive, San Jose, CA 95118 (US). BURFORD,
Neil [GB/US]; 105 Wildwood Circle, Danbury, CT 06422
(US). ISON, Graig, H. [US/US]; 1242 Weathersfield Way,
San Jose, CA 95118 (US). GURURAJAN, Rajagopal
[US/US]; 5591 Dent Avenue, San Jose, CA 95118 (US).
GANDHI, Ameena, R. [US/US]; 705 5th Avenue, San
Francisco, CA 94118 (US). ELLIOTT, Vicki, S. [US/US];
3770 Politen Place Way, San Jose, CA 95121 (US). TRAN,
Uyen, K. [US/US]; 2638 Mabury Square, San Jose, CA
95133 (US).(74) Agents: HAMLET-COX, Diana et al.; Incyte Pharma-
ceuticals, Inc., 3160 Porter Drive, Palo Alto, CA 94304
(US).(81) Designated States (national): AI, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CT, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GI,
GM, GR, GU, HD, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MY, NZ, NI, NL, NO, NZ, PT, RO, RU, SD, SE, SG, SI, SK,
SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA,
ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR,
GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent
(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR,
NE, SN, TD, TG).Published:
— with international search report(88) Date of publication of the international search report:
4 September 2003For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: MOLECULES FOR DISEASE DETECTION AND TREATMENT

(57) Abstract: The invention provides human molecules for disease detection and treatment (MDSST) and polynucleotides which
identify and encode MDDT. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The
invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of MDDT.

WO 02/064792 A3

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International Application No PCT/US 02/00254
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/12 C12N5/10 C12Q1/68 C07K14/47 C07K16/18 G01N33/50 A61K38/00 A61K39/00		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C12Q C07K G01N A61K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, EMBL, SEQUENCE SEARCH, BIOSIS, MEDLINE		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 01 57190 A (CAO YICHENG ;CHEN RUI HONG (US); GOODRICH RYLE (US); HYSEQ INC (US) 9 August 2001 (2001-08-09) SEQ ID NO: 3381 ---	1-56,82
L	DATABASE GENESEQ 6 November 2001 (2001-11-06) "Human polynucleotide SEQ ID NO 429" Database accession no. AAK51884 XP002220526 * Remark: database entry is cited to provide nucleotide sequence SEQ ID NO: 429 of WO 01 57190 * the whole document --- -/--	1-56,82
☒ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "Z" document member of the same patent family		
Date of the actual completion of the international search 14 November 2002		Date of mailing of the international search report 28. 02. 2003
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax. (+31-70) 340-3016		Authorized officer Bucka, A

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US 02/00254

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 00 55350 A (HUMAN GENOME SCIENCES INC ;ROSEN CRAIG A (US); RUBEN STEVEN M (US)) 21 September 2000 (2000-09-21) SEQ ID NO: 239 & 1081 ---	1-56,82
Y	RAMOS-MORALES FRANCISCO ET AL: "NA14 is a novel nuclear autoantigen with a coiled-coil domain." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, no. 3, 16 January 1998 (1998-01-16), pages 1634-1639, XP002220525 ISSN: 0021-9258 cited in the application the whole document ---	1-56,82
X	DATABASE EMBL 1 September 1999 (1999-09-01) "wq72f09.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2476841 3' similar to SW:UBP4 HUMAN Q13107 UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 4 ;, mRNA sequence" Database accession no. AI978591 XP002220527 the whole document ---	1-55
A	DATABASE EMBL30 April 2000 (2000-04-30) "Mus musculus brain cDNA, clone MNCb-2711, similar to Homo sapiens Sjogren's syndrome nuclear autoantigen 1 (SSNA1), mRNA" Database accession no. AB041656 XP002220528 the whole document ---	1-55
A	WO 00 73509 A (BANVILLE STEVEN C; GREENAWALT LILA B (US); INCYTE GENOMICS INC (US) 7 December 2000 (2000-12-07) the whole document -----	1-55

INTERNATIONAL SEARCH REPORT	international application No. PCT/US 02/00254
Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)	
This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:	
1. ☐ Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:	
2. ☐ Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:	
3. ☐ Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).	
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)	
This International Searching Authority found multiple inventions in this international application, as follows:	
see additional sheet	
1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.	
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.	
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:	
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: invention 1: claims 1-55, all partially; claims 56 and 82, both completely	
Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest. ☐ No protest accompanied the payment of additional search fees.	

International Application No. PCT/US 02/00254

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

Invention 1: claims 1-55,
all partially; claims 56 and 82, both completely

an isolated polypeptide, the isolated nucleic acid encoding that polypeptide, antibodies specifically binding to this protein, methods for identifying test compounds that bind to this protein, methods for detecting said nucleic acid, arrays containing said nucleic acid, methods for screening of modulators of the expression of said nucleic acid, methods for assessing toxicity of test compounds utilizing said nucleic acid, wherein the isolated polypeptide has the sequence of SEQ ID NO: 1, which is encoded by the polynucleotide sequence SEQ ID NO: 27

Inventions 2 to 26: claims 1-55,
all partially; claims 57-81, 83-107,
completely and in so far as applicable

an isolated polypeptide, the isolated nucleic acid encoding that polypeptide, antibodies specifically binding to this protein, methods for identifying test compounds that bind to this protein, methods for detecting said nucleic acid, arrays containing said nucleic acid, methods for screening of modulators of the expression of said nucleic acid, methods for assessing toxicity of test compounds utilizing said nucleic acid, wherein invention 2 is represented by SEQ ID NO: 2 and 28, invention 3 is represented by SEQ ID NO: 3 and 29, continuing to invention 26, which is represented by SEQ ID NO: 26 and 52, and wherein SEQ ID NO: 2 to 26 are amino acid sequences and SEQ ID NO: 28 to 52 are polynucleotide sequences, respectively

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/US 02/00254

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0157190	A	09-08-2001	US 2002128187 A1
			12-09-2002
			AU 2591801 A
			31-07-2001
			AU 3128801 A
			14-08-2001
			AU 3297101 A
			07-08-2001
			AU 3300301 A
			07-08-2001
			AU 3329301 A
			14-08-2001
			AU 3484701 A
			14-08-2001
			AU 3484801 A
			14-08-2001
			AU 3486501 A
			14-08-2001
			AU 3494401 A
			14-08-2001
			AU 3665801 A
			14-08-2001
			AU 3666001 A
			14-08-2001
			AU 3666301 A
			14-08-2001
			AU 3672101 A
			14-08-2001
			AU 4314201 A
			14-08-2001
			EP 1261735 A2
			04-12-2002
			EP 1254265 A1
			06-11-2002
			EP 1254266 A1
			06-11-2002
			EP 1254268 A1
			06-11-2002
			EP 1254269 A1
			06-11-2002
			EP 1254270 A1
			06-11-2002
			WO 0153326 A1
			26-07-2001
			WO 0155334 A2
			02-08-2001
			WO 0155335 A2
			02-08-2001
			WO 0157255 A1
			09-08-2001
			WO 0157260 A1
			09-08-2001
			WO 0157175 A2
			09-08-2001
			WO 0157261 A1
			09-08-2001
			WO 0157262 A1
			09-08-2001
			WO 0157187 A2
			09-08-2001
			WO 0157265 A1
			09-08-2001
			WO 0157188 A2
			09-08-2001
			WO 0157266 A1
			09-08-2001
			WO 0157267 A1
			09-08-2001
			WO 0157190 A2
			09-08-2001
			US 2002137044 A1
			26-09-2002
			US 2002127199 A1
			12-09-2002
WO 0055350	A	21-09-2000	AU 3395900 A
			04-10-2000
			AU 3617600 A
			04-10-2000
			AU 3617700 A
			04-10-2000
			AU 3619400 A
			04-10-2000
			AU 3619500 A
			04-10-2000
			AU 3869400 A
			04-10-2000
			EP 1168917 A2
			09-01-2002
			EP 1165588 A1
			02-01-2002
			EP 1169469 A1
			09-01-2002
			EP 1165589 A1
			02-01-2002
			EP 1159420 A1
			05-12-2001
			EP 1163358 A1
			19-12-2001
			WO 0055173 A1
			21-09-2000
			WO 0055350 A1
			21-09-2000
			WO 0055351 A1
			21-09-2000
			WO 0055180 A2
			21-09-2000
			WO 0055174 A1
			21-09-2000
			WO 0055320 A1
			21-09-2000
			US 2002081659 A1
			27-06-2002
			US 2002039764 A1
			04-04-2002

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/US 02/00254

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0055350	A	US 2002055627 A1	09-05-2002
		US 2002151681 A1	17-10-2002
		US 2002052308 A1	02-05-2002
		US 2002044941 A1	18-04-2002
WO 0073509	A	07-12-2000	AU 5320900 A
			18-12-2000
			EP 1268758 A2
			02-01-2003
			WO 0073509 A2
			07-12-2000
			AU 5319400 A
			28-12-2000
			EP 1181357 A2
			27-02-2002
			WO 0075298 A2
			14-12-2000
			AU 5316800 A
			05-03-2001
			EP 1200571 A1
			02-05-2002
			WO 0111032 A1
			15-02-2001

フロントページの続き

(51)Int.Cl.⁷

F I

テーマコード (参考)

A 6 1 P 1/18
 A 6 1 P 3/10
 A 6 1 P 5/02
 A 6 1 P 5/14
 A 6 1 P 5/40
 A 6 1 P 7/00
 A 6 1 P 7/06
 A 6 1 P 7/08
 A 6 1 P 9/10
 A 6 1 P 11/00
 A 6 1 P 11/06
 A 6 1 P 17/00
 A 6 1 P 17/02
 A 6 1 P 17/06
 A 6 1 P 19/02
 A 6 1 P 19/06
 A 6 1 P 19/10
 A 6 1 P 21/02
 A 6 1 P 21/04
 A 6 1 P 25/00
 A 6 1 P 25/02
 A 6 1 P 25/08
 A 6 1 P 25/14
 A 6 1 P 25/16
 A 6 1 P 25/18
 A 6 1 P 25/22
 A 6 1 P 25/28
 A 6 1 P 27/06
 A 6 1 P 27/12
 A 6 1 P 27/16
 A 6 1 P 29/00
 A 6 1 P 31/04
 A 6 1 P 31/10
 A 6 1 P 31/12
 A 6 1 P 31/18
 A 6 1 P 33/00
 A 6 1 P 33/02
 A 6 1 P 33/10
 A 6 1 P 35/00
 A 6 1 P 35/02
 A 6 1 P 37/00
 A 6 1 P 37/02
 A 6 1 P 37/08
 A 6 1 P 43/00
 C 0 7 K 1/22
 C 0 7 K 14/47
 C 0 7 K 16/18

A 6 1 P 3/10
 A 6 1 P 5/02
 A 6 1 P 5/14
 A 6 1 P 5/40
 A 6 1 P 7/00
 A 6 1 P 7/06
 A 6 1 P 7/08
 A 6 1 P 9/10
 A 6 1 P 11/00
 A 6 1 P 11/06
 A 6 1 P 17/00
 A 6 1 P 17/02
 A 6 1 P 17/06
 A 6 1 P 19/02
 A 6 1 P 19/06
 A 6 1 P 19/10
 A 6 1 P 21/02
 A 6 1 P 21/04
 A 6 1 P 25/00
 A 6 1 P 25/02
 A 6 1 P 25/08
 A 6 1 P 25/14
 A 6 1 P 25/16
 A 6 1 P 25/18
 A 6 1 P 25/22
 A 6 1 P 25/28
 A 6 1 P 27/06
 A 6 1 P 27/12
 A 6 1 P 27/16
 A 6 1 P 29/00
 A 6 1 P 31/04
 A 6 1 P 31/10
 A 6 1 P 31/12
 A 6 1 P 31/18
 A 6 1 P 33/00
 A 6 1 P 33/02
 A 6 1 P 33/10
 A 6 1 P 35/00
 A 6 1 P 35/02
 A 6 1 P 37/00
 A 6 1 P 37/02
 A 6 1 P 37/08
 A 6 1 P 43/00
 C 0 7 K 1/22
 C 0 7 K 14/47
 C 0 7 K 16/18

1 0 3

1 0 5

4 B 0 6 5
 4 C 0 8 4
 4 H 0 4 5

C 1 2 M	1/00	C 1 2 M	1/00	A
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 N	5/10	C 1 2 P	21/02	C
C 1 2 N	15/02	C 1 2 P	21/08	
C 1 2 P	21/02	C 1 2 Q	1/02	
C 1 2 P	21/08	C 1 2 Q	1/68	A
C 1 2 Q	1/02	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	D
G 0 1 N	33/50	C 1 2 N	15/00	F
G 0 1 N	33/53	C 1 2 N	5/00	A
		C 1 2 N	15/00	C
		A 6 1 K	37/02	

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT, BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN, TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EC,EE,ES,FI,GB,GD,GE, GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,R O,RU,SD,SE,SG,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

- (72)発明者 ワレン、ブリジット・エイ
アメリカ合衆国カリフォルニア州95014・クーペルティノー・#2・パーウッドドライブ 1
0130
- (72)発明者 ホンチェル、シンシア・ディー
アメリカ合衆国カリフォルニア州94070・サンカルロス・#203・ローレルストリート 4
00
- (72)発明者 スー、ユーミング
アメリカ合衆国カリフォルニア州94040・マウンテンビュー・ウォルナットドライブ 173
9
- (72)発明者 チョーラ、ナリンダー・ケイ
アメリカ合衆国カリフォルニア州94587・ユニオンシティ・#712・ユニオンスクエア 3
3
- (72)発明者 ランクマール、ジャヤラクシミ
アメリカ合衆国カリフォルニア州94555・フレモント・メイバードサークル 34359
- (72)発明者 ヤオ、モニーク・ジー
アメリカ合衆国インディアナ州46033・カーメル・ウッドゲートドライブ 1189
- (72)発明者 リュ、ヤン
アメリカ合衆国カリフォルニア州94303・パロアルト・コリーナウェイ 3885
- (72)発明者 ユエ、ヘンリー
アメリカ合衆国カリフォルニア州94087・サニーベイル・ルイスアベニュー 826
- (72)発明者 サンガベル、カピサ
アメリカ合衆国カリフォルニア州94043・マウンテンビュー・23・モンテシトアベニュー
1950
- (72)発明者 タング、トム・ワイ
アメリカ合衆国カリフォルニア州95118・サンノゼ・ランウィックコート 4230
- (72)発明者 ディング、リー
アメリカ合衆国カリフォルニア州94306・パロアルト・#146・アルマストリート 335
3

- (72)発明者 ボロースキー、マーク・エル
アメリカ合衆国カリフォルニア州9 4 0 6 1・レッドウッドシティー・オーチャードアベニュー
1 2 2
- (72)発明者 ハファリア、エープリル・ジェイ・エイ
アメリカ合衆国カリフォルニア州9 5 0 5 4・サンタクララ・コーレデブリマベール 2 2 2 7
- (72)発明者 リュ、デュング・アイナ・エム
アメリカ合衆国カリフォルニア州9 5 1 2 3・サンノゼ・コイドライブ 2 3 3
- (72)発明者 アジムザイ、ヤルダ
アメリカ合衆国カリフォルニア州9 4 5 5 2・カストロバレー・ボールダーキャニオンドライブ
5 5 1 8
- (72)発明者 トラン、バオ
アメリカ合衆国カリフォルニア州9 5 0 5 1・サンタクララ・サルバークアベニュー 7 5 0
- (72)発明者 ニュエン、ダニエル・ビー
アメリカ合衆国カリフォルニア州9 5 1 1 8・サンノゼ・リッジウッドドライブ 1 4 0 3
- (72)発明者 バーフォード、ニール
アメリカ合衆国コネチカット州0 6 4 2 2・ダラム・ワイルドウッドサークル 1 0 5
- (72)発明者 アイソン、クレイグ・エイチ
アメリカ合衆国カリフォルニア州9 5 1 1 8・サンノゼ・ウェザーズフィールドウェイ 1 2 4 2
- (72)発明者 ガルラジャン、ラジャゴパル
アメリカ合衆国カリフォルニア州9 5 1 1 8・サンノゼ・デントアベニュー 5 5 9 1
- (72)発明者 ガンディー、アミーナ・アール
アメリカ合衆国カリフォルニア州9 4 1 1 8・サンフランシスコ・フィフスアベニュー 7 0 5
- (72)発明者 エリオット、ビッキー・エス
アメリカ合衆国カリフォルニア州9 5 1 2 1・サンノゼ・ポルトンプレイスウェイ 3 7 7 0
- (72)発明者 トラン、ユエン・ケイ
アメリカ合衆国カリフォルニア州9 5 1 3 3・サンノゼ・メイブリースクエア 2 6 3 8

Fターム(参考) 2G045 AA34 AA35 BB01 BB20 BB46 CB01 DA13 DA36 FB02 FB03
FB05 GC15 JA01
4B024 AA01 AA11 AA19 BA53 BA61 BA80 CA03 CA04 CA07 CA09
CA12 CA20 DA02 DA03 DA06 EA02 EA04 FA02 HA03 HA11
HA13 HA14
4B029 AA07 AA21 AA23 BB20 CC01 CC02 CC03 CC08 FA12 FA15
4B063 QA01 QA05 QA19 QQ02 QQ08 QQ21 QQ41 QQ43 QQ53 QQ61
QQ79 QQ89 QQ91 QR08 QR32 QR35 QR40 QR42 QR56 QR62
QR77 QR80 QS16 QS25 QS33 QS34 QS36 QX01 QX02
4B064 AG01 AG27 CA02 CA10 CA19 CA20 CC01 CC24 CE12 DA01
DA13
4B065 AA26X AA58X AA72X AA90X AA91X AA93X AA93Y AB01 AB05 AC14
BA02 BA08 CA24 CA25 CA43 CA44 CA46
4C084 AA02 AA07 AA16 BA22 BA23 NA14 ZA01 ZA02 ZA06 ZA12
ZA16 ZA18 ZA20 ZA22 ZA24 ZA33 ZA36 ZA40 ZA55 ZA59
ZA66 ZA68 ZA75 ZA81 ZA89 ZA94 ZA96 ZA97 ZB05 ZB11
ZB13 ZB15 ZB21 ZB26 ZB27 ZB33 ZB35 ZB37 ZB38 ZB39
ZC06 ZC08 ZC35 ZC55
4H045 AA10 AA11 AA20 AA30 BA10 BA41 CA40 DA75 DA76 EA20
EA50 FA71 FA72 FA74 GA26

专利名称(译)	疾病检测和治疗分子		
公开(公告)号	JP2005500818A	公开(公告)日	2005-01-13
申请号	JP2002565105	申请日	2002-01-04
[标]申请(专利权)人(译)	洞察Genomics公司		
申请(专利权)人(译)	洞察基因组公司		
[标]发明人	ボーグンマライアアール ワレンブリジットエイ ホンチェルシンシアディー スーユーミング チョーラナリンダーケイ ランクマールジャヤラクシミ ヤオモニークジー リュヤン ユエヘンリー サンガベルカピサ タングトムワイ ディングリー ボロースキーマークエル ハファリアエープリルジェイエイ リュデュングアイナエム アジムザイヤルダ トランバオ ニュエンダニエルビー バーフォードニール アイソクレイグエイチ ガルラジャンラジャゴバル ガンディーアミーナアール エリオットビッキーエス トランユエンケイ		
发明人	ボーグン、マライア・アール ワレン、ブリジット・エイ ホンチェル、シンシア・ディー スー、ユーミング チョーラ、ナリンダー・ケイ ランクマール、ジャヤラクシミ ヤオ、モニーク・ジー リュ、ヤン ユエ、ヘンリー サンガベル、カピサ タング、トム・ワイ ディング、リー ボロースキー、マーク・エル ハファリア、エープリル・ジェイ・エイ リュ、デュング・アイナ・エム アジムザイ、ヤルダ トラン、バオ ニュエン、ダニエル・ビー バーフォード、ニール アイソン、クレイグ・エイチ		

ガルラジャン、ラジャゴバル
ガンディー、アミーナ・アール
エリオット、ビッキー・エス
トラン、ユエン・ケイ

IPC分类号	G01N33/50 A61K38/00 A61K45/00 A61P1/04 A61P1/16 A61P1/18 A61P3/10 A61P5/02 A61P5/14 A61P5/40 A61P7/00 A61P7/06 A61P7/08 A61P9/10 A61P11/00 A61P11/06 A61P17/00 A61P17/02 A61P17/06 A61P19/02 A61P19/06 A61P19/10 A61P21/02 A61P21/04 A61P25/00 A61P25/02 A61P25/08 A61P25/14 A61P25/16 A61P25/18 A61P25/22 A61P25/28 A61P27/06 A61P27/12 A61P27/16 A61P29/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P33/02 A61P33/10 A61P35/00 A61P35/02 A61P37/00 A61P37/02 A61P37/08 A61P43/00 C07K1/22 C07K14/47 C07K16/18 C12M1/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N15/02 C12N15/09 C12N15/12 C12P21/02 C12P21/08 C12Q1/02 C12Q1/68 G01N33/15 G01N33/53
CPC分类号	A61P1/04 A61P1/16 A61P1/18 A61P3/10 A61P5/02 A61P5/14 A61P5/40 A61P7/00 A61P7/06 A61P7/08 A61P9/10 A61P11/00 A61P11/06 A61P17/00 A61P17/02 A61P17/06 A61P19/02 A61P19/06 A61P19/10 A61P21/02 A61P21/04 A61P25/00 A61P25/02 A61P25/08 A61P25/14 A61P25/16 A61P25/18 A61P25/22 A61P25/28 A61P27/06 A61P27/12 A61P27/16 A61P29/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P33/02 A61P33/10 A61P35/00 A61P35/02 A61P37/00 A61P37/02 A61P37/08 A61P43/00 C07K14/47
FI分类号	C12N15/00.ZNA.A A61K45/00 A61P1/04 A61P1/16 A61P1/18 A61P3/10 A61P5/02 A61P5/14 A61P5/40 A61P7/00 A61P7/06 A61P7/08 A61P9/10 A61P11/00 A61P11/06 A61P17/00 A61P17/02 A61P17/06 A61P19/02 A61P19/06 A61P19/10 A61P21/02 A61P21/04 A61P25/00 A61P25/02 A61P25/02.103 A61P25/08 A61P25/14 A61P25/16 A61P25/18 A61P25/22 A61P25/28 A61P27/06 A61P27/12 A61P27/16 A61P29/00 A61P31/04 A61P31/10 A61P31/12 A61P31/18 A61P33/00 A61P33/02 A61P33/10 A61P35/00 A61P35/02 A61P37/00 A61P37/02 A61P37/08 A61P43/00.105 C07K1/22 C07K14/47 C07K16/18 C12M1/00.A C12N1/15 C12N1/19 C12N1/21 C12P21/02.C C12P21/08 C12Q1/02 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D C12N15/00.F C12N5/00.A C12N15/00.C A61K37/02
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/BB01 2G045/BB20 2G045/BB46 2G045/CB01 2G045/DA13 2G045/DA36 2G045/FB02 2G045/FB03 2G045/FB05 2G045/GC15 2G045/JA01 4B024/AA01 4B024/AA11 4B024/AA19 4B024/BA53 4B024/BA61 4B024/BA80 4B024/CA03 4B024/CA04 4B024/CA07 4B024/CA09 4B024/CA12 4B024/CA20 4B024/DA02 4B024/DA03 4B024/DA06 4B024/EA02 4B024/EA04 4B024/FA02 4B024/HA03 4B024/HA11 4B024/HA13 4B024/HA14 4B029/AA07 4B029/AA21 4B029/AA23 4B029/BB20 4B029/CC01 4B029/CC02 4B029/CC03 4B029/CC08 4B029/FA12 4B029/FA15 4B063/QA01 4B063/QA05 4B063/QA19 4B063/QQ02 4B063/QQ08 4B063/QQ21 4B063/QQ41 4B063/QQ43 4B063/QQ53 4B063/QQ61 4B063/QQ79 4B063/QQ89 4B063/QQ91 4B063/QR08 4B063/QR32 4B063/QR35 4B063/QR40 4B063/QR42 4B063/QR56 4B063/QR62 4B063/QR77 4B063/QR80 4B063/QS16 4B063/QS25 4B063/QS33 4B063/QS34 4B063/QS36 4B063/QX01 4B063/QX02 4B064/AG01 4B064/AG27 4B064/CA02 4B064/CA10 4B064/CA19 4B064/CA20 4B064/CC01 4B064/CC24 4B064/CE12 4B064/DA01 4B064/DA13 4B065/AA26X 4B065/AA58X 4B065/AA72X 4B065/AA90X 4B065/AA91X 4B065/AA93X 4B065/AA93Y 4B065/AB01 4B065/AB05 4B065/AC14 4B065/BA02 4B065/BA08 4B065/CA24 4B065/CA25 4B065/CA43 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA07 4C084/AA16 4C084/BA22 4C084/BA23 4C084/NA14 4C084/ZA01 4C084/ZA02 4C084/ZA06 4C084/ZA12 4C084/ZA16 4C084/ZA18 4C084/ZA20 4C084/ZA22 4C084/ZA24 4C084/ZA33 4C084/ZA36 4C084/ZA40 4C084/ZA55 4C084/ZA59 4C084/ZA66 4C084/ZA68 4C084/ZA75 4C084/ZA81 4C084/ZA89 4C084/ZA94 4C084/ZA96 4C084/ZA97 4C084/ZB05 4C084/ZB11 4C084/ZB13 4C084/ZB15 4C084/ZB21 4C084/ZB26 4C084/ZB27 4C084/ZB33 4C084/ZB35 4C084/ZB37 4C084/ZB38 4C084/ZB39 4C084/ZC06 4C084/ZC08 4C084/ZC35 4C084/ZC55 4H045/AA10 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA10 4H045/BA41 4H045/CA40 4H045/DA75 4H045/DA76 4H045/EA20 4H045/EA50 4H045/FA71 4H045/FA72 4H045/FA74 4H045/GA26
優先権	60/260168 2001-01-05 US 60/262857 2001-01-19 US 60/262736 2001-01-19 US
外部リンク	Espacenet

摘要(译)

本发明提供了用于疾病检测和治疗的人类分子（MSST）和鉴定并编码MDDT的多核苷酸。 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MDDT异常表达有关的疾病的方法。

Inocyte プログラム ID	MSST ID	Inocyte プログラム ID	MSST ID	Inocyte プログラム ID	MSST ID
6242606	1	6242606	27	6242606	51
997105	2	997105	28	997105	52
1482843	3	1482843	29	1482843	53
3218219	4	3218219	30	3218219	54
7371678	5	7371678	31	7371678	55
7473863	6	7473863	32	7473863	56
7473862	7	7473862	33	7473862	57
7650474	8	7650474	34	7650474	58
8092178	9	8092178	35	8092178	59
1262178	10	1262178	36	1262178	60
3535417	11	3535417	37	3535417	61
7473555	12	7473555	38	7473555	62
7473985	13	7473985	39	7473985	63
7473117	14	7473117	40	7473117	64
5036986	15	5036986	41	5036986	65
1375644	16	1375644	42	1375644	66
1356761	17	1356761	43	1356761	67
1413179	18	1413179	44	1413179	68
3509272	19	3509272	45	3509272	69
708425	20	708425	46	708425	70
7469966	21	7469966	47	7469966	71
6920129	22	6920129	48	6920129	72
71514704	23	71514704	49	71514704	73
7715945	24	7715945	50	7715945	74
7025168	25	7025168	51	7025168	75
7500954	26	7500954	52	7500954	76