

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2004-518407

(P2004-518407A)

(43) 公表日 平成16年6月24日(2004.6.24)

(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 38/43	A 6 1 K 39/395	D 4 B O 2 4
A 6 1 K 38/44	A 6 1 K 39/395	N 4 B O 5 0
A 6 1 K 38/45	A 6 1 K 45/00	4 B O 6 3
A 6 1 K 38/46	A 6 1 K 49/00	A 4 B O 6 4
	審査請求 未請求 予備審査請求 有	(全 311 頁) 最終頁に続く

(21) 出願番号	特願2002-517758 (P2002-517758)	(71) 出願人	301005050 インサイト・ゲノミックス・インコーポレイテッド アメリカ合衆国カリフォルニア州94304・パロアルト・ポータードライブ 3160
(86) (22) 出願日	平成13年8月3日(2001.8.3)	(74) 代理人	100089266 弁理士 大島 陽一
(85) 翻訳文提出日	平成15年2月3日(2003.2.3)	(72) 発明者	ボーグン、マライア・アール アメリカ合衆国カリフォルニア州94577・サンレアンドロ・サンティアゴロード 14244
(86) 国際出願番号	PCT/US2001/024382		
(87) 国際公開番号	W02002/012467		
(87) 国際公開日	平成14年2月14日(2002.2.14)		
(31) 優先権主張番号	60/223,055		
(32) 優先日	平成12年8月4日(2000.8.4)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/224,728		
(32) 優先日	平成12年8月11日(2000.8.11)		
(33) 優先権主張国	米国 (US)		
(31) 優先権主張番号	60/226,440		
(32) 優先日	平成12年8月18日(2000.8.18)		
(33) 優先権主張国	米国 (US)		

最終頁に続く

(54) 【発明の名称】 薬物代謝酵素

(57) 【要約】

本発明はヒト薬物代謝酵素(DME)およびDME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、DMEの異常発現に関連する疾患を診断、治療または予防する方法をも提供する。

【特許請求の範囲】

【請求項 1】

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

(a) SEQ ID NO: 1 - 19 (配列番号 1 乃至 19) からなる群から選択したアミノ酸配列を含むポリペプチド

(b) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列と少なくとも 90% が同一であるような天然のアミノ酸配列を含むポリペプチド

(c) SEQ ID NO: 1 - 19 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 - 19 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 - 19 からなる群から選択した請求項 1 に記載の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 4】

請求項 2 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 5】

SEQ ID NO: 20 - 38 からなる群から選択した請求項 4 に記載の単離されたポリヌクレオチド。 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結合したプロモーター配列を含む組換えポリヌクレオチド。

【請求項 7】

請求項 6 に記載の組換えポリヌクレオチドを用いて形質転換した細胞。

【請求項 8】

請求項 6 に記載の組換えポリヌクレオチドを含む遺伝形質転換体。

【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、以下の過程を含む方法。 30

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

(b) そのように発現した前記ポリペプチドを回収する過程。

【請求項 10】

請求項 1 に記載のポリペプチドと特異的に結合する単離された抗体。

【請求項 11】

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

(a) SEQ ID NO: 20 - 38 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 40

(b) SEQ ID NO: 20 - 38 からなる群から選択したポリヌクレオチド配列と少なくとも 90% が同一であるような天然のポリヌクレオチドを含むポリヌクレオチド

(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド

(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド

(e) (a) ~ (d) の RNA 等価物

【請求項 12】

請求項 11 に記載のポリヌクレオチドの少なくとも 60 の連続したヌクレオチドを含む単離されたポリヌクレオチド。

【請求項 13】

請求項 11 に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中 50

から検出する方法であって、

(a) 前記サンプル中の前記標的ポリヌクレオチドに相補的な配列を含む少なくとも 2 0 の連続したヌクレオチドを含むプローブを用いて前記サンプルをハイブリダイズする過程 (ただし、前記プローブと前記標的ポリヌクレオチドまたは断片の間でハイブリダイゼーション複合体が形成されるような条件下で、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズする) と、

(b) 前記ハイブリダイゼーション複合体の存在・不存在を検出し、該複合体が存在する場合にはオプションでその量を検出する過程を含むことを特徴とする方法。

【請求項 1 4】

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 3 に記載の方法。 10

【請求項 1 5】

請求項 1 1 に記載のポリヌクレオチドの配列を有する標的ポリヌクレオチドをサンプル中から検出する方法であって、

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

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

【請求項 1 6】

請求項 1 のポリペプチドと、薬剤として許容できる賦形剤とを含む組成物。 20

【請求項 1 7】

前記ポリペプチドが、SEQ ID NO : 1 - 1 9 からなる群から選択したアミノ酸配列を有することを特徴とする請求項 1 6 に記載の成分。

【請求項 1 8】

機能的な D M E の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 1 6 の組成物を投与することを特徴とする治療方法。

【請求項 1 9】

請求項 1 に記載のポリペプチドのアゴニストとして有効性を確認するために化合物をスクリーニングする方法であって、 30

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝露する過程と、

(b) 前記サンプルにおいてアゴニスト活性を検出するステップとを含む。

【請求項 2 0】

請求項 1 9 に記載の方法によって同定したアゴニスト化合物と薬剤として許容できる賦形剤とを含む組成物。

【請求項 2 1】

機能的な D M E の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 2 0 の組成物を投与することを特徴とする治療方法。

【請求項 2 2】

請求項 1 に記載のポリペプチドのアンタゴニストとして有効性を確認するために化合物をスクリーニングする方法であって、 40

(a) 請求項 1 のポリペプチドを含むサンプルを化合物に曝露する過程と、

(b) 前記サンプルにおいてアンタゴニスト活性を検出するステップとを含む。

【請求項 2 3】

請求項 2 2 に記載の方法によって同定したアンタゴニスト化合物と、薬剤として許容できる賦形剤とを含む組成物。

【請求項 2 4】

機能的な D M E の過剰な発現に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 2 3 の組成物を投与することを特徴とする治療方法。

【請求項 2 5】

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

(a) 適切な条件下で請求項 1 に記載のポリペプチドを少なくとも 1 つの試験化合物に結合させるステップと、

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

【請求項 2 6】

請求項 1 に記載のポリペプチドの活性を調節する化合物をスクリーニングする方法であって、

(a) 請求項 1 に記載のポリペプチドの活性が許容された条件下で、請求項 1 に記載のポリペプチドを少なくとも 1 つの試験化合物に結合させる過程と、

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

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

【請求項 2 7】

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

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

(b) 前記標的ポリヌクレオチドの変異発現を検出する過程と、

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

【請求項 2 8】

試験化合物の毒性を算定する方法であって、

(a) 核酸を含む生物学的サンプルを前記試験化合物で処理する過程と、

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

(c) ハイブリダイゼーション複合体の収量を定量するステップと、

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

【請求項 2 9】

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

(a) 前記抗体が前記ポリペプチドに結合し、抗体とポリペプチドとの複合体が形成されるのに適した条件下で、前記生物学的サンプルを請求項 1 0 に記載の抗体と結合する過程と、

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

【請求項 3 0】

前記抗体が、

(a) キメラ抗体

10

20

30

40

50

- (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請求項 10 に記載の抗体。

【請求項 31】

請求項 10 に記載の抗体と、許容できる賦形剤とを含む化合物。

【請求項 32】

被検者の D M E の発現に関連する病状又は疾患の診断方法であって、請求項 31 に記載の化合物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。

【請求項 33】

前記抗体が標識されることを特徴とする請求項 31 に記載の化合物。

【請求項 34】

被検者の D M E の発現に関連する病状又は疾患の診断方法であって、請求項 33 に記載の化合物の有効量を前記被検者に投与する過程を含むことを特徴とする方法。

【請求項 35】

請求項 10 に記載の抗体の特異性を有するポリクローナル抗体を調製する方法であって、
(a) 抗体反応を誘発する条件下で、S E Q I D N O : 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチドまたはその免疫原性断片を用いて動物を免疫化する過程と、

(b) 前記動物から抗体を単離する過程と、

(c) 前記単離された抗体をポリペプチドでスクリーニングし、それによって、S E Q I D N O : 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するようなポリクローナル抗体を同定する過程とを含むことを特徴とする方法。

【請求項 36】

請求項 35 に記載の方法で産出した抗体。

【請求項 37】

請求項 36 に記載の抗体及び適切なキャリアを含む化合物。

【請求項 38】

請求項 10 に記載の抗体の特異性を有するモノクローナル抗体を製造する方法であって、
(a) 抗体反応を誘発する条件下で、S E Q I D N O : 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチドまたはその免疫原性断片を用いて動物を免疫化する過程と、

(b) 前記動物から抗体産出細胞を単離する過程と、

(c) 前記抗体産出細胞を不死化の細胞と融合して、モノクローナル抗体を産出するハイブリドーマ細胞を形成する過程と、

(d) 前記ハイブリドーマ細胞を培養する過程と、

(e) S E Q I D N O : 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチドに特異結合するような前記培養モノクローナル抗体から単離する過程とを含むことを特徴とする方法。

【請求項 39】

請求項 38 に記載の方法で産出したモノクローナル抗体。

【請求項 40】

請求項 39 に記載の抗体及び適切なキャリアを含む化合物。

【請求項 41】

F a b 発現ライブラリのスクリーニングにより前記抗体を産出することを特徴とする請求項 10 に記載の抗体。

【請求項 42】

組換え免疫グロブリンライブラリのスクリーニングにより前記抗体を産出することを特徴とする請求項 10 に記載の抗体。

【請求項 43】

10

20

30

40

50

SEQ ID NO: 1 - 30 からなる群から選択したアミノ酸配列を有するポリペプチドをサンプル中から検出する方法であって、

(a) 前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項10に記載の抗体をインキュベートする過程と、

(b) 特異結合を検出する過程とを含み、該特異結合が、SEQ ID NO: 1 - 19 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表示することを特徴とする方法。

【請求項44】

SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチドを精製する方法であって、

(a) 前記抗体と前記ポリペプチドの特異結合を許容する条件下で、サンプルを用いて請求項10に記載の抗体をインキュベートする過程と、

(b) 前記サンプルから前記抗体を分離し、SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有する精製ポリペプチドを得る過程とを含むことを特徴とする方法。

【請求項45】

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

【請求項46】

SEQ ID NO: 2 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項47】

SEQ ID NO: 3 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項48】

SEQ ID NO: 4 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項49】

SEQ ID NO: 5 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項50】

SEQ ID NO: 6 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項51】

SEQ ID NO: 7 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項52】

SEQ ID NO: 8 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項53】

SEQ ID NO: 9 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項54】

SEQ ID NO: 10 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項55】

SEQ ID NO: 11 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項56】

SEQ ID NO: 12 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項57】

SEQ ID NO: 13 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項58】

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

【請求項59】

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

【請求項60】

SEQ ID NO: 16 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項61】

SEQ ID NO: 17 のアミノ酸配列を含む請求項1に記載のポリペプチド。

【請求項62】

10

20

30

40

50

SEQ ID NO : 18 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 63】

SEQ ID NO : 19 のアミノ酸配列を含む請求項 1 に記載のポリペプチド。

【請求項 64】

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

【請求項 65】

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

【請求項 66】

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

【請求項 67】

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

【請求項 68】

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

【請求項 69】

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

【請求項 70】

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

【請求項 71】

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

【請求項 72】

SEQ ID NO : 28 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 73】

SEQ ID NO : 29 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 74】

SEQ ID NO : 30 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 75】

SEQ ID NO : 31 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 76】

SEQ ID NO : 32 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 77】

SEQ ID NO : 33 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 78】

SEQ ID NO : 34 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 79】

SEQ ID NO : 35 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオ

10

20

30

40

50

チド。

【請求項 80】

SEQ ID NO : 36 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 81】

SEQ ID NO : 37 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

【請求項 82】

SEQ ID NO : 38 のポリヌクレオチド配列を含む請求項 11 に記載のポリヌクレオチド。

10

【発明の詳細な説明】

【0001】

(技術分野)

本発明は、薬剤代謝酵素の核酸配列及びアミノ酸配列に関する。本発明はまた、これらの配列を自己免疫/炎症の疾患、細胞増殖異常、発生または発達障害、内分泌障害、眼の疾患、代謝障害、および肝臓の疾患を含む胃腸疾患の診断・治療・予防に利用することに関する。本発明はさらに、薬剤代謝酵素の核酸配列及びアミノ酸配列の発現における外来性化合物の効果についての評価に関する。

【0002】

(発明の背景)

薬剤の代謝および薬剤の体内での移動(薬物動態学)は、その効果、毒性、および他の薬剤との相互作用を決定する上で重要である。薬剤の吸収、様々な組織への薬剤の分布、および薬剤代謝産物の排泄の3つのプロセスが薬物動態を支配している。様々な代謝調節によって、溶解性、受容体との結合性、および排泄の速度を含む薬剤の物理化学的特性および薬理学的特性の大部分を変えるため、これらのプロセスは薬剤代謝に密接に関係する。薬剤を変える代謝経路はまた、ステロイド、脂肪酸、プロスタグランジン、ロイコトリエン、およびビタミン等の様々な天然の基質を受容する。従って、これらの経路における酵素は、天然の化合物、薬剤、発癌物質、変異誘発物質、および生体異物の間の生化学的および薬理学的相互作用の重要な部位となる。

20

【0003】

薬剤代謝における遺伝的な差異が、個人間において薬剤効果および毒性のレベルの著しい違いを引き起こすことが以前から知られている。治療指数が狭い薬剤、またはコデイン等の生理活性を必要とする薬剤の場合、このような遺伝子多型は極めて重要である。更に、有望な新規の薬剤が、一部の患者群のみに副作用を引き起こすという毒性から臨床試験で排除されることがよくある。薬剤代謝酵素が重要な部分を占める薬理ゲノミクス研究が進歩すれば、薬剤の効果および毒性の疑問に耐え得るツールが発展し、そのような情報が拡大されるであろう(Evans, W. E. および R. V. Reeling (1999) *Science* 286 : 487 - 491 を参照)。

30

【0004】

薬剤代謝反応は、薬剤分子を機能させて更なる代謝のために準備するフェーズ I、およびそれに引き続くフェーズ II に分類される。一般に、フェーズ I の反応生成物は部分的或いは完全に不活性であり、フェーズ II の反応生成物は主に排泄種である。しかしながら、フェーズ I 反応生成物は、投与された元の薬剤よりも活性が高い場合があり、この代謝活性の原理がプロドラッグとして利用される(例えば、L-ボドパ)。加えて、或る種の毒性化合物(例えば、アフラトキシン、ベンゾ-a-ピレン)は、これらの経路を経て毒性中間体に代謝される。フェーズ I 反応は通常、薬剤代謝における律速段階である。化合物或いは複数の化合物への事前の暴露によって、フェーズ I 酵素の発現させることができるが、それによってこの代謝経路を介する基質の流入が増大する(Klaassen, C. D., Amdur, M. O. および J. Doull (1996) *Casarett and Doull's Toxicology: The Basic Science*

40

50

n ce of Poisons, McGraw-Hill, New York, NY, 113-186頁; B. G. Katzung (1995) Basic and Clinical Pharmacology, Appleton および Lange, Norwalk, CT, 48-59頁; G. G. Gibson および P. Skett (1994) Introduction to Drug Metabolism, Blackie Academic and Professional, Londonを参照)。

【0005】

薬剤代謝酵素 (DME) は幅広い基質特性を有する。これは、多種多様な抗体がそれらの抗原に対して高い特異性を有する免疫系とは対照的である。多様な分子を代謝するDMEの能力によって、ある代謝レベルにおいて薬剤の相互作用の可能性が生まれる。例えば、或る化合物のDMEの誘導により、その酵素によって別の化合物が代謝され得る。 10

【0006】

DMEは、それらが触媒する反応の種類および関係する補助因子に従って分類することができる。フェーズI酵素の主なクラスには、限定するものではないがチトクロームP450およびフラビン含有モノオキシゲナーゼが含まれる。フェーズI型触媒サイクルおよび反応に関与するその他の酵素のクラスには、限定するものではないが、NADPHチトクロームP450還元酵素 (CPR)、ミクロソームチトクロームb5/NADHチトクロームb5レダクターゼ系、フェレドキシン/フェレドキシンレダクターゼレドックス対、アルド/ケト還元酵素、およびアルコールデヒドロゲナーゼが含まれる。フェーズII酵素の主なクラスには、限定するものではないが、UDPグルクロニルトランスフェラーゼ、スルホトランスフェラーゼ、グルタチオンSトランスフェラーゼ、Nアシルトランスフェラーゼ、およびNアセチルトランスフェラーゼが含まれる。 20

【0007】

チトクロームP450およびP450触媒サイクル関連酵素

酵素チトクロームP450のスーパーファミリーのメンバーは、様々な基質の酸化的代謝を触媒する。そのような基質には、ステロイド、脂肪酸、プロスタグランジン、ロイコトリエン、およびビタミン等の天然の化合物や、薬剤、発癌物質、変異誘発物質、および生体異物が含まれる。P450ヘム-チオレートタンパク質としても知られるチトクロームP450は通常、P450含有モノオキシゲナーゼ系と呼ばれる多成分電子伝達鎖における末端酸化酵素として作用する。触媒される特定の反応には、ヒドロキシル化、エポキシ化、N-酸化、スルホキシド化、N-脱アルキル、S-脱アルキル、およびO-脱アルキル、脱硫酸化、脱アミノ化、並びにアゾ、ニトロ、およびN-オキシド基の還元が含まれる。これらの反応は、動物における糖質コルチコイド、コルチゾール、エストロゲン、およびアンドロゲンのステロイド産生や、昆虫における殺虫剤耐性や、植物における除草剤耐性および花の発色や、微生物による環境浄化バイオレメディエーションに関係する。薬剤、発癌物質、変異誘発物質、および生体異物にチトクロームP450が作用して、物質の解毒或いはより毒性の強い生成物への変換を引き起こし得る。チトクロームP450は肝臓に豊富に存在するが、その他の組織にも存在し、その酵素はミクロソームに存在する (EXPASYZYME EC 1.14.14.1; Prosite PDO C00081 Cytochrome P450 cysteine heme-iron ligand signature; PRINTS EP450IE-Class P450 Group I signature; Graham-Lorence, S. および Peterson, J. A. (1996) FASEB J. 10: 206-214. を参照)。 40

【0008】

400種類のチトクロームP450が、細菌、真菌、植物、および動物を含む様々な生物において同定されている (Graham-Lorence, 前出)。Bクラスは原核生物および真菌に見られ、Eクラスは細菌、植物、昆虫、脊椎動物、および哺乳動物に見られる。5つのサブクラス即ちグループが、EクラスチトクロームP450の大きなファミリ 50

ーの中に含まれる (PRINTS EP 450 I E - C l a s s P 4 5 0 グループIシグネチャ)。

【0009】

全てのチトクロームP450はヘム補助因子を用いており、構造的特性を共有する。ほとんどのチトクロームP450は、400~530のアミノ酸の長さである。酵素の二次構造は、約70%のヘリックスと約22%のシートである。タンパク質のC末端部におけるヘム結合部位の周りの領域は、全てのチトクロームP450に保存されている。このヘム-鉄結合領域におけるアミノ酸10個のシグネチャ配列が同定され、この配列には第5の配位部位にあるヘム-鉄の結合に関与する保存されたシステインが含まれる。真核生物チトクロームP450では、通常は膜貫通領域がタンパク質の初めの15~20のアミノ酸において見られ、通常は約15の疎水性残基およびそれに続く正に帯電した残基からなる (Prositate PDOC00081前出 ; Graham-Lorence前出を参照)。

10

【0010】

チトクロームP450酵素は、細胞増殖および発達に関係する。この酵素は、化学物質を代謝してDNAと付加物を形成する反応性中間体にすることによって化学的な変異誘発および発癌において或る役割を果たす (Nebert, D. W. および Gonzalez, F. J. (1987) Ann. Rev. Biochem. 56 : 945-993)。これらの付加物が、発癌を引き起こすヌクレオチドの改変およびDNAの再編成を引き起こし得る。肝臓およびその他の組織におけるチトクロームP450の発現は、多環式芳香族炭化水素、ペルオキシソーム増殖因子、フェノバルビタール、および糖質コルチコイドデキサメタゾン等の生体異物によって引き起こされる (Dogra, S. C. 他 (1998) Clin. Exp. Pharmacol. Physiol. 25 : 1-9)。チトクロームP450タンパク質は、P450遺伝子CYP1B1における突然変異が原発性先天性緑内障を引き起こすように、目の発達に関与し得る (Online Mendelian Inheritance in Man (OMIM) *601771 Cytochrome P450, subfamily I (dioxin-inducible), polypeptide 1 ; CYP1B1)。

20

【0011】

チトクロームP450は炎症および感染に関係する。肝チトクロームP450活性は、様々な感染および炎症性の刺激によって著しく増減される (Morgan, E. T. (1997) Drug Metab. Rev. 29 : 1129-1188)。in vivoで観察される効果は、炎症誘発性のサイトカインおよびインターフェロンによって模倣され得る。2つのチトクロームP450タンパク質に対する自己抗体は、自己免疫性多腺性内分泌不全症I型 (APECED) 即ち多腺性自己免疫症候群の患者に見られた (OMIM *240300 Autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy)。

30

【0012】

チトクロームP450における突然変異は、副腎皮質過形成 (乳児期および幼児期の副腎機能不全において最も一般的)、偽ビタミンD-欠損くる病、脳腱黄色腫症 (進行性の神経障害、早期アテローム性動脈硬化症、および白内障によって特徴付けられる脂質貯蔵病)、抗凝血薬であるクマリンおよびワルファリンに対する遺伝性の耐性を含む代謝障害に関係する (Isselbacher, K. J. 他 (1994) Harrison's Principles of Internal Medicine, McGraw-Hill, Inc. New York, NY, 1968-1970頁 ; Takeyama, K. 他 (1997) Science 277 : 1827-1830 ; Kitanaka, S. 他 (1998) N. Engl. J. Med. 338 : 653-661 ; OMIM *213700 Cerebrotendinous xanthomatosis ; および OMIM *122700 Coumarin resistance)。チトクロームP450タンパク質アロマトラーゼの極端に高い発現レ

40

50

ベルが、重度の女性化乳房（女性化）を有する少年の繊維層肝細胞癌（fibrolamellar hepatocellular carcinoma）に見られた（Agarwal, V. R. (1998) J. Clin. Endocrinol. Metab. 83 : 1797 - 1800）。

【0013】

チトクロームP450触媒サイクルは、NADPHチトクロームP450レダクターゼ（CPR）によるチトクロームP450の還元によって完了する。チトクロームb5およびNADPHチトクロームb5レダクターゼからなる別のミクロソーム電子伝達系からのチトクロームP450触媒サイクルへの電子の寄与は、小規模であると一般的に見られている。しかしながら、Lamb, D. C.他（1999 ; FEBS Lett. 462 : 283 - 8）による近年の研究報告から、ミクロソームチトクロームb5/NADPHチトクロームb5レダクターゼ系によって効果的に還元され、支持され得る Candida albicans チトクロームP450（CYP51）が確認された。従って、この別の電子供与系によって支持される多くのチトクロームP450が存在すると思われる。

10

【0014】

チトクロームb5レダクターゼはまた、赤血球細胞における酸化型ヘモグロビン（酸素を保持することができないメトヘモグロビン）を活性型ヘモグロビン（ferrohemoglobin）へ還元する。酸化剤、または十分に還元されていない異常なヘモグロビン（ヘモグロビンM）が高いレベルで存在すると、メトヘモグロビン血症が引き起こされる。メトヘモグロビン血症はまた、赤血球チトクロームb5レダクターゼの先天的な欠損症からも起こり得る（Mansour, A. およびLurie, A. A. (1993) Am. J. Hematol. 42 : 7 - 12を参照）。

20

【0015】

チトクロームP450ファミリーのメンバーもまた、ビタミンDの合成および異化に密接に関係している。ビタミンDは、植物組織で生成されるエルゴカルシフェロール（ビタミンD2）および動物組織で生成されるコレカルシフェロール（ビタミンD3）の2つに生物学的に等価なプロホルモンとして存在する。後者のコレカルシフェロールは、7-デヒドロコレステロールが近紫外線（例えば、290~310 nm）に曝露されると形成される。通常は、皮膚が短時間日光に曝されると生成される（Miller, W. L. およびPortale, A. A. (2000) Trends in Endocrinol. Metab. 11 : 315 - 319を参照）。

30

【0016】

両方のホルモン型は更に、肝臓において酵素25-ヒドロキシラーゼによって25-ヒドロキシビタミンD（25(OH)D）に代謝される。25(OH)DはビタミンDの最も豊富に存在する前駆体であって、腎臓において酵素25-ヒドロキシビタミンD1-ヒドロキシラーゼ（1-ヒドロキシラーゼ）によって、活性型である1,25-ジヒドロキシビタミンD（1,25(OH)₂D）に更に代謝される。1,25(OH)₂D生成の調節は主に合成経路の最終ステップで行われる。1-ヒドロキシラーゼ活性は、酵素産物（1,25(OH)₂D）の循環レベル、並びに副甲状腺ホルモン（PTH）、カルシトニン、インスリン、カルシウム、リン、成長ホルモン、およびプロラクチンのレベルを含む幾つかの生理学的因子に左右される。更に、腎臓外の1-ヒドロキシラーゼ活性が報告され、組織特異的かつ局所的な1,25(OH)₂D生成の調節が生物学的に重要であると考えられる。1,25(OH)₂Dの24,25-ジヒドロキシビタミンD（24,25(OH)₂D）への触媒作用は、酵素25-ヒドロキシビタミンD24-ヒドロキシラーゼ（24-ヒドロキシラーゼ）を伴い腎臓でも起こる。24-ヒドロキシラーゼはまた、基質として25(OH)Dを利用することができる（Shinki, T.他（1997）Proc. Natl. Acad. Sci. U. S. A. 94 : 12920 - 12925 ; Miller, W. L. およびPortale, A. A.、前出、を参照）。

40

50

【0017】

ビタミンD25 - ヒドロキシラーゼ、1 - ヒドロキシラーゼ、および24 - ヒドロキシラーゼは全て、NADPH依存性I型(ミトコンドリア)チトクロームP450酵素であって、そのファミリーの他のメンバーと高い相同性を有する。ビタミンD25 - ヒドロキシラーゼはまた、幅広い基質特異性を有し、胆汁酸中間体の26 - ヒドロキシル化およびコレステロールの25 - ヒドロキシル化、26 - ヒドロキシル化、および27 - ヒドロキシル化を行う(Dilworth, F. J. 他(1995) J. Biol. Chem. 270 : 16766 - 16774 ; Miller, W. L. およびPortale, A. A. 前出、を参照)。

【0018】

ビタミンDの活性型(1, 25(OH)₂D)は、カルシウムおよびリン酸の恒常性に関与し、骨髄細胞および皮膚細胞の分化を促進する。ビタミンDの代謝に関与する酵素(例えば、1 - ヒドロキシラーゼ)の欠損によって生じるビタミンDの欠損は、低カルシウム血症、低リン酸血症、ビタミンD依存性(感受性)くる病、骨密度の低下並びによたよた歩行を伴う内反膝およびO脚という症状を示す疾患を引き起こす。ビタミンD25 - ヒドロキシラーゼの欠損は、脳腱黄色腫症や、アキレス腱、脳、肺、およびその他の多くの組織におけるコレステロールおよびコレスタノールの蓄積という特徴をもつ脂質貯蔵病を引き起こす。この疾患は、思春期後の小脳性運動失調症、アテローム性動脈硬化症、および白内障を含む進行性の神経障害を示す。ビタミンD25 - ヒドロキシラーゼの欠損によってくる病が起きないことから、25(OH)D合成の別の経路が存在することが推定される(Griffin, J. E. およびZerwekh, J. E. (1983) J. Clin. Invest. 72 : 1190 - 1199 ; Gambelin, G. T. 他(1985) J. Clin. Invest. 75 : 954 - 960 ; およびW. L. およびPortale, A. A. 前出)。

【0019】

フェレドキシンおよびフェレドキシンレダクターゼは電子伝達アクセサリタンパク質であって、少なくとも1つのヒトチトクロームP450種であって、CYP27遺伝子によってコードされるチトクロームP450c27を支持する(Dilworth, F. J. 他(1996) Biochem. J. 320 : 267 - 71)。ストレプトマイセス グリセウスチトクロームP450であるCYP104D1は、大腸菌において異種と共に発現され、内在性フェレドキシンおよびフェレドキシンレダクターゼ酵素によって還元されている(Taylor, M. 他(1999) Biochem. Biophys. Res. Commun. 263 : 838 - 42)。このことから、多くのチトクロームP450種が、フェレドキシン/フェレドキシンレダクターゼ対によって支持されていることが推定される。フェレドキシンレダクターゼはまた、モデル薬剤代謝系に見られ、抗腫瘍性抗生物質であるアクチノマイシンDを反応性フリーラジカル種に還元することが知られている(Flitter, W. D. およびMason, R. P. (1988) Arch. Biochem. Biophys. 267 : 632 - 9)。

【0020】

フラビン含有モノオキシゲナーゼ(FMO)

フラビン含有モノオキシゲナーゼは、求核性の窒素、硫黄、およびリンのヘテロ原子の異例な範囲の基質を酸化する。チトクロームP450と同様に、FMOはミクロソーム酵素であり、NADPHおよびO₂を用い、その基質がチトクロームP450の基質と広範に重なる。FMOは、肝臓、腎臓、および肺などの組織に分布している。

【0021】

組織特異的に発現される5つの異なった既知のFMOのアイソフォーム(FMO1、FMO2、FMO3、FMO4、およびFMO5)が哺乳動物に見られる。これらのアイソフォームは組織特異性が異なり、更に様々な化合物による阻害および反応の立体特異性等のその他の特性も異なる。FMOは、アミノ酸13個のシグネチャ配列を有し、その成分は配列のN末端の3分の2にまたがり、多くのNヒドロキシル化酵素に見られるFATGY

10

20

30

40

50

モチーフおよびFAD結合領域を含む(Ste hr, M.他(1998) Trends Biochem. Sci. 23 : 56 - 57 ; PRINTS FMOXYGEN ASE Flavin - containing monooxygenase signature)。

【0022】

特異的な反応には、求核性三級アミンのNオキシドへの酸化、二級アミンのヒドロキシルアミンおよびニトロンへの酸化、一級アミンのヒドロキシルアミンおよびオキシムへの酸化、および硫黄含有化合物およびホスフィンのS - オキシドおよびP - オキシドへの酸化が含まれる。ヒドラジン、ヨウ化物、セレン化物、および硼素含有化合物も基質である。FMOは化学的にチトクロームP450に類似しているが、FMOはその熱不安定性およびチトクロームP450の非イオン界面活性剤感受性に基づいて*in vitro*のチトクロームP450から通常は区別することができる。しかしながら、FMOのアイソフォームによって熱安定性および界面活性剤感受性が異なるため、これらの特性を区別に用いることは複雑である。

10

【0023】

FMOは、幾つかの薬剤および生体異物の代謝に重要な役割を果たしている。FMO(肺FMO3)は、尿中に排出される(S)ニコチンの(S)ニコチンN - 1' - オキシドへの代謝において主要な役割を果たす。FMOはまた、胃潰瘍の治療に広く用いられているH₂ - アンタゴニストであるシメチジンのS - 酸素化に関係する。FMOの肝臓発現型は、チトクロームP450と同じ調節制御下でない。例えば、ラットにおいて、フェノバルビタール治療によりチトクロームP450が生成されるが、FMO1は抑制される。

20

【0024】

FMOの内因性基質には、ジスルフィドに酸化されるシステアミン、シスタミンおよびトリメチルアミンN - オキシドに代謝されるトリメチルアミン(TMA)が含まれる。TMAは腐った魚のような臭いがし、FMO3アイソフォームの突然変異によって、悪臭がする遊離アミンが大量に汗、尿、および呼気から排出されるようになる。このような現象は、魚臭症候群引き起こす(OMIM 602079 Trimethylaminuria)。

【0025】

リシルオキシダーゼ

リシルオキシダーゼ(リシン6 - オキシダーゼLO)は、コラーゲンとエラスチンの架橋結合による結合組織マトリクスの形成に参与する銅依存性アミンオキシダーゼである。LOは、約50 kDaのNグリコシル化前駆体タンパク質として分泌され、この前駆体も活性ではあるが、メタロプロテアーゼによって酵素の成熟型に切断される。LOの銅原子は、酸素へ電子を伝達したりその反対に酸素から電子を取り除いたりすることに関係し、これらの細胞外マトリクスタンパク質におけるリシン残基の酸化的脱アミノ反応を促進する。銅の配位がLO活性に必須であるが、食事によって銅が摂取されなくてもアポ酵素の発現はその影響を受けない。しかしながら、機能的なLOの不在は、食事による銅の欠損に関係する骨格組織および血管組織の疾患に関係する。LOはまた、様々なセミカルバジド、ヒドラジン、および亜硝酸アミノ、並びにヘパリンによって阻害される。 - アミノプロピオニトリルは一般にインヒビターとして用いられる。LOの活性は、オゾン、カドミウム、および局所組織外傷に反応して放出されるホルモンのレベルの上昇に反応して増大する。このようなホルモンには、トランスフォーミング成長因子 - 、血小板由来成長因子、アンジオテンシンII、および線維芽成長因子等が含まれる。LO活性における異常は、メンケス症候群および後角症候群(occipital horn syndrome)に関係する。サイトゾル型の酵素は異常な細胞増殖に関係する(Rucker, R. B.他.(1998) Am. J. Clin. Nutr. 67 : 996S - 1002SおよびSmith - Mungo, L. I.およびKagan, H. M. (1998) Matrix Biol. 16 : 387 - 398を参照)。

30

40

【0026】

50

ジヒドロ葉酸レダクターゼ

ジヒドロ葉酸レダクターゼ (DHFR) は遍在性の酵素であって、ジヒドロ葉酸のテトラヒドロ葉酸への NADPH 依存性の還元を触媒するが、この反応はグリシンおよびプリンの新規な合成並びにデオキシウリジン-リン酸 (dUMP) のデオキシチミジン-リン酸 (dTMP) への変換における重要な過程である。この基本的な反応は、
 $7, 8\text{-ジヒドロ葉酸} + \text{NADPH} \rightarrow 5, 6, 7, 8\text{-テトラヒドロ葉酸} + \text{NADP}^+$ である。

【0027】

この酵素は、トリメトプリムおよびメトトレキセートを含む様々なジヒドロ葉酸類似体によって阻害され得る。豊富な TMP が DNA の合成に必要であるため、迅速な細胞の分裂には DHFR 活性が必要である。DNA ウイルス (例えば、ヘルペスウイルス) の複製はまた、高いレベルの DHFR 活性を必要とする。そのため、DHFR を標的とする薬剤が、DNA ウイルスの複製を抑制するために癌の化学療法に用いられている (同様の理由で、チミジル酸シンターゼが標的酵素である)。DHFR を抑制する薬剤は、急激に分裂する細胞 (または DNA ウイルス感染細胞) に対して細胞毒であるのが好ましいが、特異性を持たず、分裂する細胞を無差別に破壊する。更に、癌細胞は、獲得性の輸送障害または 1 つ以上の DHFR 遺伝子の複製の結果として、メトトレキセート等の薬剤に対して耐性を有するようになり得る (Stryer, L (1988) Biochemistry. W. H. Freeman および Co., Inc. New York. 511-5619 頁)。

【0028】

アルド/ケト還元酵素

アルド/ケト還元酵素は、単量体 NADPH 依存性オキシドレダクターゼであって幅広い基質特異性を有する (Bohren, K. M. 等 (1989) J. Biol. Chem. 264 : 9547-51)。これらの酵素は、カルボニル含有糖および芳香族化合物を含むカルボニル含有化合物の対応するアルコールへの還元を触媒する。従って、様々なカルボニル含有薬剤および生体異物はこのクラスの酵素によって代謝されると思われる。

【0029】

ファミリーメンバーであるアルドースレダクターゼによって触媒される既知の或る反応は、グルコースがソルビトールへ還元され、更にソルビトールデヒドロゲナーゼによってフルクトースに代謝される。通常の条件下で、グルコースのソルビトールへの還元は主な経路ではない。しかしながら、高血糖状態では、ソルビトールの蓄積は糖尿病合併症の発症に関係する (OMIM *103880 Aldo-ketoreductase family 1, member B1)。この酵素のファミリーのメンバーはまた、或る種の肝癌で高発現される (Cao, D. 他 (1998) J. Biol. Chem. 273 : 11429-35)。

【0030】

アルコールデヒドロゲナーゼ

アルコールデヒドロゲナーゼ (ADH) は、単純アルコールを対応するアルデヒドに酸化する。ADH はサイトゾル酵素であって、補助因子 NAD⁺ を好み亜鉛イオンを結合させる。ADH のレベルは肝臓において最も高く、腎臓、肺、および胃粘膜においては低い。

【0031】

既知の ADH アイソフォームは、40 kDa のサブユニットからなる二量体タンパク質である。これらのサブユニット (a, b, g, p, c) をコードする 5 つの既知の遺伝子座が存在し、その内の幾つかは特徴的な対立遺伝子変異体 (b1, b2, b3, g1, g2) を有する。サブユニットはホモ二量体およびヘテロ二量体を形成することができ、サブユニットの構成によって活性な酵素の特異的特性が決定される。従って、ホロ酵素がクラス I (サブユニット構成、aa, ab, ag, bg, gg)、クラス II (pp)、およびクラス III (cc) として分類される。クラス I ADH アイソ酵素はエタノールおよ

10

20

30

40

50

びその他の小さい脂肪族アルコールを酸化し、ピラゾールによって阻害される。クラス I I イソ酵素は長鎖の脂肪族アルコールおよび芳香族アルコールを好み、メタノールを酸化できない。また、ピラゾールによって阻害されない。クラス I I I イソ酵素は、更に長い長鎖脂肪族アルコール（5炭素およびそれ以上）および芳香族アルコールを好み、ピラゾールによって阻害されない。

【0032】

短鎖アルコールデヒドロゲナーゼには、様々な基質特異性を有する幾つかの関連酵素が含まれる。このグループに含まれるものは哺乳動物酵素である D - - ヒドロキシ酪酸デヒドロゲナーゼ、(R) - 3 - ヒドロキシ酪酸デヒドロゲナーゼ、15 - ヒドロキシプロスタグランジンデヒドロゲナーゼ、NADPH - 依存性カルボニルレダクターゼ、コルチコステロイド 11 - - デヒドロゲナーゼ、エストラジオール - 17 - - デヒドロゲナーゼ、細菌酵素であるアセトアセチル - CoAレダクターゼ、グルコース 1 - デヒドロゲナーゼ、3 - - ヒドロキシステロイドデヒドロゲナーゼ、20 - - ヒドロキシステロイドデヒドロゲナーゼ、リビトールデヒドロゲナーゼ、3 - オキソアシルレダクターゼ、2, 3 - dihydro - 2, 3 - dihydroxybenzoate dehydrogenase、ソルビトール - 6 - リン酸 2 - デヒドロゲナーゼ、7 - - ヒドロキシステロイドデヒドロゲナーゼ、cis - 1, 2 - dihydroxy - 3, 4 cyclohexadiene - 1 - carboxylate dehydrogenase、シス - トルエンジヒドロジオールデヒドロゲナーゼ、シス - ベンゼングリコールデヒドロゲナーゼ、ピフェニル - 2, 3 - ジヒドロ - 2, 3 - ジオールデヒドロゲナーゼ、N - acylmannosamine 1 dehydrogenase、および 2 - デオキシ - D - グルコン酸 3 - デヒドロゲナーゼがある (Krozowski, Z. (1994) J. Steroid Biochem. Mol. Biol. 51 : 125 - 130 ; Krozowski, Z. (1992) Mol. Cell Endocrinol. 84 : C25 - 31 ; および Marks, A. R. 他 (1992) J. Biol. Chem. 267 : 15459 - 15463)。

10

20

【0033】

UDPグルクロニルトランスフェラーゼ

UDPグルクロニルトランスフェラーゼファミリー (UGT) のメンバーは、補助因子ウリジン二リン酸 - グルクロン酸 (UDP - グルクロン酸) から基質へのグルクロン酸基の転移を触媒する。この転移は、通常は求核性ヘテロ原子 (O、N、または S) に対して行われる。基質には、フェーズ I 反応によって機能するようになる生体異物、並びにピリルビン、ステロイドホルモン、および甲状腺ホルモン等の内因性化合物が含まれる。グルクロン酸抱合の生成物は、基質の分子量が約 250 g/mol 未満であれば尿中に排泄されるが、グルクロン酸抱合された基質がそれより大きい場合は胆汁に排泄される。

30

【0034】

UGT は肝臓、腎臓、小腸、皮膚、脳、脾臓、および鼻粘膜のミクロソームに局在する。これらの局在化位置が、チトクローム P 450 酵素およびフラビン含有モノオキシゲナーゼと同じ小胞体膜の側であるため、フェーズ I 薬剤代謝の生成物への到達に理想的な位置である。UGT は、その UGT を小胞体膜に固着する C 末端膜貫通ドメイン、並びにその C 末端部分における約 50 のアミノ酸残基の保存されたシグネチャドメインを有する (Prositate PDOC00359 UDP - glycosyltransferase signature)。

40

【0035】

薬剤代謝に関係する UGT は、UGT 1 および UGT 2 の 2 つの遺伝子ファミリーによってコードされる。UGT 1 ファミリーのメンバーは、補助因子結合および膜挿入に関係する定常領域および可変基質結合ドメインを有する単一の遺伝子座の選択的スプライシングによって得られる。UGT 2 ファミリーのメンバーは異なる遺伝子座によってコードされ、UGT 2 A および UGT 2 B の 2 つのファミリーに分類される。2 A サブファミリーは嗅上皮で発現し、2 B サブファミリーは肝臓ミクロソームで発現する。UGT 遺伝子にお

50

ける突然変異は、高ビリルビン血症 (OMIM# 143500 Hyperbilirubinemia I)、出生児からの重度の高ビリルビン血症によって特徴付けられるクリグラー-ナジャー症候群 (OMIM# 218800 Crigler-Najjar syndrome)、ギルバール病と呼ばれる軽度の高ビリルビン血症 (OMIM * 191740 UGT1) に関係する。

【0036】

スルホトランスフェラーゼ

硫酸抱合は多くの同じ基質において起こり、O-グルクロン酸抱合により高水溶性の硫酸エステルが生成される。スルホトランスフェラーゼ (ST) は、補助因子 3'-ホスホアデノシン-5'-ホスホ硫酸 (PAPS) から基質へ SO_3 を転移してこの反応を触媒する。STの基質は主にフェノールおよび脂肪族アルコールであるが、抱合して対応するスルファミン酸を生成する芳香族アミンおよび脂肪族アミンも含まれる。これらの反応による生成物は主に尿中に排泄される。

10

【0037】

STは、肝臓、腎臓、腸管、肺、血小板、および脳を含む様々な組織に見られる。これらの酵素は、一般にサイトゾル酵素であって、多数の型が同時に発現される場合が多い。例えば、12種類を超えるSTの型がラットの肝サイトゾルに存在する。これらの生化学的に特徴付けられるSTは、それらの基質選択性に基づいて、アリアルスルホトランスフェラーゼ、アルコールスルホトランスフェラーゼ、エストロゲンスルホトランスフェラーゼ、チロシンエステルスルホトランスフェラーゼ、および胆汁酸塩スルホトランスフェラーゼの5つのクラスに分類される。

20

【0038】

ST酵素活性は、ラットの性別および年齢によって著しく異なる。発生の合図および性関連ホルモンとを組み合わせた効果が、ST発現プロファイルの差異並びにチトクロームP450等の他のDMEのプロファイルの差異を生じさせると考えられている。注目すべきは、ネコにおけるSTの高発現が、UDPグルクロナルトランスフェラーゼ活性のレベルの低下を部分的に補償する。

【0039】

STの幾つかの型は、ヒト肝サイトゾルから精製されクローニングされた。熱安定性および基質選択性が異なった2つのフェノールスルホトランスフェラーゼが存在する。熱安定酵素は、パラニトロフェノール、ミノキシジル、およびアセトアミノフェン等のフェノールの硫酸化を触媒し、熱不安定酵素は、ドーパミン、エピネフリン、およびlevodopa等のモノアミン基質を好む。その他のクローニングされたSTには、エストロゲンスルホトランスフェラーゼおよびN-アセチルグルコサミン-6-O-スルホトランスフェラーゼが含まれる。このN-アセチルグルコサミン-6-O-スルホトランスフェラーゼ酵素は、細胞生化学におけるSTのその他の重要な役割を示す。即ち、細胞分化およびプロテオグリカンの成熟に重要となり得る炭化水素基質の修飾である。実際に、スルホトランスフェラーゼの先天性の異常は、成熟したケラタン硫酸プロテオグリカンの合成不良という特徴をもつ障害である斑状角膜変性症に関係する (Nakazawa, K. 他 (1984) J. Biol. Chem. 259: 13751-7; OMIM * 217800 Macular dystrophy, corneal)。

30

40

【0040】

ガラクトシルトランスフェラーゼ

ガラクトシルトランスフェラーゼは、溶液に遊離している糖脂質または糖タンパク質の一部であるN末端アセチルグルコサミン (GlcNAc) オリゴ糖鎖にガラクトース (Gal) を転移するグリコシルトランスフェラーゼのサブセットである (Kolbinger, F. 他 (1998) J. Biol. Chem. 273: 433-440; Amado, M. 他 (1999) Biochim. Biophys. Acta 1473: 35-53)。ガラクトシルトランスフェラーゼは細胞表面で見出された可溶性細胞外タンパク質であり、ゴルジ体にも存在する。 1, 3 - ガラクトシルトランスフェ

50

ラーゼは、Gal (1 - 3) GlcNAc 結合を有する I 型炭化水素鎖を形成する。既知のヒトおよびマウス 1, 3 - ガラクトシルトランスフェラーゼは、短いサイトゾルドメイン、1つの膜貫通ドメイン、および8つの保存された領域を有する触媒ドメインを有すると考えられる (Kolbinger, F., 前出および Hennet, T. 他 (1998) J. Biol. Chem. 273 : 58 - 65)。マウス UDP - ガラクトースである - N - アセチルグルコサミン 1, 3 - ガラクトシルトランスフェラーゼ - I の領域 1 がアミノ酸残基の 78 - 83、領域 2 がアミノ酸残基の 93 - 102、領域 3 がアミノ酸残基 116 - 119、領域 4 がアミノ酸残基 147 - 158、領域 5 がアミノ酸残基の 172 - 183、領域 6 がアミノ酸残基の 203 - 206、領域 7 がアミノ酸残基の 236 - 246、領域 8 がアミノ酸残基の 264 - 275 に位置する。マウス UDP - ガラクトース内に見られる配列の変異体である - N - アセチルグルコサミン 1, 3 - ガラクトシルトランスフェラーゼ - I の領域 8 が、細菌ガラクトシルトランスフェラーゼにも見られることから、この配列がガラクトシルトランスフェラーゼ配列モチーフを決定すると考えられる (Hennet, T. 前出)。近年の研究により、brai ni ac タンパク質が 1, 3 - ガラクトシルトランスフェラーゼであると報告された (Yuan, Y. 他 (1997) Cell 88 : 9 - 11 ; および Hennet, T. 前出)。

10

【0041】

UDP - Gal である GlcNAc - 1, 4 - ガラクトシルトランスフェラーゼ (- 1, 4 - GalT) (Sato, T. 他 (1997) EMBO J. 16 : 1850 - 1857) が、Gal (1 - 4) GlcNAc 結合を有する II 型炭化水素鎖の形成を触媒する。1, 3 - ガラクトシルトランスフェラーゼの場合と同様に、可溶性の酵素が、膜結合型の切断によって形成される。1, 4 - ガラクトシルトランスフェラーゼに保存されているアミノ酸には、ジスルフィド結合によって連結された2つのシステイン残基および触媒ドメインにおける推定上の UDP - ガラクトース結合部位が含まれる (Yada v, S. および Brew, K. (1990) J. Biol. Chem. 265 : 14163 - 14169 ; Yada v, S. P. および Brew, K. (1991) J. Biol. Chem. 266 : 698 - 703 ; および Shaper, N. L. 他 (1997) J. Biol. Chem. 272 : 31389 - 31399)。

20

1, 4 - ガラクトシルトランスフェラーゼは糖タンパク質または糖脂質の炭化水素鎖を合成するほかにもいくつかの特定の役割を有する。哺乳動物において、 β - ラクトアルブミンを有するヘテロ二量体の一部として 1, 4 - ガラクトシルトランスフェラーゼが、乳汁分泌哺乳動物の乳汁の生成に作用する。精子の表面上の 1, 4 - ガラクトシルトランスフェラーゼは、卵を特異的に認識する受容体として働く。細胞表面 1, 4 - ガラクトシルトランスフェラーゼはまた、細胞接着、細胞 / 基板相互作用、正常な細胞遊走、および転移性腫瘍細胞遊走に作用する (Shur, B. (1993) Curr. Opin. Cell Biol. 5 : 854 - 863 ; および Shaper, J. (1995) Adv. Exp. Med. Biol. 376 : 95 - 104)。

30

【0042】

グルタチオンSトランスフェラーゼ

グルタチオンSトランスフェラーゼ (GST) によって触媒される基本的な反応は、還元型グルタチオン (GSH) と求電子体の抱合である。GST はホモ二量体またはヘテロ二量体タンパク質であって、主にサイトゾルに局在化するが、ある程度の活性がミクロソームにも見られる。主なイソ酵素は、共通の構造および触媒特性を有し、ヒトにおいてそれらは、主に α 、 μ 、 π 、および θ の4つのクラスに分類される。2つの最も大きなクラスである α および μ は、それぞれのタンパク質等電点 (α は pI が 7.5 - 9.0 まで、 μ は pI が 6.6 まで) によって同定される。それぞれの GST は、GSH に対する共通の結合部位および可変性疎水性結合部位を有する。それぞれのイソ酵素における疎水性結合部位は、特定の求電子性基質に対して特異的である。GST 内の特定のアミノ酸残基が、これらの結合部位および触媒活性に重要であるとして同定された。残基 Q67、T68、

40

50

D101、E104、およびR131は、GSHの結合に重要である(Lee, H-C 他(1995) J. Biol. Chem. 270:99-109)。残基R13、R20、およびR69はGSTの触媒活性に重要である(Stenberg G 他(1991) Biochem. J. 274:549-55)。

【0043】

殆どの場合、GSTは、潜在的な変異誘発性および発癌性の化学物質を不活性化し解毒するなど好都合に作用する。しかしながら、場合によってはそれらの作用が有害なもので、化学物質を活性化させて変異誘発性および発癌性になる。ラットおよびヒトGSTの或る型は、発癌の検出を助ける信頼性の高い新生物発生前マーカーとなる。変異誘発性を調べるための良く知られたエイムス試験に用いられるサルモネラチフィウムなどの細菌株におけるヒトGSTの発現は、変異誘発におけるこれらの酵素の役割を決定するのに役立つ。マウスにおいて肝腫瘍を引き起こす二ハロゲン化メタン(dihalomethanes)は、GSTによって活性化されると考えられている。この考えは、二ハロゲン化メタンが非感染細胞よりもヒトGSTを発現する細菌細胞において突然変異誘発性が高いということから支持される(Thier, R. 他(1993) Proc. Natl. Acad. Sci. USA 90:8567-80)。(1993) Proc. Natl. Acad. Sci. USA 90:856780)。二臭化エチレンおよび二塩化エチレンの変異誘発性が、ヒトGST, A1-1を発現する細菌細胞において上昇し、アフラトキシンB1の変異誘発がGSTの発現が促進されることで実質的に減少する(Simula, T. P. 他(1993) Carcinogenesis 14:1371-6)。6)。従って、GST活性の制御が、変異誘発および発癌の制御において有用であろう。

【0044】

GSTは、多剤耐性(MDR)として知られる多くの癌の薬剤治療に対する耐性獲得に係る。MDRは、シクロホスファミドなどの細胞毒で治療を受けた癌患者に起こり、その薬剤に対する耐性を有するようになった後、更にその他の様々な細胞毒に対しても耐性を有するようになる。GSTレベルの上昇は、これらの薬剤耐性癌に係り、薬剤に反応してGSTレベルが上昇した後、GSTによって触媒されたGSH抱合反応によりその薬剤が不活化されると考えられる。次に上昇したGSTレベルにより癌細胞がその他のGSTに結合する細胞毒から保護される。腫瘍におけるA1-1のレベルの上昇は、シクロホスファミド治療によって生じた薬剤耐性に係る(Dirven H. A. 他(1994) Cancer Res. 54:6215-20)。従って、癌組織におけるGST活性の調節は、癌患者のMDR治療に有用であろう。

【0045】

- グルタミルトランスぺプチダーゼ

- グルタミルトランスぺプチダーゼは広範に発現する酵素であって、- グルタミルアミド結合を切断して細胞外のグルタチオン(GSH)の分解を開始させる。GSHの分解は、生合成経路のために局所的なシステインプールを細胞に提供する。- グルタミルトランスぺプチダーゼはまた、細胞の酸化防止に寄与し、その発現は酸化ストレスによって引き起こされる。細胞表面局在糖タンパク質は、癌細胞において高いレベルで発現される。研究により、癌細胞の表面における- グルタミルトランスぺプチダーゼの高レベルの活性が、プロドラッグの活性化に利用され、抗癌治療薬が局所的に高いレベルで蓄積され得る(Hanigan, M. H. (1998) Chem. Biol. Interact. 111-112:333-42; Taniguchi, N. および Ikeda, Y. (1998) Adv. Enzymol. Relat. Areas Mol. Biol. 72:239-78; Chikhi, N. 他(1999) Comp. Biochem. Physiol. B. Biochem. Mol. Biol. 122:367-80)。

【0046】

アシルトランスフェラーゼ

10

20

30

40

50

N - アシルトランスフェラーゼ酵素は、アミノ酸抱合体の活性化カルボキシル基への転移を触媒する。内因性化合物および生体異物は、サイトゾル、ミクロソーム、およびミトコンドリアにおけるアシル - CoAシンセターゼによって活性化される。次に、アシル - CoA 中間体は、サイトゾルまたはミトコンドリアにおける N - アシルトランスフェラーゼによってアミノ酸（通常はグリシン、グルタミン、またはタウリンであるが、オルニチン、アルギニン、ヒスチジン、セリン、アスパラギン酸、およびいくつかのジペプチドを含む）と抱合し、アミド結合を有する代謝産物を形成する。この反応は、O - グルクロン酸抱合に相補的であるが、アミノ酸抱合では、グルクロン酸抱合によって生じる場合が多い反応性および毒性の代謝産物を生成されない。

【0047】

胆汁酸 - CoA であるアミノ酸 N - アシルトランスフェラーゼ (BAT) は、このクラスの良く特徴づけられた酵素の1つである。BAT は、腸管において界面活性剤として作用する胆汁酸抱合体の生成に関係する (Falany, C. N. 他 (1994) J. Biol. Chem. 269:19375-9; Johnson, M. R. 他 (1991) J. Biol. Chem. 266:10227-33)。BAT はまた、部分肝切除の後の肝臓癌患者の予後の予測マーカーとして有用である (Furutani, M. 他 (1996) Hepatology 24:1441-5)。

【0048】

アセチルトランスフェラーゼ

アセチルトランスフェラーゼは、ヒストンのアセチル化におけるその役割について非常に良く研究されている。ヒストンのアセチル化により、真核細胞におけるクロマチン構造が弛緩し、それによって転写因子が、ゲノムの影響を受けた領域（またはゲノム全体）における DNA の鑄型のプロモーターエレメントに到達できるようになる。これとは対照的に、ヒストンの脱アセチル化により、クロマチン構造が閉じ転写因子への到達が制限されて転写物が減少する。この最後に、ヒストンの脱アセチル化を阻害する化学薬品（例えば、酪酸ナトリウム）を細胞転写の刺激の一般的な手段として用いると、人為的結果であるが遺伝子発現が全体的に上昇する。アセチル化による遺伝子発現の調節はまた、限定するものではないが、p53、GATA-1、MyoD、ACTR、TFIIIE、TFIIF、および高移動グループ (HMG) タンパク質を含むその他のタンパク質のアセチル化によっても行うことができる。p53 の場合、アセチル化により DNA への結合が増大し、p53 によって調節される遺伝子の転写が刺激される。プロトタイプヒストンアセチラーゼ (HAT) は、サッカロミセス セレビジエ (酵母) に由来する Gcn5 である。Gcn5 は、テトラヒメナ p55、ヒト Gcn5、およびヒト p300/CBP を含むアセチラーゼのファミリーメンバーである。ヒストンのアセチル化については文献 (Cheung, W. L. 他 (2000) Curr. Opin. Cell Biology 12:326-333 および Berger, S. L (1999) Curr. Opin. Cell Biol 11:336-341) を参照されたい。ある種のアセチルトランスフェラーゼ酵素は、限定するものではないが、アセチルコリンエステラーゼおよびカルボキシルエステラーゼを含む幾つかの他の酵素の主なクラスに一般的である / ヒドロラーゼフォールド (/ hydrolase fold) (Center of Applied Molecular Engineering Inst. of Chemistry and Biochemistry University of Salzburg, <http://predict.sanger.ac.uk/irbm-course97/Docs/ms/>) を有する (Structural Classification of Proteins, <http://scop.mrc-lmb.cam.ac.uk/scop/index.html>)。

【0049】

N - アセチルトランスフェラーゼ

芳香族アミンおよびヒドラジン含有化合物は、肝臓およびその他の組織の N - アセチルトランスフェラーゼ酵素によって N アセチル化される。ある種の生体異物は、同じ酵素によ

10

20

30

40

50

ってある程度のアセチル化される。N - アセチルトランスフェラーゼは、補助因子アセチル - 補酵素 A (アセチル - CoA) を用いて 2 つの過程でアセチル基を転移するサイトゾル酵素である。第 1 の過程では、アセチル基がアセチル - CoA から活性部位システイン残基に転移され、第 2 の過程でアセチル基が基質アミノ基に転移され酵素が再生される。

【0050】

他のほとんどの DME クラスとは対照的に、既知の N - アセチルトランスフェラーゼの数が少ない。ヒトの場合、2 つの高度に類似した酵素である NAT 1 および NAT 2 が存在し、マウスには第 3 型酵素である NAT 3 が存在する。N - アセチルトランスフェラーゼのヒト型は、独立した調節 (NAT 1 は広範に発現されるが、NAT 2 は肝臓および臓器のみに発現される) および重複した基質特異性を有する。両方の酵素はある程度までほとんどの基質を受容するが、NAT 1 はある種の基質 (パラアミノ安息香酸、パラアミノサリチル酸、スルファメトキサゾール、およびスルファニルアミド) を好み、一方 NAT 2 は他の基質 (イソニアジド、ヒドララジン、プロカインアミド、ダブソン、アミノグルテチミド、およびスルファメチアジン) を好む。

10

【0051】

1950 年代に、抗結核薬であるイソニアジドを投与された患者の臨床試験から、化合物のアセチル化が速い人 (rapid acetylator) および遅い人 (slow acetylator) が報告された。これらの表現型は後に、酵素の活性または安定性に影響を与える NET 2 遺伝子における突然変異によることが示された。イソニアジドのアセチル化が遅い表現型が、中東の集団に優性 (約 70%) であり、白人ではやや劣り (約 50%)、アジアの集団では 25% 未満である。近年になって、NET 1 における機能的な多型が検出され、検査を受けた集団の約 8% が、アセチル化が遅い表現型を有する (Butcher, N. J. 他 (1998) Pharmacogenetics 8: 67 - 72)。NAT 1 がある種の既知の芳香族アミン発癌物質を活性化することができるため、広範に発現される NAT 1 酵素における多型が癌のリスクの決定に重要であると考えられる (OMIM *108345 N-acetyltransferase 1)。

20

【0052】

アミノトランスフェラーゼ

アミノトランスフェラーゼは、ピリドキサル 5' - リン酸 (PLP) 依存性酵素のファミリーを含み、アミノ酸のトランスフォーメーションを触媒する。アスパラギン酸アミノトランスフェラーゼ (AspAT) は最も研究された PLP 含有酵素である。AspAT は、ジカルボキシル - アミノ酸、アスパラギン酸およびグルタミン酸、および対応する 2 - オキソ酸、オキサロ酢酸、および 2 - オキソグルタル酸の可逆的なアミノ基転移反応を触媒する。このファミリーの他のメンバーには、推定アミノトランスフェラーゼ、分枝鎖アミノ酸アミノトランスフェラーゼ、チロシンアミノトランスフェラーゼ、芳香族アミノトランスフェラーゼ、アラニン - グリオキシル酸アミノトランスフェラーゼ (AGT) およびキヌレニンアミノトランスフェラーゼが含まれる (Vacca, R. A. 他 (1997) J. Biol. Chem. 272: 21932 - 21937)。

30

【0053】

原発性高シュウ酸尿症 I 型は、常染色体性劣性疾患であって、肝特異的ペルオキシソーム酵素であるアラニン - グリオキシル酸アミノトランスフェラーゼ - 1 の欠損が起こる。この疾患の表現型は、グリオキシル酸代謝の欠損である。AGT が存在しない場合、グリオキシル酸はグリシンに転移されるのではなく、シュウ酸に酸化される。その結果、腎臓および尿管に不溶性カルシウムシュウ酸が蓄積し、最終的に腎不全が起こる (Lumb, M. J. 他 (1999) J. Biol. Chem. 274: 20587 - 20596)。

40

【0054】

ヌレニンアミノトランスフェラーゼは、L - トリプトファン代謝産物 L - キヌレニンの不可逆的なアミノ転移を触媒してキヌレニン酸を形成する。この酵素はまた、L - 2 - アミノアジピン酸から 2 - オキソグルタル酸およびその逆への可逆的なアミノ基転移反応を触

50

媒して2-オキソアジピン酸およびL-グルタミン酸を生成し得る。キヌレイン酸は、グルタミン酸作動性神経伝達の推定上のモジュレーターであるため、キヌレインアミノトランスフェラーゼの欠損が多面発現性(pleiotropic)効果に関係すると思われる(Buchli, R. 他(1995) J. Biol. Chem. 270:29330-29335)。

【0055】

銅-亜鉛スーパーオキシドジスムターゼ

銅-亜鉛スーパーオキシドジスムターゼはコンパクトな二量体金属酵素であって、酸化的な傷害に対する細胞防御に関係する。この酵素は1つの亜鉛原子および1つの銅原子を各サブユニットに含み、スーパーオキシドアニオンの O_2 および H_2O_2 への不均化反応を触媒する。この不均化反応の速度は、拡散によって律速されるため、基質と酵素活性部位との間の好ましい静電的相互作用によって促進される。この酵素のクラスの例が、全ての真核細胞の細胞質において同定され、或る細菌種のペリプラズムでも同定された。銅-亜鉛スーパーオキシドジスムターゼは強壯な酵素であって、蛋白分解、および尿素およびSDSによる変性に高度な耐性を有する。この酵素のコンパクトな構造に加えて、金属イオンおよび内部サブユニットのジスルフィド結合が酵素の安定性に寄与していると考えられている。酵素は70 もの高温でも可逆的に変性される(Battistoni, A. 他(1998) J. Biol. Chem. 273:5655-5661)。

10

【0056】

スーパーオキシドジスムターゼの過剰な発現は、遺伝子組み換えアルファルファの耐凍性を高めるのに関係しており、ジフェニルエーテル除草剤であるアシフルオリフェン(acifluorfen)などの環境毒に対する耐性を与える(McKersie, B. D. 他(1993) Plant Physiol. 103:1155-1163)。加えて、酵母細胞が過酸化水素への暴露の後に凍結融解損傷に対する耐性が高まる。これは、曝露によるスーパーオキシドジスムターゼの発現の上方制御により、酵母細胞が更なる過酸化ストレスに適応するようになるためである。この研究により、酵母スーパーオキシドジスムターゼ遺伝子の突然変異は、冷凍保存過程を経て生物が存在するか否かを決定するのに重要であると長い間考えられていたグルタチオン代謝の調節に影響を及ぼす突然変異よりも、冷凍融解耐性に悪影響を与える(Jong-In Park, J-I. 他(1998) J. Biol. Chem. 273:22921-22928)。

20

30

【0057】

スーパーオキシドジスムターゼの発現はまた、結核を引き起こす生物である結核菌に関連する。スーパーオキシドジスムターゼは、結核菌によって排泄される10の主なタンパク質の内の1つであり、酸化ストレスに応じてその発現が約5倍上方制御される。結核菌は、非病原性ミコバクテリアM. smegmatisよりスーパーオキシドジスムターゼをほぼ100倍多く発現し、極めて高い割合で発現酵素を分泌する。この結果、結核菌はM. smegmatisよりも最大350倍多い酵素を分泌し、酸化ストレスに対する実質的な耐性を与える(Harth, G. およびHorwitz, M. A. (1999) J. Biol. Chem. 274:4281-4292)。

【0058】

銅-亜鉛スーパーオキシドジスムターゼの発現の低下、並びに抗酸化能力を有するその他の酵素の発現の低下は初期の癌に関係する。銅-亜鉛スーパーオキシドジスムターゼの発現レベルは、正常な前立腺組織と比べ前立腺の上皮新生物および前立腺癌において低い(Bostwick, D. G. (2000) Cancer 89:123-134)。

40

【0059】

ホスホジエステラーゼ

ホスホジエステラーゼは、ホスホジエステル化合物の2つのエステル結合の一方の加水分解を触媒する酵素のクラスを構成する。従って、ホスホジエステラーゼは様々な細胞プロセスにとって重要である。ホスホジエステラーゼには、細胞増殖および複製に必須であるDNAおよびRNAのエンドヌクレアーゼおよびエクソヌクレアーゼや、DNAのトポロ

50

ジ再編成中の核酸鎖の分解および再形成をするトポイソメラーゼが含まれる。Tyr-DNAホスホジエステラーゼは、トポイソメラーゼI型およびDNAとの間に形成されたデッドエンド共有結合中間体を加水分解することでDNAの修復に作用する(Pouliot, J. J.他(1999) Science 286:552-555; Yang, S.-W. (1996) Proc. Natl. Acad. Sci. USA 93:11534-11539)。

【0060】

酸性スフィンゴミエリナーゼは、膜リン脂質スフィンゴミエリンを加水分解してセラミドおよびホスホリルコリンを生成するホスホジエステラーゼである。ホスホリルコリンは、様々な細胞内シグナル伝達経路に関与するホスファチジルコリンの合成に用いられ、一方のセラミドは神経組織に高濃度で見られる膜脂質であるガングリオシドの生成のための必須前駆体である。酸性スフィンゴミエリナーゼが欠損すると、リソソームにおいてスフィンゴミエリン分子が蓄積され、それによってニーマン-ピック病が引き起こされる(Schuchman, E. H.およびS. R. Miranda (1997) Genet. Test. 1:13-19)。

10

【0061】

グリセロホスホリルジエステルホスホジエステラーゼ(glycerophosphoryl diester phosphodiesterase)(グリセロホスホジエステルホスホジエステラーゼとも呼ばれる)は、ジアセチル化リン脂質グリセロホスホジエステル(deacetylated phospholipid glycerophosphodiesters)を加水分解してsn-グリセロール-3-リン酸およびアルコールを生成するホスホジエステラーゼである。グリセロホスホコリン、グリセルホスホエタノールアミン(glycerophosphoethanolamine)、グリセロホスホグリセロール(glycerophosphoglycerol)、およびグリセロホスホイノシトール(glycerophosphoinositol)は、グリセロホスホリルジエステルホスホジエステラーゼのための基質の例である。大腸菌由来のグリセロホスホリルジエステルホスホジエステラーゼは、グリセロホスホジエステル(glycerophosphodiester)基質に対して広範な特異性を有する(Larson, T. J.他(1983) J. Biol. Chem. 248:5428-5432)。

20

30

【0062】

サイクリックヌクレオチドホスホジエステラーゼ(PDE)は、サイクリックヌクレオチドcAMPおよびcGMPの調節に極めて重要な酵素である。cAMPおよびcGMPは、ホルモン、光、および神経伝達物質を含む様々な細胞外シグナルを伝達する細胞内セカンドメッセンジャーとして機能する。PDEはサイクリックヌクレオチドをそれらの対応する一リン酸に分解し、それによってサイクリックヌクレオチドの細胞内濃度およびシグナル伝達におけるそれらの効果を調節する。それらの役割がシグナル伝達の制御因子であることから、PDEは化学療法の標的として大規模に研究された(Perry, M. J.およびG. A. Higgs (1998) Curr. Opin. Chem. Biol. 2:472-481; Torphy, J. T. (1998) Am. J. Resp. Crit. Care Med. 157:351-370)。

40

【0063】

哺乳動物PDEのファミリーは、それらの基質特異性および親和性、補助因子に対する感受性、および抑制剤に対する感受性に基づいて分類される(Beavo, J. A. (1995) Physiol. Rev. 75:725-748; Conti, M.他(1995) Endocrine Rev. 16:370-389)。これらのファミリーのいくつかは、固有の遺伝子を含み、その多くは様々な組織でスプライス変異体として発現される。PDEファミリーの中には、多数のイソ酵素およびこれらのイソ酵素の多数のスプライス変異体が存在する(Conti, M.およびS.-L. C. Jin (1999) Prog. Nucleic Acid Res. Mol. Biol. 63

50

: 1 - 38)。多数のPDEファミリー、イソ酵素、およびスプライス変異体の存在は、サイクリックヌクレオチドを伴う調節経路の多様性および複雑性を示すものである (Hoslay, M. D. および G. Milligan (1997) Trends Biochem. Sci. 22: 217 - 224)。

【0064】

PDE1型 (PDE1) は Ca^{2+} / カルモジュリン依存性であって、それぞれが少なくとも2つの異なるスプライス変異体を有する少なくとも3つの異なる遺伝子によってコードされると思われる (Kakkar, R. 他. (1999) Cell Mol. Life Sci. 55: 1164 - 1186)。PDE1は、肺、心臓、および脳で見出された。ある種のPDE1イソ酵素は、in vitroでリン酸化/脱リン酸化によって調節される。これらのPDE1イソ酵素のリン酸化すると、カルモジュリンに対するこの酵素の親和性が低下し、PDE活性が低下し、cAMPのレベルが安定する (Kakkar, 前出)。PDE1は、PDE1がサイクリックヌクレオチドおよびカルシウムのシグナル伝達の両方に関与することによる中枢神経系および心血管、免疫系の疾患のための有用な治療標的を提供し得る (Perry, M. J. および G. A. Higgs (1998) Curr. Opin. Chem. Biol. 2: 472 - 481)。

10

【0065】

PDE2は、小脳、新皮質、心臓、腎臓、肺、肺動脈、および骨格筋に見られるcGMP刺激PDE (cGMP-stimulated PDE) である (Sadhu, K. 他 (1999) J. Histochem. Cytochem. 47: 895 - 906)。

20

PDE2は、カテコールアミン分泌におけるcAMPの効果を仲介し、アルドステロンの調節に関与し (Beavo, 前出)、更に嗅覚シグナル伝達においても役割を果たしていると思われる (Juilfs, D. M. 他 (1997) Proc. Natl. Acad. Sci. USA 94: 3388 - 3395)。

【0066】

PDE3はcGMPおよびcAMPの両方に対して高い親和性を有するため、これらのサイクリックヌクレオチドがPDE3の競合的基質として作用する。PDE3は、心筋の収縮の刺激、血小板凝集の抑制、血管および気道の平滑筋の弛緩、Tリンパ球および血管平滑筋培養細胞の増殖抑制、脂肪組織からのカテコールアミン誘導性遊離脂肪酸放出の調節に作用する。ホスホジエステラーゼのPDE3ファミリーは、シロスタマイド (cilostamide)、エノキシモン (enoximone)、およびlixazinoneなどの特定のインヒビターに対する感受性を有する。PDE3のイソ酵素は、cAMP依存性プロテインキナーゼまたはインスリン依存性キナーゼによって抑制され得る (Degerman, E. 他 (1997) J. Biol. Chem. 272: 6823 - 6826)。

30

【0067】

PDE4はcAMPに特異的であって、気道平滑筋、血管上皮、および全ての炎症細胞に局在化し、cAMP依存性リン酸化によって活性化され得る。cAMPのレベルの上昇によって、炎症細胞の活性が抑制され気管平滑筋が弛緩し得るため、PDE4は、喘息治療の発見に重点をおいて新規の抗炎症剤の可能性のある標的として広範に研究された (Barad, M. 他 (1998) Proc. Natl. Acad. Sci. USA 95: 15020 - 15025)。PDE4インヒビターもまた、急性肺傷害、内毒素血症、リウマチ様関節炎、多発性硬化症、および様々な神経や胃腸の疾患に対する可能性のある治療薬として研究された (Doherty, A. M. (1999) Curr. Opin. Chem. Biol. 3: 466 - 473)。

40

【0068】

PDE5は、基質としてのcGMPに対して高い選択性を有し (Turko, I. V. 50

50

他(1998) *Biochemistry* 37:4200-4205)、2つのアロステリックcGMP特異的結合部位を有する(McAllister-Lucas, L. M. 他.(1995) *J. Biol. Chem.* 270:30671-30679)。cGMPのこれらのアロステリック結合部位への結合は、触媒活性の直接的な調節よりもcGMP依存性プロテインキナーゼによるPDE5のリン酸化にとって重要であると思われる。PDE5が高いレベルで、血管平滑筋、血小板、肺、および腎臓に見られる。インヒビターザプリナスト(zaprinast)はPDE5およびPDE1に対して効果がある。PDE5に対する特異性を得るためにザプリナストを改良してシルデナフィル(sildenafil)を生成した(VIAGRA; Pfizer, Inc., New York NY)。このsildenafilは男性勃起不全の治療薬である(Terrett, N. 他.(1996) *Bioorg. Med. Chem. Lett.* 6:1819-1824)。PDE5のインヒビターは、心血管治療薬として現在研究されている(Perry, M. J. および G. A. Higgs (1998) *Curr. Opin. Chem. Biol.* 2:472-481)。

10

【0069】

光受容体サイクリックヌクレオチドホスホジエステラーゼであるPDE6は、光伝達カスケードの重要な要素である。PDE6はGタンパク質トランスデュクションと結合して、cGMPを加水分解して光受容体膜におけるcGMP作動性陽イオンチャネルを調節する。cGMP結合活性部位に加えて、PDE6はまた、PDE6の機能における調節的な役割を果たすと考えられる2つの高親和性cGMP結合部位を有する(Artemyev, N. O. 他(1998) *Methods* 14:93-104)。PDE6の欠損は網膜の疾患に関係する。rdマウスの網膜変性症(Yan, W. 他(1998) *Invest. Ophthalmol. Vis. Sci.* 39:2529-2536)、ヒトの常染色体性劣性色素性網膜炎(Danciger, M. 他(1995) *Genomics* 30:1-7)、およびアイリッシュセッター犬の杆状体/錐状体異形成1型(Suber, M. L. 他(1993) *Proc. Natl. Acad. Sci. USA* 90:3968-3972)は、PDE6B遺伝子における突然変異が原因である。

20

【0070】

PDEのPDE7ファミリーは、複数のスプライス変異体を有する唯一つの既知のメンバーから成る(Bloom, T. J. および J. A. Beavo (1996) *Proc. Natl. Acad. Sci. USA* 93:14188-14192)。PDE7はcAMP特異的であるが、その他の生理的機能については殆ど知られていない。PDE7をコードするmRNAが骨格筋、心臓、脳、肺、腎臓、および膵臓で見られるが、PDE7タンパク質の発現は特定の組織型に限定される(Han, P. 他(1997) *J. Biol. Chem.* 272:16152-16157; Perry, M. J. および G. A. Higgs (1998) *Curr. Opin. Chem. Biol.* 2:472-481)。PDE7はPDE4ファミリーに密接に関連するが、PDE4の特異的なインヒビターであるロリプラムによって阻害されない(Beavo, 前出)。

30

【0071】

PDE8はcAMP特異的であり、PDE4ファミリーに密接に関連する。PDE8は、甲状腺、精巣、眼、肺、骨格筋、心臓、腎臓、卵巣、および脳において発現される。PDE8のcAMP加水分解活性は、PDEのインヒビターであるロリプラム、ビンボセチン、ミルリノン、IBMX(3-イソブチル-1-メチルキサンチン)、またはザプリナストによって抑制されないが、PDE8はジピリダモールによって抑制される(Fisher, D. A. 他(1998) *Biochem. Biophys. Res. Commun.* 246:570-577; Hayashi, M. 他(1998) *Biochem. Biophys. Res. Commun.* 250:751-756; Soderling, S. H. 他(1998) *Proc. Natl. Acad. Sci. USA* 95:8991-8996)。

40

50

【0072】

PDE9はcAMP特異的であって、PDEのPDE8ファミリーに最も類似している。PDE9は腎臓、肺、肝臓、脳、脾臓、および小腸で発現される。PDE9はsildenafil (VIAGRA; Pfizer, Inc., New York NY)、ロリプラム、ビンボセチン、ジピリダモール、またはIBMX (3-イソブチル-1-メチルキサンチン)によって抑制されないが、PDE5インヒビターであるザプリナストに対して感受性を有する (Fisher, D. A. 他 (1998) J. Biol. Chem. 273:15559-15564; Soderling, S. H. 他 (1998) J. Biol. Chem. 273:15553-15558)。

【0073】

PDE10は二重基質 (dual-substrate) PDEであって、cAMPおよびcGMPの両方を加水分解する。PDE10は、脳、甲状腺、及び精巣に発現する。(Soderling, S. H. 他 (1999) Proc. Natl. Acad. Sci. USA 96:7071-7076; Fujishige, K. 他 (1999) J. Biol. Chem. 274:18438-18445; Loughney, K. 他 (1999) Gene 234:109-117)。

【0074】

PDEは、約270-300のアミノ酸の触媒ドメイン、および補助因子の結合に必要なN末端調節ドメインを含み、場合によっては機能が未知の親水性C末端ドメインを含む (Conti, M. および S.-L. C. Jin (1999) Prog. Nucl. 20
eic Acid Res. Mol. Biol. 63:1-38)。保存された推定上の垂鉛結合モチーフであるHDXXHXGXXNが、全てのPDEの触媒ドメインにおいて同定された。N末端調節ドメインは、PDE2、PDE5、およびPDE6における非触媒cGMP結合ドメイン、PDE1におけるカルモジュリン結合ドメイン、およびPDE3およびPDE4におけるリン酸化部位を含むドメインを有する。PDE5では、N末端cGMP結合ドメインが約380のアミノ酸残基にまたがり、保存された配列モチーフN(R/K)XnFX₃DEのタンデムリピートを含む (McAllister-Lucas, L. M. 他 (1993) J. Biol. Chem. 268:22863-22873)。NKXnDモチーフは変異誘発によって見られ、cGMP結合に重要である (Turko, I. V. 他 (1996) J. Biol. Chem. 271:222 30
40-22244)。PDEファミリーは、触媒ドメイン内において約30%のアミノ酸同一性を有するが、同じファミリー内のイソ酵素は通常約85から95%のこの領域における同一性を示す (例えばPDE4AとPDE4B)。更に、あるファミリー内の触媒ドメイン外の類似性は高いが (60%を超える)、ファミリー間のこのドメイン外の配列類似性は殆ど存在しない。

【0075】

免疫反応および炎症反応を構成する作用の多くは、細胞内のcAMPのレベルを上昇させる薬剤によって阻害される (Verghese, M. W. 他 (1995) Mol. Pharmacol. 47:1164-1171)。様々な疾患がPDE活性の上昇が原因で起こり、サイクリックヌクレオチドのレベルの低下に関係する。例えば、マウスにおける尿崩症の或る型はPDE4活性の上昇に関係し、低いK_mをもったcAMP PDE活性の上昇がアトピー患者の白血球に見られ、PDE3が心疾患に関連する。 40

【0076】

PDEの多くのインヒビターが同定され、臨床試験が行われている (Perry, M. J. および G. A. Higgs (1998) Curr. Opin. Chem. Biol. 2:472-481; Torphy, T. J. (1998) Am. J. Respir. Crit. Care Med. 157:351-370)。PDE3インヒビターは、血小板凝集阻止薬、血圧降下薬、および鬱血性心不全の治療に有用な強心薬として開発された。PDE4インヒビターであるロリプラムは、抑鬱症の治療に用いられ、PDE4のその他のインヒビターは抗炎症薬として評価が行われている。ロリプラム 50

はまた、in vitroでHIV-1の複製を促すことが認められたリポ多糖(LPS)誘導性TNF-aを阻害することが分かった。従って、ロリプラムはHIV-1の複製を阻害すると考えられる(Angel, J. B.他(1995) AIDS 9:1137-1144)。更に、ロリプラムが、TNF-a、TNF-b、およびインターフェロン γ などのサイトカインの生成を抑制する能力に基づいて脳髄膜炎の治療に有効であることが示された。ロリプラムはまた、遅発性ジスキネジーに有効であると考えられ、実験動物モデルにおける多発性硬化症の治療に効果があった(Sommer, N.他(1995) Nat. Med. 1:244-248; Sasaki, H.他(1995) Eur. J. Pharmacol. 282:71-76)。

【0077】

テオフィリンは、気管支喘息およびその他の呼吸器疾患の治療に用いられる非特異的PDEインヒビターである。テオフィリンは、気道平滑筋の機能に作用し、呼吸器疾患の治療における抗炎症能力即ち免疫調節能力があると考えられる(Banner, K. H.およびC. P. Page(1995) Eur. Respir. J. 8:996-1000)。ペントキシフィリンは、間欠性跛行および糖尿病性末梢血管疾患の治療に用いられる別の非特異的PDEインヒビターである。ペントキシフィリンはまた、TNF-aの生成を阻止し、HIV-1の複製を阻害し得る(Angel他, 前出)。

【0078】

PDEは、様々な細胞型の細胞増殖に影響を与え(Conti他(1995) Endocrine Rev. 16:370-389)、様々な癌に関係すると報告された。前立腺癌細胞株DU145およびLNCaPの成長は、cAMP誘導体およびPDEインヒビターの送達によって抑制された(Bang, Y. J.他(1994) Proc. Natl. Acad. Sci. USA 91:5330-5334)。これらの細胞はまた、上皮からニューロン形態への表現型における永久的な変換を示した。また、PDEインヒビターがメサングウム細胞の増殖を調節する可能性があり(Matousovich, K.他(1995) J. Clin. Invest. 96:401-410)、またリンパ球の増殖を調節する可能性もある(Joulain, C.他(1995) J. Lipid Mediat. Cell Signal. 11:63-79)ことが示唆された。PDEを腫瘍の特定の細胞区画に送達し、細胞死を導く癌治療が報告されている(Deonaraian, M. P.およびA. A. Epenetos(1994) Br. J. Cancer 70:786-794)。

【0079】

ホスホトリエステラーゼ

ホスホトリエステラーゼ(PTE, paraoxonases)は、毒性有機リン化合物を加水分解する酵素であって、様々な組織から単離された。この酵素は、哺乳動物には豊富に存在するが鳥や昆虫では不足していると思われ、鳥や昆虫の有機リン化合物に対する耐性の低さの説明となる(Vilanova, E.およびSogorb, M. A.(1999) Crit. Rev. Toxicol. 29:21-57)。ホスホトリエステラーゼは、哺乳動物による殺虫剤の解毒において中心的な役割を果たす。ホスホトリエステラーゼ活性は個人によって差があり、大人より幼児の方が低い。ノックアウトマウスは、有機リン系の毒素であるダイアジノン オキシソン(diazoxon)およびクロルピリホス オキシソン(chlorpyrifos oxon)に対して顕著な感受性を有する(Furlong, C. E.,他(2000) Neurotoxicology 21:91-100)。PTEは、有機リン含有化学廃棄物および化学兵器(例えば、パラチオン)、並びに農薬や殺虫剤の解毒能力を有する酵素として注目されている。ある研究により、ホスホトリエステラーゼがアテローム性動脈硬化症およびリポタンパク代謝に関係する疾患に関与することが示された。

【0080】

チオエステラーゼ

脂肪酸生合成に関係する2つの可溶性チオエステラーゼが、哺乳動物組織から単離された

10

20

30

40

50

。その内の一方は、長鎖脂肪酸アシルチオエステルに対してのみ活性であり、他方は様々な長さの脂肪酸アシル鎖を有するチオエステルに対して活性である。これらのチオエステラーゼは、脂肪酸の新規合成における脂肪酸鎖の終結段階を触媒する。脂肪酸鎖の終結段階は、脂肪酸アシル鎖を脂肪酸シターゼのアシルキャリアタンパク質 (ACP) のサブユニットの 4' - ホスホパンテテイン補欠分子族と結合させるチオエステル結合の加水分解を伴う (Smith, S. (1981a) *Methods Enzymol.* 71:181-188; Smith, S. (1981b) *Methods Enzymol.* 71:188-200)。

【0081】

大腸菌は、長鎖アシルチオエステルに対してのみ活性なチオエステラーゼ I 型および様々な長さの鎖に対して特異性を有するチオエステラーゼ II 型 (TEII) の 2 つの可溶性チオエステラーゼを含む (Nagget, J. 他 (1991) *J. Biol. Chem.* 266:11044-11050)。大腸菌 TEII は、新規の脂肪酸生合成における脂肪酸鎖終結酵素として機能する哺乳動物チオエステラーゼの 2 つの型のいずれとも配列類似性を有していない。哺乳動物チオエステラーゼとは異なり、大腸菌 TEII は、特徴的なセリン活性部位 gly-X-ser-X-gly 配列モチーフを含まず、セリン変性剤であるジソプロピルフルオロリン酸によって不活化されない。しかしながら、ヨードアセトアミドおよびジエチルピロカルボネートによるヒスチジン 58 の修飾によって PEII 活性が失われる。TEII の過剰な発現は大腸菌に含まれる脂肪酸を変化させない。これは、脂肪酸生合成において脂肪酸鎖終結酵素として機能していないことを示すものである (Nagget 他, 前出)。このような理由から、Nagget 他 (前出) は、大腸菌 TEII の生理学的基質が、ACP - ホスホパンテテイン脂肪酸エステルではなく、補酵素 A (CoA) 脂肪酸エステルであると提案した。

【0082】

カルボキシルエステラーゼ

哺乳動物カルボキシルエステラーゼは、様々な組織および細胞型で発現される多重遺伝子ファミリーを構成する。イソ酵素は有意な配列相同性を有し、主にアミノ酸配列に基づいて分類される。アセチルコリンエステラーゼ、ブチリルコリンエステラーゼ、およびカルボキシルエステラーゼは、エステラーゼのセリンスーパーファミリー (B エステラーゼ) に分類される。その他のカルボキシルエステラーゼには、チログロブリン、トロンピン、IX 因子、gliotactin、およびプラスミノゲンがある。カルボキシルエステラーゼは、分子のエステル基およびアミド基の加水分解を触媒し、薬剤、環境毒、および発癌物質の解毒に関与する。カルボキシルエステラーゼの基質には短鎖および長鎖アシルグリセロール、アシルカルニチン、炭酸塩、ジピペフリン塩酸塩 (dipivefrin hydrochloride)、コカイン、サリチル酸塩、カプサイシン、パルミトイル-CoA、イミダプリル、ハロペリドール、ピロリジジンアルカロイド、ステロイド、p-ニトロフェニル酢酸、マラチオン、butanilicaine、およびイソカルボキサジドが含まれる。この酵素は低い基質特異性を示すことがよくある。カルボキシルエステラーゼはまた、プロドラッグを対応する遊離酸に変換するために重要である。

対応する遊離酸は、例えば、血中コレステロールを低下させるために用いられるロバスタチンなどのそのプロドラッグの活性型であり得る (Sato, T. および Hosokawa, M. (1998) *Annu. Rev. Pharmacol. Toxicol.* 38:257-288 を参照)。

【0083】

neuroligins は、(i) N 末端シグナル配列を有し、(ii) 細胞表面受容体に類似し、(iii) カルボキシルエステラーゼドメインを含み、(iv) 脳で高発現され、(v) カルシウム依存的にニューレキシンと結合する分子のクラスである。カルボキシルエステラーゼと相同性を有するにもかかわらず、neuroligins は活性部位セリン残基を含まず、触媒としてではなく基質結合において役割を果たすと考えられる (Ichtchenko, K. 他 (1996) *J. Biol. Chem.* 271:2

10

20

30

40

50

676 - 2682)。

【0084】

スクアレンエポキシダーゼ

スクアレンエポキシダーゼ(スクアレンモノオキシゲナーゼ、SE)は、ミクロソーム膜に結合しFADに依存性のオキシドレクターゼであって、真核細胞のステロール生合成経路における初めの酸素添加ステップを触媒する。コレステロールは、LDL受容体仲介経路若しくは生合成経路によって獲得される細胞質膜の必須の構成成分である。後者の場合、コレステロール分子における27全ての炭素原子がアセチル-CoAに由来する(Stryer, L., 前出)。SEはスクアレンをまず2,3(S)-オキシドスクアレンに変換し、次にラノステロールに変換し、更にコレステロールに変換する。コレステロール生合成に係るステップを以下に要約する(Stryer, L (1988) Biochemistry. W. H FreemanおよびCo., Inc. New York. .554-560頁およびSakakibara, J.他(1995) 270:17-20)。アセテート(アセチル-CoA由来) 3ヒドロキシ-3-メチル-グルタリルCoA メバロン酸 5-ホスホメバロン酸 5-ピロホスホメバロン酸 イソペンテニルピロリン酸 ジメチルアリルピロリン酸 グラニルピロリン酸 ファルネシルピロリン酸 スクアレン スクアレンエポキシド ラノステロール コレステロール

10

【0085】

コレステロールは真核細胞の生存に必須であるが、血清コレステロールのレベルが過度に上昇すると高等生物の動脈においてアテローム斑が形成されるようになる。例えば冠状動脈などの必須の血管壁部に不溶性の脂質が蓄積されると、血流が減少して十分な血液が組織に流れなくなり組織壊死が起こる可能性がある。HMG-CoAレダクターゼは、3-ヒドロキシ-3-メチル-グルタリルCoA(HMG-CoA)のメバロン酸への変換に必要であり、この変換がコレステロール生合成の第1のステップである。HMG-CoAは、血漿コレステロールレベルを低下させるようにデザインされた様々な医薬化合物の標的である。しかしながら、MHG-CoAの阻害はまた、その他の生合成経路に必要な非ステロール中間体(例えば、メバロン酸)の合成が減少する。SEは、ステロール合成経路の後の方で起こる律速反応を触媒し、コレステロールはSEによる触媒ステップの後の経路の最終産物である。従って、SEは、その他の必要な中間体を減少させない抗高脂血症薬をデザインするための理想的な標的である(Nakamura, Y.他(1996) 271:8053-8056)。

20

30

【0086】

エポキシドヒドロラーゼ

エポキシドヒドロラーゼは、エポキシド含有化合物の水の添加を触媒し、それによってエポキシドがその対応する1,2-ジオールに加水分解される。これらは、細菌ハロアルカンデハロゲナーゼ(haloalkane dehalogenase)に関連し、酵素の / ヒドロラーゼフォールド(/ hydrolase fold)ファミリーのその他のメンバーと配列類似性を有する(例えば、黄色ブドウ球菌(Streptomyces aureofaciens)由来ブロモペルオキシダーゼA2(bromoperoxidase A2)、緑膿菌(Pseudomonas putida)由来のヒドロキシムコン酸セミアルデヒド ヒドロラーゼ、および水素細菌(Xanthobacter autotrophicus)由来ハロアルカンデハロゲナーゼ)。エポキシドヒドロラーゼは遍在性であって、哺乳動物、脊椎動物、植物、真菌、細菌に見られる。この酵素のファミリーは、生物内に導入されると求電子性が高く破壊性である場合が多い生体異物エポキシド化合物の解毒にとって重要である。エポキシドヒドロラーゼ反応の例には、cis-9,10-epoxyoctadec-9(Z)-enoic acid(ロイコトキシン)からその対応するジオール(threo-9,10-dihydroxyoctadec-12(Z)-enoic acid(ロイコトキシンジオール))への加水分解、およびcis-12,13-epoxyoctadec-9(Z)-enoic acid(イソロイコトキシン)からその対応するジオール(diol th

40

50

reo-12,13-dihydroxyoctadec-9(Z)-enoic acid (イソロイコトキシジオール)への加水分解が含まれる。ロイコトキシンは膜の透過性を変えてイオンを輸送し炎症反応を引き起こす。加えて、エポキシド発癌物質は、薬剤および環境毒素の解毒における中間体としてチトクロームP450によって生成されることが知られている。

【0087】

この酵素は、Asp(求核性)、Asp(ヒスチジンを支持する酸)、およびHis(水活性化ヒスチジン)の3つの触媒部分を有する。エポキシドヒドロラーゼの反応のメカニズムは、標的分子のエポキシド環の第1の炭素原子に対するAsp残基の1つへの求核攻撃によって開始される共有結合エステル中間体によって始まり、共有結合エステル中間体が形成される(Michael Arand, M.他(1996) J. Biol. Chem. 271:4223-4229; Rink, R.他(1997) J. Biol. Chem. 272:14650-14657; Argiriadi, M. A. 他(2000) J. Biol. Chem. 275:15265-15270)。

10

【0088】

カテコール-O-メチルトランスフェラーゼ

カテコール-O-メチルトランスフェラーゼ(COMT)は、カテコール基質における1つのヒドロキシル基(例えば、エルドーパ、ドーパミン、またはDBA)へのS-アデノシル-L-メチオニン(AdoMet; SAM)供与体のメチル基の転移を触媒する。3'-ヒドロキシル基のメチル化は4'-ヒドロキシル基のメチル化より優先され、COMTの膜結合アイソフォームが可溶性よりも位置特異的である。この酵素の可溶性の翻訳は、完全長mRNA(1.5kb)の内部の開始コドンの利用によるか、或いは内部のプロモーターから転写された短いmRNA(1.3kb)の翻訳による。提案されたS_N2様メチル化反応には、Mg²⁺が必要であり、Ca²⁺によって抑制される。供与体および基質のCOMTへの結合が連続的に起こる。まず、AdoMetがMg²⁺非依存的にCOMTに結合し、Mg²⁺の結合およびカテコール基質の結合がそれに続く。

20

組織におけるCOMTの量は活性のために通常必要とされる量より多く存在するため阻害が困難である。しかしながら、インヒビターが*in vitro*での使用(例えば、没食子酸、トロポロン、U-0521、および3',4'-dihydroxy-2-methyl-propionophetrolone)および臨床での使用(例えば、nitrocatechol系化合物およびtolcapone)のために開発された。これらのインヒビターを投与すると、エルドーパの半減期が長くなり、続いてドーパミンの生成が起こる。COMTの阻害により、エピネフリン/ノルエピネフリン、イソプレナリン、リミテロール、ドブタミン、フェノルドパム(fenoldopam)、アボモルフィン、および-メチルドパを含む他の様々なカテコール構造化合物の半減期が長くなると思われる。ノルエピネフリンの欠損は臨床的抑鬱症に繋がるため、COMTインヒビターの使用は抑鬱性の治療に有用であると思われる。COMTインヒビターは通常、副作用が最小限であり、最終的に肝臓で代謝され、体内にはわずかな代謝物が残るのみである(Mannisto, P. T.およびKaakkola, S. (1999) Pharmacological Reviews 51:593-628)。

30

40

複数の新規の薬剤代謝酵素、および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、自己免疫/炎症の疾患、細胞増殖異常、発生または発達障害、内分泌障害、眼の疾患、代謝障害、および肝臓の疾患を含む胃腸疾患の診断・治療・予防において有用であり、また、薬剤代謝酵素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。

【0089】

(発明の概要)

本発明は、総称して「DME」、個別にはそれぞれ「DME-1」、「DME-2」、「DME-3」、「DME-4」、「DME-5」、「DME-6」、「DME-7」、「

50

DME - 8」、「DME - 9」、「DME - 10」、「DME - 11」、「DME - 12」、「DME - 13」、「DME - 14」、「DME - 15」、「DME - 16」、「DME - 17」、「DME - 18」、および「DME - 19」と呼ぶ薬剤代謝酵素である精製されたポリペプチドを提供する。或る実施態様において本発明は、(a) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列からなるポリペプチド、(b) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を含む群から選択した単離されたポリペプチドを提供する。一実施態様では、SEQ ID NO: 1 - 19 のアミノ酸配列を含む単離されたポリペプチドを提供する。 10

【0090】

また、本発明は(a) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列と少なくとも90%の相同性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列の生物学的活性断片、または(d) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列の免疫原性断片を含む群から選択されたポリペプチドをコードするような単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドはSEQ ID NO: 1 - 19 からなる群から選択したポリペプチドをコードする。別の実施態様では、ポリヌクレオチドはSEQ ID NO: 20 - 38 からなる群から選択される。 20

【0091】

本発明は更に、(a) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列と少なくとも90%の相同性のある天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 19 を有する群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、または(d) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片を有する群から選択したポリペプチドをコードするようなポリヌクレオチドと機能的に結合したプロモーター配列を有する組換えポリヌクレオチドを提供する。一実施態様では、本発明は組換えポリヌクレオチドを用いて形質転換した細胞を提供する。別の実施態様では、本発明はこの組換えポリヌクレオチドを含む遺伝形質転換体を提供する。 30

【0092】

また、本発明は、(a) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列と少なくとも90%が同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチドの生物学的活性断片、および(d) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片をから構成される群から選択されたポリペプチドを製造する方法を提供する。製造方法は、(a) 組換えポリヌクレオチドを用いて形質転換した細胞をポリペプチドの発現に適した条件下で培養する過程と、(b) そのように発現したポリペプチドを回収する過程とを有し、組換えポリヌクレオチドはポリペプチドをコードするポリヌクレオチドに機能的に結合したプロモーター配列を有する。 40

【0093】

本発明は更に、(a) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列と少なくとも90%同一である天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 19 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 50

の生物学的活性断片、および (d) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチドの免疫原性断片から構成される群から選択されたポリペプチドに特異結合するような単離された抗体を提供する。

【0094】

本発明は更に、(a) SEQ ID NO: 20 - 38 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ポリヌクレオチド、(b) SEQ ID NO: 20 - 38 からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を有するポリヌクレオチド、(c) (a) に相補的なポリヌクレオチド配列、(d) (b) に相補的なポリヌクレオチド配列、および (e) (a) ~ (d) のRNA等価物からなる群から選択された単離されたポリヌクレオチドを提供する。一実施態様では、ポリヌクレオチドは少なくとも60の連続したヌクレオチドを含む。

10

【0095】

本発明は更に、サンプル中の標的ポリヌクレオチドを検出する方法を提供する。ここで、標的ポリヌクレオチドは (a) SEQ ID NO: 20 - 38 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: 20 - 38 からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c) (a) に相補的なポリヌクレオチド、(d) (b) に相補的なポリヌクレオチド配列、および (e) (a) ~ (d) のRNA等価物からなる群から選択されたポリヌクレオチド配列を含む。検出方法は、(a) サンプル中の標的ポリヌクレオチドに相補的な配列を含む少なくとも20の連続したヌクレオチドを含むプローブを用いて該サンプルをハイブリダイズする過程と、(b) ハイブリダイゼーション複合体の存在・不存在を検出し、複合体が存在する場合にはオプションでその量を検出する過程からなり、プローブと標的ポリヌクレオチドあるいはその断片の間でハイブリダイゼーション複合体が形成されるような条件下で、プローブは標的ポリヌクレオチドに特異的にハイブリダイズする。一実施態様では、プローブは少なくとも60の連続したヌクレオチドを含む。

20

【0096】

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: 20 - 38 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: 20 - 38 からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(c) (a) のポリヌクレオチドに相補的なポリヌクレオチド、(d) (b) のポリヌクレオチドに相補的なポリヌクレオチド、および (e) (a) ~ (d) のRNA等価物からなる群から選択された配列のポリヌクレオチドを有する。検出方法は、(a) ポリメラーゼ連鎖反応増幅を用いて標的ポリヌクレオチドまたはその断片を増幅する過程と、(b) 標的ポリヌクレオチドまたはその断片の存在・不存在を検出し、該標的ポリヌクレオチドまたはその断片が存在する場合にはオプションでその量を検出する過程を含む。

30

【0097】

本発明は更に、有効量のポリペプチドと薬剤として許容できる賦形剤とを含む成分を提供する。有効量のポリペプチドは、(a) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を含むポリペプチドの生物学的活性断片、および (d) SEQ ID NO: 1 - 19 を有する群から選択したアミノ酸配列の免疫原性断片からなる群から選択される。一実施例では、SEQ ID NO: 1 - 19 からなる一群から選択されたアミノ酸配列を含む組成物を提供する。更に、本発明は、患者にこの組成物を投与することを含む、機能的DMEの発現の低下に関連した疾患やその症状の治療方法を提供する。

40

【0098】

50

本発明はまた、(a) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を有するポリペプチド、(b) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列の生物学的活性断片、および(d) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列の免疫原性断片からなる群から選択されたポリペプチドのアゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドを有するサンプルを化合物に曝す過程と、(b) サンプル中のアゴニスト活性を検出する過程とを含む。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的DMEの発現の低下に関連した疾患やその症状の治療方法を提供する。

10

【0099】

本発明は更に、(a) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列の生物学的活性断片、および(d) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列のポリペプチドの免疫原性断片からなる群から選択されたポリペプチドのアンタゴニストとしての有効性を確認するために化合物をスクリーニングする方法を提供する。スクリーニ

20

【0100】

本発明は更に、(a) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列を含むポリペプチド、(b) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列と少なくとも90%の同一性を有する天然のアミノ酸配列を含むポリペプチド、(c) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列の生物学的活性断片、または(d) SEQ ID NO: 1 - 19 からなる群から選択したアミノ酸配列の免疫原性断片を含むポリペプチドに特異結合する化合物をスクリーニングする方法を提供する。スクリーニング方法は、(a) ポリペプチドを適切な条件下で少なくとも1つの試験化合物に結合させる過程と、(b) 試験化合物とのポリペプチドの結合を検出し、それによってポリペプチドに特異結合する化合物を同定する過程とを含む。

30

【0101】

更に本発明は、(a) SEQ ID NO: 1 - 19 からなる一群から選択されたアミノ酸配列からなるポリペプチドと、(b) SEQ ID NO: 1 - 19 からなる一群から選択されたアミノ酸配列と90%以上の配列同一性を有する天然のアミノ酸配列からなるポリペプチドと、(c) SEQ ID NO: 1 - 19 からなる一群から選択されたアミノ酸配列を有するポリペプチドの生物学的活性断片と、(d) SEQ ID NO: 1 - 19 からなる一群から選択されたアミノ酸配列を有するポリペプチドの免疫原性断片とで構成される群から選択されたポリペプチドの活性を調節する化合物をスクリーニングする方法を提供する。その方法は、(a) ポリペプチドの活性が許容された条件下で、ポリペプチドを少なくとも1つの試験化合物と結合させる過程と、(b) ポリペプチドの活性を試験化合物の存在下で算定する過程と、(c) 試験化合物の存在下でのポリペプチドの活性を試験化合物の不存在下でのポリペプチドの活性と比較する過程とを含み、試験化合物の存在下でのポリペプチドの活性の変化は、ポリペプチドの活性を調節する化合物であることを意味する。

40

【0102】

50

本発明は更に、標的ポリヌクレオチドの変異発現の有効性を確認するために化合物をスクリーニングする方法を提供する。標的ポリヌクレオチドは、SEQ ID NO: 20 - 38を有する群から選択した配列を含む。スクリーニング方法は、(a)標的ポリヌクレオチドを含むサンプルを化合物に曝す過程と、(b)標的ポリヌクレオチドの変異発現を検出する過程とを含む。

【0103】

本発明は更に、(a)核酸を含む生物学的サンプルを試験化合物で処理する過程と、(b)(i)SEQ ID NO: 20 - 38からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(ii)SEQ ID NO: 20 - 38からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(iii)(i)に相補的な配列を有するポリヌクレオチド、(iv)(ii)のポリヌクレオチドに相補的なポリヌクレオチド、(v)(i)~(iv)のRNA等価物からなる群から選択したポリヌクレオチドの少なくとも20の連続したヌクレオチドを含むプローブを用いて、処理した生物学的サンプルの核酸をハイブリダイズする過程とを含む試験化合物の毒性の算定方法を提供する。ハイブリダイゼーションは、上記プローブと生物学的サンプル中の標的ポリヌクレオチドの間に特定のハイブリダイゼーション複合体が形成されるような条件下で発生し、上記標的ポリヌクレオチドは、(i)SEQ ID NO: 20 - 38からなる群から選択したポリヌクレオチド配列を含むポリヌクレオチド、(ii)SEQ ID NO: 20 - 38からなる群から選択したポリヌクレオチド配列と少なくとも90%の同一性を有する天然のポリヌクレオチド配列を含むポリヌクレオチド、(iii)(i)のポリヌクレオチドに相補的な配列を有するポリヌクレオチド、(iv)(ii)のポリヌクレオチドに相補的なポリヌクレオチド、および(v)(i)~(iv)のRNA等価物からなる群から選択する。或いは、標的ポリヌクレオチドは、上記(i)~(v)からなる群から選択したポリヌクレオチド配列の断片と、(c)ハイブリダイゼーション複合体の量を定量する過程と、(d)処理した生物学的サンプルのハイブリダイゼーション複合体の量を、非処理の生物学的サンプルのハイブリダイゼーション複合体の量と比較する過程を含み、処理した生物学的サンプルのハイブリダイゼーション複合体の量の差は、試験化合物の毒性を示す。

【0104】

(発明を実施するための形態)

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

【0105】

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

【0106】

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

【0107】

(定義)

用語「DME」は、天然、合成、半合成或いは組換え体など全ての種（特にウシ、ヒツジ、ブタ、マウス、ウマ及びヒトを含む哺乳動物）から得られる実質的に精製されたDMEのアミノ酸配列を指す。

【0108】

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

10

【0109】

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

【0110】

DMEをコードする「変異」核酸配列は、様々なヌクレオチドの欠失、挿入、或いは置換が起こっても、DMEと同じポリペプチド或いはDMEの機能特性の少なくとも1つを備えるポリペプチドを指す。この定義には、DMEをコードするポリヌクレオチド配列の正常な染色体の遺伝子座ではない位置での対立遺伝子変異配列との不相当或いは予期しないハイブリダイゼーション、並びにDMEをコードするポリヌクレオチドの特定のオリゴヌクレオチドプローブを用いて容易に検出可能な或いは検出困難な多型性を含む。コードされたタンパク質も変異され得り、サイレント変化を生じDMEと機能的に等価となるアミノ酸残基の欠失、挿入、或いは置換を含み得る。意図的なアミノ酸置換は、生物学的或いは免疫学的にDMEの活性が保持される範囲で、残基の極性、電荷、溶解度、疎水性、親水性、及び/または両親媒性についての類似性に基づいて成され得る。例えば、負に帯電したアミノ酸にはアスパラギン酸及びグルタミン酸があり、正に帯電したアミノ酸にはリジン及びアルギニンがある。親水性値が近似している非荷電極性側鎖を有するアミノ酸には、アスパラギンとグルタミン、セリンとスレオニンがある。親水性値が近似している非荷電側鎖を有するアミノ酸には、ロイシンとイソロイシンとバリン、グリシンとアラニン、フェニルアラニンとチロシンがある。

20

30

【0111】

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

40

【0112】

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

【0113】

用語「アンタゴニスト」は、DMEの生物学的活性を阻害或いは減弱する分子である。アンタゴニストは、DMEに直接相互作用するか、或いはDMEが関与する生物学的経路の成分と作用して、DMEの活性を調節する抗体、核酸、糖質、小分子、任意の他の化合物や組成物などのタンパク質を含み得る。

【0114】

「抗体」の語は、エピトープの決定基と結合することができる、そのままの免疫グロブリ

50

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

【0115】

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

【0116】

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

【0117】

「生物学的に活性」の語は、天然分子の構造的機能、調節機能または生化学的機能を有するタンパク質を指す。同様に、用語「免疫学的に活性」または「免疫原性」は、天然或いは組換え体のD M E、合成のD M Eまたはそれらの任意のオリゴペプチドが、適当な動物或いは細胞の特定の免疫応答を誘発して特定の抗体と結合する能力を指す。

【0118】

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

【0119】

「所定のポリヌクレオチド配列を含む成分」及び「所定のアミノ酸配列を含む成分」は、所定のポリヌクレオチド配列またはアミノ酸配列を含む広範囲の任意の成分を指す。この成分には、乾燥製剤または水溶液が含まれ得る。D M E 若しくはD M E の断片をコードするポリヌクレオチド配列を含む組成物は、ハイブリダイゼーションプローブとして使用され得る。このプローブは、凍結乾燥状態で保存可能であり、糖質などの安定化剤と結合させることが可能である。ハイブリダイゼーションにおいては、塩（例えばN a C l）、界面活性剤（例えばドデシル硫酸ナトリウム；S D S）及びその他の構成エレメント（例えばデンハート液、脱脂粉乳、サケの精子のD N A等）を含む水溶液中にプローブを分散させることができる。

【0120】

「コンセンサス配列」は、不要な塩基を分離するためにD N A 配列の解析を繰り返し行い、X L - P C R キット（P E B i o s y s t e m s , F o s t e r C i t y C A）を用いて5'及び/または3'の方向に伸長され、再度シーケンシングされた核酸配列、

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

【0121】

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

10

元残基	保存的置換
Aa	GySt
Ag	HsLys
An	ApGnHs
Ap	AnGu
Gs	AaSer
Gn	AnGuHs
Gu	ApGnHs
Gy	Aa
Hs	AnAgGnGu
Ile	LeuVal
Leu	IleVal
Lys	Ag
Met	Leu
Phe	Hs
Ser	Gs
Thr	Ser
Tyr	Phe
Tyr	Hs
Val	IleLeu

20

30

40

【0122】

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

【0123】

50

「欠失」は、結果的に1個若しくは数個のアミノ酸残基またはヌクレオチドが失われてなくなるようなアミノ酸またはヌクレオチド配列における変化を指す。

【0124】

「誘導體」の語は、ポリペプチドまたはポリヌクレオチドの化学修飾を指す。例えば、アルキル基、アシル基、ヒドロキシル基またはアミノ基による水素の置換は、ポリヌクレオチド配列の化学修飾に含まれ得る。ポリヌクレオチド誘導體は、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。ポリペプチド誘導體は、グリコシル化、ポリエチレングリコール化 (pegylation)、或いは任意の同様なプロセスであって誘導起源のポリペプチドから少なくとも1つの生物学的若しくは免疫学的機能を保持しているプロセスによって、修飾されたポリペプチドである。

【0125】

「検出可能な標識」は、測定可能な信号を生成することができ、ポリヌクレオチドまたはポリペプチドに共有結合または非共有結合するようなレポーター分子または酵素を指す。

【0126】

「示差発現」は少なくとも2つの異なったサンプルを比較することによって決められる、増加、または非調節、あるいは減少、下方調節、または欠損遺伝子またはタンパク発現を指す。このような比較は例えば、治療後サンプルと未治療のサンプルまたは病態のサンプルと正常サンプルの間で行われ得る。

【0127】

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

【0128】

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

【0129】

SEQ ID NO: 20 - 38の断片には、固有のポリヌクレオチド配列領域が含まれる。この領域は、SEQ ID NO: 20 - 38を特異的に同定するものであり、例えばこの断片を得たゲノム中のSEQ ID NO: 4 - 6以外の配列とは異なるものである。SEQ ID NO: 20 - 38のある断片は、例えば、ハイブリダイゼーションや増幅技術、またはSEQ ID NO: 3 - 38を関連ポリヌクレオチド配列から区別する類似の方法に有用である。SEQ ID NO: 20 - 38の断片の正確な長さ及び断片に対応するSEQ ID NO: 20 - 38の領域は、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

【0130】

SEQ ID NO: 1 - 19の断片は、SEQ ID NO: 20 - 38の断片によってコードされる。SEQ ID NO: 1 - 19の断片には、SEQ ID NO: 1 - 19

を特異的に同定する固有のアミノ酸配列領域が含まれている。例えば、SEQ ID NO: 1 - 19 の断片は、SEQ ID NO: 1 - 19 を特異認識する抗体を産出するための免疫原性ペプチドとして有用である。SEQ ID NO: 1 - 19 の断片及び断片に対応するSEQ ID NO: 1 - 19 の領域の正確な長さは、断片に対する意図した目的に基づき当業者が慣例的に決定することが可能である。

【0131】

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

【0132】

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

【0133】

ポリヌクレオチド配列に適用される「一致率」または「一致%」の語は、標準化されたアルゴリズムを用いてアラインメントされた少なくとも2つ以上のポリヌクレオチド配列間で一致する残基の割合を意味する。このようなアルゴリズムは、2配列間のアラインメントを最適化するために比較する配列において、標準化された再現性のある方法でギャップを挿入するので、2つの配列をより有意に比較できる。

【0134】

ポリヌクレオチド配列間の一致率は、MEGALIGN version 3.12e 配列アラインメントプログラムに組込まれているようなCLUSTAL Vアルゴリズムのデフォルトのパラメータを用いて決定できる。このプログラムは、LASERGENE ソフトウェアパッケージ（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. 及び P.M. Sharp (1989) CABIOS 5:151-153、Higgins, D.G. 他 (1992) CABIOS 8:189-191に記載されている。ポリヌクレオチド配列を2つ1組でアラインメントする際のデフォルトパラメータは、Ktuple = 2、gap penalty = 5、window = 4、「diagonals saved」= 4と設定する。Ktuple = 2、gap penalty = 5、window = 4、「diagonals saved」= 4と設定する。残基「重み付け」表をデフォルトとして選択する。CLUSTAL Vは、アラインメントされたポリヌクレオチド配列対間の「類似率」として一致率を報告する。

【0135】

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

10

20

30

40

50

」ツール Version 2.0.12 (2000年4月21日) を用いて `blastn` を実行してもよい。デフォルトパラメータの設定例を以下に示す。

【0136】

`Matrix:HCBLM2`

`Reward:match:11`

`Penalty:mismatch:-2`

`OpenGap:5` 及び `ExtensionGap:2penalties`

`Gapxdropoff:50`

`Expect:10`

`WordSize:11`

`Filter:n`

10

【0137】

一致率は、完全に画定された（例えば特定の配列番号で画定された）配列長さと比較して測定し得る。或いは、より短い長さ、例えばより大きな画定された配列から得られた断片（例えば少なくとも20、30、40、50、70、100または200の連続したヌクレオチドの断片）の長さと比較して一致率を測定してもよい。ここに挙げた長さは単なる例示的なものに過ぎず、表、図及び配列リストを含めた本明細書に記載された配列に裏付けられた任意の配列長さの断片を用いて、一致率を測定し得る長さを説明し得ることを理解されたい。

20

【0138】

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

30

【0139】

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

【0140】

ポリペプチド配列間の一致率は、`MEGALIGN version 3.12e` 配列アラインメントプログラムに組込まれているような `CLUSTAL V` アルゴリズムのデフォルトのパラメータを用いて決定できる（既に説明したのでそれを参照されたい）。`CLUSTAL V` を用いて、ポリペプチド配列を2つ1組でアラインメントする際のデフォルトパラメータは、`Ktuple=1`、`gap penalty=3`、`window=5`、`「diagonals saved」=5` と設定する。デフォルトの残基重み付け表として `PAM250` マトリクスを選択する。ポリヌクレオチドアラインメントと同様に、`CLUSTAL V` は、アラインメントされたポリペプチド配列対間の「類似率」として一致率を報告する。

40

【0141】

或いは、`NCBI BLAST` ソフトウェア一式を用いてもよい。例えば、2つのポリペ

50

ブチド配列を対で比較をする場合、ある者は、デフォルトパラメータで設定された「BLAST 2 Sequences」ツール Version 2.0.12 (Apr - 21 - 2000) で blastp を使用するであろう。デフォルトパラメータの設定例を以下に示す。

【0142】

MixHCSUM2

OpenCap 11 及 ExtensionCap 1 penalties

Capxdcpoff 50

Expect 10

WordSize 3

Filter m

10

【0143】

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

20

【0144】

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

【0145】

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

30

【0146】

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

40

【0147】

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

50

及び pH における特定の配列の融点 (T_m) より約 5 ~ 20 低くなるように選択する。この T_m は、所定のイオン強度及び pH の下で、完全に一致するプローブに標的配列の 50% がハイブリダイズする温度である。T_m を計算する式及び核酸のハイブリダイゼーション条件はよく知られており、Sambrook ら (1989) Molecular Cloning: A Laboratory Manual, 第 2 版, 1 - 3 巻, Cold Spring Harbor Press, Plainview NY に記載されており、特に 2 巻の 9 章を参照されたい。

【0148】

本発明のポリヌクレオチド間の高いストリンジェンシーのハイブリダイゼーションでは、約 0.2 x SSC 及び約 1% の SDS の存在の下、約 68 で 1 時間の洗浄過程を含む。別法では、65、60、55、42 の温度で行う。SSC 濃度は、約 0.1% の SDS 存在下で、約 0.1 ~ 2 x SSC の範囲で変化し得る。通常は、ブロッキング剤を用いて非特異ハイブリダイゼーションを阻止する。このような遮断剤には、例えば、約 100 ~ 200 µg/ml の変性サケ精子 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に有機溶剤、例えば約 35 ~ 50% v/v の濃度のホルムアミドを用いることもできる。洗浄条件の有用なバリエーションは、当業者には自明であろう。ハイブリダイゼーションは、特に高ストリンジェント条件下では、ヌクレオチド間の進化的な類似性を示唆し得る。このような類似性は、ヌクレオチド及びヌクレオチドにコードされるポリペプチドに対する類似の役割を強く示唆している。

10

【0149】

「ハイブリダイゼーション複合体」の語は、相補的塩基対間の水素結合の形成力によって 2 つの核酸配列間に形成された複合体を指す。ハイブリダイゼーション複合体は、溶解状態で形成し得る (C₀t または R₀t 解析等)。或いは、一方の核酸配列が溶解状態で存在し、もう一方の核酸配列が固体支持体 (例えば紙、膜、フィルタ、チップ、ピンまたはガラススライド、或いは他の適切な基質であって細胞若しくはその核酸が固定される基質) 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。

20

【0150】

「挿入」及び「付加」の語は、1 個若しくは数個のアミノ酸残基またはヌクレオチドを各々付加するようなアミノ酸またはヌクレオチド配列における変化を指す。

【0151】

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

30

【0152】

用語「免疫原性断片」は、例えば哺乳動物などの生きている動物に導入すると、免疫反応を引き起こす G V R E D のポリペプチド断片またはオリゴペプチド断片を指す。用語「免疫原性断片」はまた、本明細書で開示するまたは当分野で周知のあらゆる抗体生産方法に有用な D M E のポリペプチド断片またはオリゴペプチド断片を含む。

【0153】

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

40

【0154】

「エレメント」または「アレイエレメント」の語は、マイクロアレイ上に定義された固有の位置にあるハイブリダイズ可能なポリヌクレオチド、ポリペプチドその他の化合物を指す。

【0155】

用語「調節」は、D M E の活性の変化を指す。例えば、調節によって、D M E のタンパク質活性、或いは結合特性、またはその他の生物学的特性、機能的特性或いは免疫学的特性の変化が起こる。

50

【0156】

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

【0157】

「機能的に結合した」は、第1核酸配列が第2核酸配列と機能的な関係があるように配置された状態を指す。例えば、プロモーターがコード配列の転写または発現に影響を及ぼす場合には、そのプロモーターはそのコード配列に機能的に結合している。機能的に結合したDNA配列は非常に近接するか、或いは連続し得る。そして、2つのタンパク質コード領域を結合する必要がある場合は、同一のリーディングフレーム内にある。

10

【0158】

「ペプチド核酸」(PNA)は、アンチセンス分子または抗遺伝子物質であって、リジンを末端とするアミノ酸残基のペプチドバックボーンに結合した、少なくとも約5ヌクレオチドの長さのオリゴヌクレオチドからなるものを指す。末端のリジンは、成分に溶解性を与える。PNAは、相補的一本鎖DNAまたはRNAに優先的に結合して転写の拡張を停止するものであり、ポリエチレングリコール化して細胞におけるPNAの寿命を延長し得る。

【0159】

DMEの「翻訳後修飾」には、脂質化、グリコシル化、リン酸化、アセチル化、ラセミ化、蛋白分解性切断及びその他の当分野で既知の修飾を含まれ得る。これらのプロセスは、合成或いは生化学的に生じ得る。生化学的修飾は、DMEの酵素環境に依存し、細胞の種類によって異なり得る。

20

【0160】

「プローブ」とは、同一配列或いは対立遺伝子核酸配列、関連する核酸配列の検出に用いる、DMEやそれらの相補配列、またはそれらの断片をコードする核酸配列のことである。プローブは、単離されたオリゴヌクレオチドまたはポリヌクレオチドであって、検出可能な標識またはレポーター分子に結合したものである。典型的な標識には、放射性アイソトープ、リガンド、化学発光試薬及び酵素がある。「プライマー」は、短い核酸、通常はDNAオリゴヌクレオチドであり、相補的塩基対を形成することで標的ポリヌクレオチドにアニーリングされ得る。プライマーは次に、DNAポリメラーゼ酵素によって標的DNA鎖に延在し得る。プライマー対は、例えばポリメラーゼ連鎖反応(PCR)による核酸配列の増幅(及び同定)に用い得る。

30

【0161】

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

40

【0162】

プローブ及びプライマーの調製及び使用方法については、Sambrook, J. ら (1989) Molecular Cloning: A Laboratory Manual, 第2版, 1-3巻, Cold Spring Harbor Press, Plainview NY、Ausubel, F.M. ら, (1987) Current Protocols in Molecular Biology, Greene Publi. Assoc. & Wiley-Intersciences, New York NY、Innisら (1990) PCR Protocols, A Guide to

50

Methods and Applications Academic Press, San Diego CA等を参照されたい。PCRプライマー対は、その目的のためのコンピュータプログラム、例えばPrimer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA)を用いるなどして既知の配列から得ることができる。

【0163】

プライマーとして用いるオリゴヌクレオチドの選択は、そのような目的のために本技術分野でよく知られているソフトウェアを用いて行う。例えばOLIGO 4.06ソフトウェアは、各100ヌクレオチドまでのPCRプライマー対の選択に有用であり、オリゴヌクレオチド及び最大5,000ヌクレオチドまでの大きめのポリヌクレオチドであって32キロベースまでのインプットポリヌクレオチド配列から得たものを分析するのにも有用である。類似のプライマー選択プログラムには、拡張能力のための追加機能が組み込まれている。例えば、PrimOUプライマー選択プログラム(テキサス州ダラスにあるテキサス大学南西部医療センターのゲノムセンターから一般向けに入手可能)は、メガベース配列から特定のプライマーを選択することが可能であり、従ってゲノム全体の範囲でプライマーを設計するのに有用である。Primer3プライマー選択プログラム(マサチューセッツ州ケンブリッジのWhitehead Institute/MITゲノム研究センターから一般向けに入手可能)ではユーザーが「ミスプライミング・ライブラリ」をインプットすることができ、ここでプライマー結合部位として避けたい配列はユーザーが指定する。Primer3は特に、マイクロアレイのためのオリゴヌクレオチドの選択に有用である。(後二者のプライマー選択プログラムのソースコードは、各自のソースから得てユーザー固有のニーズを満たすように変更してもよい。)PrimerGenプログラム(英国ケンブリッジ市の英国ヒトゲノムマッピングプロジェクト-リソースセンターから一般向けに入手可能)は、多数の配列アラインメントに基づいてプライマーを設計し、それによって、アラインメントされた核酸配列の最大保存領域または最小保存領域の何れかとハイブリダイズするようなプライマーの選択を可能にする。従って、このプログラムは、固有な、および保存されたオリゴヌクレオチド及びポリヌクレオチドの断片の同定に有用である。上記選択方法のいずれかによって同定したオリゴヌクレオチド及びポリヌクレオチドの断片は、ハイブリダイゼーション技術において、例えばPCRまたはシー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、或いは核酸のサンプルにおいて完全または部分的相補的ポリヌクレオチドを同定する特異プローブとして有用である。オリゴヌクレオチドの選択方法は、上記の方法に限定されるものではない。

【0164】

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

【0165】

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

【0166】

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

タンパク質と相互作用する。

【0167】

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

【0168】

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

10

【0169】

「サンプル」の語は、その最も広い意味で用いられる。DME、DMEをコードする核酸、またはその断片を含むと推定されるサンプルは、体液と、細胞からの抽出物や細胞から単離された染色体や細胞内小器官、膜と、細胞と、溶液中に存在するまたは基板に固定されたゲノムDNA、RNA、またはcDNAと、組織と、組織プリント等を含み得る。

【0170】

「特異結合」または「特異的に結合する」の語は、タンパク質またはペプチドと、アゴニスト、抗体、アンタゴニスト、小分子、任意の天然成分または合成結合成分との間の相互作用を指す。この相互作用は、タンパク質の特定の構造（例えば抗原決定基即ちエピトープ）であって結合分子が認識するものが存在するか否かに依存している。例えば、抗体が

20

【0171】

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

【0172】

「置換」は、1個若しくは数個のアミノ酸残基またはヌクレオチドを各々別のアミノ酸残基またはヌクレオチドに置換することを意味する。

30

【0173】

「基質」は、任意の好適な固体または半固体の支持体を指すものであって、膜、フィルタ、チップ、スライド、ウエハ、ファイバー、磁性または非磁性ビーズ、ゲル、管、プレート、ポリマー、微細粒子、毛管が含まれる。基質は、凹み、溝、ピン、チャネル、孔等、様々な表面形態を有することができ、基質表面にはポリヌクレオチドやポリペプチドが結合する。

【0174】

「転写イメージ」は、所与の時間、条件での固有の細胞タイプまたは組織による遺伝子発現の集合的パターンを指す。

40

【0175】

「形質転換」は、外来性のDNAが受入細胞に入り込むプロセスを表す。形質転換は、本技術分野で知られている種々の方法に従って自然条件または人工条件下で生じ得るものであり、外来性の核酸配列を原核または真核宿主細胞に挿入する任意の既知の方法を基にし得る。形質転換の方法は、形質転換する宿主細胞の種類によって選択する。限定するものではないが形質転換方法には、バクテリオファージまたはウイルス感染、電気穿孔法（エレクトロポレーション）、熱ショック、リポフェクション及び微粒子銃を用いる方法がある。「形質転換された」細胞には、導入されたDNAが自律的に複製するプラスミドとして或いは宿主染色体の一部として複製可能である安定的に形質転換された細胞が含まれる。さらに、限られた時間に一時的に導入DNA若しくは導入RNAを発現する細胞も

50

含まれる。

【0176】

ここで用いる「遺伝形質転換体」とは任意の有機体であり、限定するものではないが動植物を含み、有機体の1個若しくは数個の細胞が、ヒトの関与によって、例えば本技術分野でよく知られている形質転換技術によって導入された異種核酸を有する。核酸の細胞への導入は、直接または間接的に、細胞の前駆物質に導入することによって、計画的な遺伝子操作によって、例えば微量注射法によって或いは組換えウイルスの導入によって行う。遺伝子操作の語は、古典的な交雑育種或いは *in vitro* 受精を指すものではなく、組換えDNA分子の導入を指すものである。本発明に基づいて予期される遺伝形質転換体には、バクテリア、シアノバクテリア、真菌及び動植物がある。本発明の単離されたDNAは、本技術分野で知られている方法、例えば感染、形質移入、形質転換またはトランス接合によって宿主に導入することができる。本発明のDNAをこのような有機体に移入する技術はよく知られており、前出の Sambrook ら (1989) 等の参考文献に与えられている。

10

【0177】

特定の核酸配列の「変異体」は、核酸配列1本全部の長さに対して特定の核酸配列と少なくとも40%の相同性を有する核酸配列であると定義する。その際、デフォルトパラメータに設定した「BLAST 2 Sequences」ツールVersion 2.0.9 (1999年5月7日)を用いてblastnを実行する。このような核酸対は、所定の長さに対して、例えば少なくとも50%、60%、70%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の相同性を示し得る。或る変異体は、例えば「対立遺伝子」変異体(前述)、「スプライス」変異体、「種」変異体または「多型性」変異体として説明し得る。スプライス変異体は参照分子とかなりの相同性を有し得るが、mRNAプロセッシング中のエキソンの交互スプライシングによって通常多数の或いは僅かな数のポリヌクレオチドを有することになる。対応するポリペプチドは、追加機能ドメインを有するか或いは参照分子に存在するドメインが欠落していることがある。種変異体は、種相互に異なるポリヌクレオチド配列である。結果的に生じるポリペプチドは通常、相互にかなりのアミノ酸相同性を有する。多型性変異体は、与えられた種の個体間で特定の遺伝子のポリヌクレオチド配列が異なる。また、多型性変異体は、1つのヌクレオチド塩基によってポリヌクレオチド配列が変化する「1塩基多型性」(SNP)を含み得る。SNPの存在は、例えば特定の個体群、病状または病状性向を示し得る。

20

30

【0178】

特定のポリペプチド配列の「変異体」は、ポリペプチド配列の1本の長さ全体で特定のポリペプチド配列に対して少なくとも40%の相同性を有するポリペプチド配列として画定される。ここで、デフォルトパラメータに設定した「BLAST 2 Sequence」ツールVersion 2.0.9 (1999年5月7日)を用いてblastpを実行する。このようなポリペプチド対は、所定の長さに対して、例えば少なくとも50%、60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%またはそれ以上の配列同一性を示し得る。

40

【0179】

(発明)

本発明は、新規のヒト薬剤代謝酵素(DME)及びDMEをコードするポリヌクレオチドの発見に基づき、これらの組成物を利用した自己免疫/炎症の疾患、細胞増殖異常、発生または発達障害、内分泌障害、眼の疾患、代謝障害、および肝臓の疾患を含む胃腸疾患の診断、治療、及び予防に関する。

【0180】

表1は、本発明の完全長ポリヌクレオチド配列及びポリペプチド配列の命名の概略である。各ポリヌクレオチド及びその対応するポリペプチドは、1つのIncycyteプロジェクト識別番号(IncycyteプロジェクトID)と関連する。各ポリペプチド配列は、ポリ

50

ペプチド配列識別番号 (ポリペプチドSEQ ID NO) と Incyte ポリペプチド配列番号 (Incyte ポリペプチドID) によって表示した。と Incyte ポリペプチド配列番号 (Incyte ポリペプチドID) によって表示した。各ポリヌクレオチド配列は、ポリヌクレオチド配列識別番号 (ポリヌクレオチドSEQ ID NO) と Incyte ポリヌクレオチドコンセンサス配列番号 (Incyte ポリヌクレオチドID) によって表示した。と Incyte ポリヌクレオチド配列番号 (Incyte ポリヌクレオチドID) によって表示した。

【0181】

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

10

【0182】

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

20

30

【0183】

表2および表3は共に、本発明のポリペプチドの特性を要約したものであって、これらの特性は請求するポリヌクレオチドが薬剤代謝酵素であることを立証するものである。例えば、SEQ ID NO : 3は、Basic Local Alignment Search Tool (BLAST) で決定されるようにマウスチトクロームP450 (GenBank ID g2653663) と40%同一である (表2参照)。BLAST確率スコアは $5.3e-91$ 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 : 3はまた、チトクロームP450シグネチャドメインを有し、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ドメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された (表3を参照)。BLIMPS及びPROFILESCAN解析よりのデータは、SEQ ID NO : 3 がチトクロームP450のメンバーである、さらに実証的な証拠を提供する。

40

【0184】

別の例において、SEQ ID NO : 1 はアメリカ産黄色スズキ (Perca flavescens; GenBank ID g4929199) のリシルオキシダーゼと58%の同一性を有するが、これはBLAST解析によって決定された。BLAST確率スコアは $2.2e-248$ である。SEQ ID NO : 1はまた、チトクロームP450シグネチャ配列を有する。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

50

索して決定された。

【0185】

別の例において、SEQ ID NO: 2はヒトのフラビン含有モノオキシゲナーゼ5 (GenBank ID g559046) と61%の同一性を有するがこれはBLAST解析によって決定され、確率スコアは 4.5×10^{-181} であった。また、SEQ ID NO: 2はフラビン含有モノオキシゲナーゼシグネチャ配列を含むが、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。

【0186】

別の例において、SEQ ID NO: 4はPseudomonas 2, 3-ブタンジオールデヒドロゲナーゼ (GenBank ID g529564) と61%の同一性を有するがこれはBLAST解析によって決定され、確率スコアは 2.0×10^{-61} であった。SEQ ID NO: 4はまた、1つのデヒドロゲナーゼシグネチャ配列を有するが、これは、HMMを基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。BLIMPS、MOTIFS、及びPROFILES CAN解析よりのデータは、SEQ ID NO: 4がデヒドロゲナーゼであるという、さらに確証的な証拠を提供する。

10

【0187】

SEQ ID NO: 5はBacillus キノンオキシダーゼ (GenBank ID g2633069) と54%の同一性を有するがこれはBLAST解析によって決定され、確率スコアは 7.1×10^{-96} であった。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得られたデータにより、SEQ ID NO: 5がキノンオキシダーゼであるさらに確証的な証拠が得られた。

20

【0188】

別の例において、SEQ ID NO: 6はマウスのヘパラン硫酸6スルフトランスフェラーゼ (GenBank ID g6683558) と92%の同一性を有するがこれはBLAST解析によって決定され、確率スコアは 2.3×10^{-255} であった。

【0189】

別の例において、SEQ ID NO: 7はヒトのグルタチオンSトランスフェラーゼ (GenBank ID g242749) と90%の同一性を有するがこれはBLAST解析によって決定され、確率スコアは 1.3×10^{-101} であった。また、SEQ ID NO: 7はグルタチオンSトランスフェラーゼシグネチャ配列を含むが、これは、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、またBLIMPS解析を行うことにより決定された。

30

【0190】

別の例において、SEQ ID NO: 8はヒトのステロイドデヒドロゲナーゼ (GenBank ID g5531815) と40%の同一性を有するがこれはBLAST解析によって決定され、確率スコアは 1.9×10^{-56} であった。SEQ ID NO: 8はまた、1つのデヒドロゲナーゼシグネチャ配列を有するが、これは、HMMを基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。

40

【0191】

SEQ ID NO: 9はウサギの肝臓カルボキシエステラーゼ (GenBank ID g3219695) と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 6.3×10^{-72} であり、これは観察されたポリペプチド配列アライメントが偶然に得られる確率を示している。SEQ ID NO: 9はまた、カルボキシエステラーゼドメインを有するが、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意

50

な一致を検索して決定された。(表3参照)BLIMPS、MOTIFS、及びPROFILESCAN解析よりのデータは、SEQ ID NO: 9 がカルボキシエステラーゼであるという、さらに確証的な証拠を提供する。

【0192】

SEQ ID NO: 10 はヒトのカルボキシルエステラーゼ (GenBank ID g180950) と45%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 (BLAST、表2を参照)。BLAST確率スコアは $8.7e-130$ 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: 10 はまた、カルボキシルエステラーゼドメインを有するが、これは、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された (表3を参照)。BLIMPS、MOTIFS、及びPROFILESCAN解析よりのデータは、SEQ ID NO: 10 がカルボキシエステラーゼであるという、さらに確証的な証拠を提供する。

10

【0193】

別の例において、SEQ ID NO: 11 はBasic Local Alignment Search Tool (BLAST) によって同定されるようにマウスのヘパラン硫酸 6-O-スルホトランスフェラーゼ2 前駆体 (GenBank ID g6683558) に89%同一である (BLAST、表2参照)。BLAST確率スコアは $1.8e-236$ 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。そしてまた、SEQ ID NO: 11 がDMEであり、具体的にはSEQ ID NO: 11 がスルホトランスフェラーゼである証拠を提供する。

20

【0194】

別の例として、SEQ ID NO: 12 は Bacillus subtilis エポキシドヒドラーゼ (GenBank ID g2633182) と25%の同一性を有す、これはBasic Local Alignment Search Tool (BLAST) によって確認された (BLAST、表2を参照)。BLAST確率スコアは $4.0e-11$ であり、これは観測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: 12 はまた、ヒドラーゼドメインを有するが、これは、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された (表3を参照)。BLIMPS解析よりのデータは、SEQ ID NO: 12 がヒドラーゼである、さらに実証的な証拠を提供する。

30

【0195】

別の例において、SEQ ID NO: 13 はラットのベータアラニンピルビン酸アミノトランスフェラーゼ (GenBank ID g1944136) と83%の同一性を有しているが、これはBLAST解析によって決定された (表2を参照)。BLAST確率スコアは $1.1e-234$ である。SEQ ID NO: 13 はまた、アミノトランスフェラーゼドメインを有するが、これは、隠れマルコフモデル (HMM) を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された (表3を参照)。BLIMPS及びPROFILESCAN解析よりのデータは、SEQ ID NO: 13 がアミノトランスフェラーゼであることをさらに確証する証拠を提供する。別の例において、SEQ ID NO: 14 はモルモットのヒドロキシステロイドトランスフェラーゼ (GenBank ID g1151081) と50%の同一性を有しているが、これはBLAST解析によって決定された (表2を参照)。BLAST確率スコアは $5.4e-34$ であり、これはSEQ ID NO: 14 がスルホトランスフェラーゼであるという証拠を提供する。

40

【0196】

別の例において、SEQ ID NO: 15 はモルモットの銅ノ亜鉛スーパーオキシドジスムターゼ (GenBank ID g1066120) と52%の同一性を有しているが

50

、これはBLAST解析によって決定された(表2を参照)。BLAST確率スコアは 2.1×10^{-25} である。SEQ ID NO: 15はまた、銅/亜鉛スーパーオキシドジスムターゼドメインを有する。これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。BLAST解析から得られたデータによって、さらにSEQ ID NO: 15が銅/亜鉛スーパーオキシドジスムターゼであることが裏付けられた。

【0197】

SEQ ID NO: 16はヒトの3'ホスホアデニリル硫酸ガラクトシルセラミド3'スルホトランスフェラーゼ(セレプロシドスルホトランスフェラーゼ、GenBank ID g1871141)と37%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(BLAST、表2を参照)。BLAST確率スコアは 2.8×10^{-60} 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。

10

【0198】

別の例において、SEQ ID NO: 17は線虫(*C. elegans*)のモノアミンオキシダーゼ(GenBank ID g6782275)と38%の同一性を有するがこれはBLAST解析によって決定され、確率スコアは 3.0×10^{-99} であった。SEQ ID NO: 17はまた、モノアミンオキシダーゼドメインを有するが、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された。(表3参照)BLIMPS解析から得られたデータによって、さらにSEQ ID NO: 17がモノアミンオキシダーゼであることが裏付けられた。

20

【0199】

別の例において、SEQ ID NO: 18はヒトのカテコールOメチルトランスフェラーゼ(GenBank ID g179955)と36%の同一性を有するがこれはBLAST解析によって決定され、確率スコアは 9.5×10^{-41} であった。SEQ ID NO: 18はマウスのカテコールOメチルトランスフェラーゼ(GenBank ID g3493253)と36%の同一性を有するがこれはBLAST解析によって決定され、確率スコアは 1.3×10^{-41} であった。

【0200】

別の例において、Basic Local Alignment Search Tool (BLAST)によって決定されるようにSEQ ID NO: 19はトマミチヨグ(*Fundulus heteroclitus*)のチトクロームP450 2N1(GenBank ID g5852342)に44%同一である(BLAST、表2参照)。BLAST確率スコアは 4.6×10^{-99} 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: 19はまた、チトクロームP450ドメインを有するが、これは、隠れマルコフモデル(HMM)を基にした保存されたタンパク質ファミリードメインのPFAMデータベースにおいて、統計的に有意な一致を検索して決定された(表3を参照)。BLIMPS、MOTIFS、およびPROFILES CAN解析から得られたデータによって、SEQ ID NO: 19がチトクロームP450であることが裏付けられた。

30

40

【0201】

SEQ ID NO: 1-19の解析のためのアルゴリズム及びパラメータが表7で記述されている。

【0202】

表4に示すように、本発明の完全長ポリヌクレオチド配列は、cDNA配列またはゲノムDNA由来のコード(エキソン)配列を用いて、或いはこれら2種類の配列を任意に組み合わせ構築した。列1および列2はそれぞれ、本発明の各ポリヌクレオチドのポリヌクレオチド配列識別番号(Polynucleotide SEQ ID NO:)およびそれに対応するIncyteポリヌクレオチドコンセンサス配列番号(Incyte Po

50

lynucleotide ID)を示す。列3は、各ポリヌクレオチド配列の長さ(塩基対単位)を示す。列4は、例えば、SEQ ID NO: 20-38を同定するため、或いはSEQ ID NO: 20-38と関連するポリヌクレオチド配列とを区別するためのハイブリダイゼーションまたは増幅技術に有用なポリヌクレオチド配列の断片を示す。列5はcDNA配列、ゲノムDNAから予想されたコード配列(エキソン)及び/またはcDNA及びゲノムDNAを共に有する配列集合に対応する識別番号を示している。これらの配列は、本発明の完全長ポリヌクレオチド配列を構築するのに用いた。表4の列6および列7はそれぞれ完全長配列に対して、列5の配列に対応するcDNA配列およびゲノム配列の開始ヌクレオチド(5')位置および終了ヌクレオチド(3')位置を示す。

【0203】

表4の列5の識別番号は、特に例えばIncyte cDNAとそれに対応するcDNAライブラリを指す場合もある。例えば、7690384J1はインサイトcDNA配列の識別番号であり、PROSTME06はそれが由来するcDNAライブラリの識別番号である。cDNAライブラリが示されていないIncyte cDNAは、プールされているcDNAライブラリ(例えば、55017748J1)に由来する。または、列5の識別番号は、完全長ポリヌクレオチド配列の組み立てに用いたGenBankのcDNAすなわちEST(例えば、g1203094)の識別番号の場合もある。また、列5の識別番号は、ENSEMBL(The Sanger Centre, Cambridge, UK)データベースに由来する配列(すなわち、名前に「ENST」を含む配列)の識別番号である場合もある。また、列5の識別番号は、NCBI RefSeqヌクレオチド配列記録データベースに由来する配列(すなわち、名前に「NM」または「NT」を含む配列)あるいはNCBI RefSeqタンパク質配列記録に由来する配列(すなわち、名前に「NP」を含む配列)の識別番号である場合もある。または列5の識別番号は、「エキソンスティック(exon-stitching)」アルゴリズムにより結び合わせたcDNA及びGenscan予想エキソンの両方の集合を指す場合もある。例えば、FL__XXXXXX__N₁__N₂__YYYYY__N₃__N₄はアルゴリズムが適用される配列のクラスターの識別番号がXXXXXXXXであり、アルゴリズムにより生成される予測の番号がYYYYYであり、(もし存在すれば)N₁, N₂, N₃が解析中に手動で編集された可能性のある特定のエキソンであるような「縫合された」配列である(実施例5参照)。または列5の識別番号は、「エキソストレッチ(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、FLXXXXXXXX_gAAAAA_gBBBBB__1__Nは「ストレッチ」配列の識別番号である。ここでXXXXXXXXはIncyteプロジェクト識別番号、gAAAAAは「エキソストレッチ」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識別番号である。(実施例5を参照。)RefSeq配列が「エキソストレッチ」アルゴリズム用のタンパク質相同体として津筒されているときは、RefSeq識別子(「NM」、「NP」、または「NT」で表示)をGenBank識別し(たとえばgBBBBB)の代わりに使うことができる。

【0204】

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

10

20

30

40

接頭コード	解析タイプおよびまたはプログラム例
GNJ CFG ENST	たとえばGENSCAN(Stanford University CA USA) またはGENES (Computer Graphics Group, The Sanger Centre, Cambridge UK) を使ったゲノムからのエキソン予測
GH	ゲノム配列の自動断片化
EL	縫合 (スッチ) またはストレッチされたゲノム配列の断片化を参照。
INCY	ゲノムのEST配列のクランプ由来の完全転写およびエキソン予測。ゲノム位置とEST組織データを結合してエキソンと生体転写を推定する。

10

20

【0205】

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

30

【0206】

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

【0207】

本発明はまた、DME の変異体も含む。好適な DME の変異体は、DME の機能的或いは構造的特徴の少なくともどちらか一方を有し、かつ DME 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。

40

【0208】

本発明はまた、DME をコードするポリヌクレオチドを提供する。特定の実施例において、本発明は、DME をコードする SEQ ID NO: 20 - 38 からなる一群から選択された配列を含むポリヌクレオチド配列を提供する。SEQ NO ID: 20 - 38 のポリヌクレオチド配列は、配列表に示されているように等価 RNA 配列と同等の価値を有しているが、窒素塩基チミンの出現はウラシルに置換され、糖のバックボーンはデオキシリボースではなくシリボースから構成されている。

50

【0209】

本発明はまた、DMEをコードするポリヌクレオチド配列の変異配列を含む。詳細には、このようなポリヌクレオチド配列の変異配列は、DMEをコードするポリヌクレオチド配列に対して少なくとも70%のポリヌクレオチド配列同一性、或いは少なくとも85%のポリヌクレオチド配列同一性、更には少なくとも95%ものポリヌクレオチド配列同一性を有する。本発明の特定の実施形態は、SEQ ID NO: 20 - 38からなる一群から選択された核酸配列と少なくとも70%のポリヌクレオチド配列同一性、或いは少なくとも85%のポリヌクレオチド配列同一性、更には少なくとも95%ものポリヌクレオチド配列同一性を有するSEQ ID NO: 20 - 38からなる一群から選択された配列を含むポリヌクレオチド配列の変異配列を提供する。上記したポリヌクレオチド変異配列は 10
何れも、DMEの機能的或いは構造的特徴の少なくとも1つを有するアミノ酸配列をコードする。

【0210】

遺伝暗号の縮重により作り出され得るDMEをコードする種々のポリヌクレオチド配列には、既知の自然発生する任意の遺伝子のポリヌクレオチド配列と最小の類似性しか有しないものも含まれることを、当業者は理解するであろう。したがって本発明には、可能コドン選択に基づく組合せの選択によって産出し得るようなありとあらゆる可能性のあるポリヌクレオチド配列変異体を網羅し得る。これらの組み合わせは、天然のDMEのポリヌクレオチド配列に適用される標準的なトリプレット遺伝暗号を基に作られ、全ての変異が明確に開示されていると考慮する。 20

【0211】

DMEをコードするヌクレオチド配列及びその変異配列は一般に、好適に選択されたストリンジェントな条件下で、天然のDMEのヌクレオチド配列とハイブリダイズ可能であるが、非天然のコドンを含めるなどの実質的に異なった使い方のコドンを有するDME或いはその誘導体をコードするヌクレオチド配列を作るとは有利となり得る。宿主が特定のコドンを利用する頻度に基づいて、特定の真核又は原核宿主に発生するペプチドの発現率を高めるようにコドンを選択することが可能である。コードされたアミノ酸配列を変えずに、DME及びその誘導体をコードするヌクレオチド配列を実質的に変更する別の理由は、天然の配列から作られる転写物より例えば長い半減期など好ましい特性を備えるRNA転写物を作ることにある。 30

【0212】

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

【0213】

更に本発明には、種々のストリンジェントな条件下で、請求項に記載されたポリヌクレオチド配列、特に、SEQ ID NO: 20 - 38及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる(例えば、Wahl, G. M. 及びS. L. Berger (1987) *Methods Enzymol.* 152: 399 - 407、Kimmel, A. R. (1987) *Methods Enzymol.* 152: 507 - 511等を参照)。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。 40

【0214】

DNAシーケンシングの方法は当分野でよく知られており、本発明の何れの実施例もDNAシーケンシング方法を用いて実施可能である。DNAシーケンシング方法には酵素を用いることができ、例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amersham 50

am, Pharmacia Biotech, Piscataway NJ)を用いることができる。或いは、例えばELONGASE増幅システム(Life Technologies, Gaithersburg MD)において見られるように、ポリメラーゼと校正エキソヌクレアーゼを併用することができる。好適には、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)及びABI CATALYST 800サーマルサイクラー(Applied Biosystems)等の装置を用いて配列の調製を自動化する。次に、ABI 373 或いは 377 DNAシー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)または当分野でよく知られている他の方法を用いてシーケンシングを行う。結果として得られた配列を当分野でよく知られている種々のアルゴリズムを用いて分析する。(Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7, Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, 856 - 853ページ.等を参照)。

【0215】

当分野で周知のPCR法をベースにした種々の方法で、部分的なヌクレオチド配列を利用して、DMEをコードする核酸配列を伸長し、プロモーターや調節エレメントなどの上流にある配列を検出する。例えば、使用し得る方法の1つである制限部位PCR法は、ユニバーサルプライマー及びネステッドプライマーを用いてクローニングベクター内のゲノムDNAから未知の配列を増幅する方法である(例えば、Sarkar, G. (1993) PCR Methods Applic. 2: 318322を参照。)別の方法に逆PCR法があり、これは広範な方向に伸長させたプライマーを用いて環状化した鋳型から未知の配列を増幅する方法である。鋳型は、既知のゲノム遺伝子座及びその周辺の配列を含む制限酵素断片から得る(例えば、Triglia, T. 他(1988) Nucleic Acids Res. 16: 8186。)第3の方法としてキャプチャPCR法があり、これはヒト及び酵母菌人工染色体DNAの既知の配列に隣接するDNA断片をPCR増幅する方法に参与している。(Lagerstrom, M.ら(1991) PCR Methods Applic 1: 111-119等を参照)。この方法では、PCRを行う前に複数の制限酵素の消化及び連結反応を用いて未知の配列領域内に組換え二本鎖配列を挿入することが可能である。また、未知の配列を検索するために用い得る別の方法については当分野で知られている。(Parker, J.D.ら(1991) Nucleic Acids Res. 19: 3055-3060等を参照)。更に、PCR、ネステッドプライマー及びPromoter Finder(商標)ライブラリ(Clonotech, Palo Alto CA)を用いてゲノムDNAをウォーキングすることができる。この手順は、ライブラリをスクリーニングする必要がなく、イントロン/エキソン接合部を見付けるのに有用である。全てのPCRベースの方法に対して、市販されているソフトウェア、例えばOLIGO 4.06プライマー分析ソフトウェア(National Biosciences, Plymouth MN)或いは別の好適なプログラムを用いて、長さが約22~30ヌクレオチド、GC含有率が約50%以上、温度約68~72で鋳型に対してアニーリングするようにプライマーを設計し得る。

【0216】

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

【0217】

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

10

【0218】

本発明の別の実施例では、DME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DME、その断片または機能的等価物を発現させることが可能である。遺伝暗号固有の縮重により、実質的に同じ或いは機能的に等価のアミノ酸配列をコードする別のDNA配列が作られ得り、これらの配列をDMEのクローン化及び発現に利用可能である。

【0219】

種々の目的でDMEをコードする配列を変えるために、当分野で一般的に知られている方法を用いて、本発明のヌクレオチド配列を組換えることができる。この目的には、遺伝子産物のクローン化、プロセッシング及び/または発現の調節が含まれるが、これらに限定されるものではない。遺伝子断片及び合成オリゴヌクレオチドのランダムなフラグメンテーション及びPCR再アセンブリによるDNAシャッフリングを用い、ヌクレオチド配列を組み換えることが可能である。例えば、オリゴヌクレオチド仲介特定部位突然変異誘導を利用して、新規な制限部位の生成、グリコシル化パターンの変更、コドン優先の変更、スプライス変異体の生成等を行う突然変異を導入し得る。

20

【0220】

本発明のヌクレオチドを、MOLECULAR BREEDING (Maxygen Inc., Santa Clara CA; 米国特許第5,837,458号; Chang, C.-C. 他(1999) Nat. Biotechnol. 17:793-797; Christians, F.C. 他(1999) Nat. Biotechnol. 17:259-264; Cramer, A. 他(1996) Nat. Biotechnol. 14:315-319)などのDNAシャッフリング技術を用いてシャッフリングして、DMEの生物学的または酵素的な活性、或いは他の分子や化合物と結合する能力などのDMEの生物学的特性を変更或いは改良することができる。DNAシャッフリングは、遺伝子断片のPCR仲介再組換えを用いて遺伝子変異体のライブラリを生成するプロセスである。ライブラリはその後、その遺伝子変異体を所望の特性に同定するような選択またはスクリーニングにかける。次にこれらの好適な変異体をプールし、更に反復してDNAシャッフリング及び選択/スクリーニングを行ってもよい。このように、遺伝の多様性は「人為的」品種改良及び急速な分子の進化を経て創生される。例えば、ランダムポイント突然変異を有する単一の遺伝子の断片を組み換えて、スクリーニングし、その後所望の特性が最適化されるまでシャッフリングすることができる。或いは、所定の遺伝子の断片を同種または異種のいずれかから得た同一遺伝子ファミリーの相同遺伝子と組み換えて、それによって天然に存在する複数の遺伝子の遺伝多様性を、指図された制御可能な方法で最大化させることができる。

30

40

【0221】

別の実施例によれば、DMEをコードする配列は、当分野で周知の化学的方法を用いて、全体或いは一部が合成可能である(例えば、Caruthers, M.H.ら(1980) Nucl. Acids Res. Symp. Ser. 7:215-223; 及びHorn, T. 他(1980) Nucl. Acids Res. Symp. Ser. 2:25-232を参照)。別法として、化学的方法を用いてDME自体またはその断片を合

50

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

10

【0222】

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

【0223】

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

20

30

【0224】

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

40

種々の発現ベクター/宿主系を利用して、DMEをコードする配列の保持及び発現が可能である。限定するものではないがこのような発現ベクター/宿主系には、組換えバクテリオファージ、プラスミドまたはコスミドDNA発現ベクターで形質転換させた細菌や、酵母菌発現ベクターで形質転換させた酵母菌や、ウイルス発現ベクター(例えばバキュロウイルス)に感染した昆虫細胞系や、ウイルス発現ベクター(例えばカリフラワーモザイクウイルス、CaMVまたはタバコモザイクウイルスTMV)または細菌発現ベクター(例えばTiまたはpBR322プラスミド)で形質転換させた植物細胞系、動物細胞系など

50

の微生物等がある。(前出のSambrook、前出のAusubel、Van Heeke, G. および S. M. Schuster (1989) J. Biol. Chem. 264:5503-5509、; Engelhard, E. K. ら (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227、Sandig, V. ら (1996) Hum. Gene Ther. 7:1937-1945、Takamatsu, N. (1987) EMBOJ. 6:307-311、; 『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, 191-196ページ、Logan, J. および T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659、Harrington, J. J. ら (1997) Nat. Genet. 15:345-355等を参照)レトロウイルス、アデノウイルス、ヘルペスウイルスまたはワクシニアウイルス由来の発現ベクター、または種々の細菌性プラスミド由来の発現ベクターを用いて、ヌクレオチド配列を標的器官、組織または細胞集団へ輸送することができる(Di Nicola, M. ら (1998) Cancer Gen. Ther. 5(6):350-356、Yu, M. ら (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344、Buller, R. M. ら (1985) Nature 317(6040):813-815; McGregor, D. P. ら (1994) Mol. Immunol. 31(3):219-226、Verma, I. M. および N. Somia (1997) Nature 389:239-242等を参照)。本発明は使用される宿主細胞によって限定されるものではない。

【0225】

細菌系では、多数のクローニングベクター及び発現ベクターが、DMEをコードするポリヌクレオチド配列の使用目的に応じて選択可能である。例えば、DMEをコードするポリヌクレオチド配列の日常的なクローニング、サブクローニング、増殖には、PBLUESCRIPT(Stratagene, La Jolla CA)またはpSPORT1プラスミド(GIBCO BRL)などの多機能の大腸菌ベクターを用いることができる。ベクターの多数のクローニング部位にDMEをコードする配列をライゲーションするとlacZ遺伝子が破壊され、組換え分子を含む形質転換された細菌の同定のための比色スクリーニング法が可能となる。更にこれらのベクターは、クローニングされた配列における*in vitro*転写、ジデオキシのシーケンシング、ヘルパーファージによる一本鎖の救出、入れ子状態の欠失の生成にも有用であろう(例えば、Van Heeke, G. および S. M. Schuster (1989) J. Biol. Chem. 264:5503-5509を参照)。例えば、抗体の産生のためなどに多量のDMEが必要な場合は、DMEの発現をハイレベルで誘導するベクターが使用できる。例えば、強力な誘導SP6バクテリオファージプロモーターまたは誘導T7バクテリオファージプロモーターを含むベクターが使用できる。

【0226】

DMEの発現に酵母の発現系の使用が可能である。因子、アルコールオキシダーゼ、PGHプロモーター等の構成型或いは誘導型のプロモーターを含む多数のベクターが、酵母菌サッカロミセス-セレビジエまたは*Pichia pastoris*に使用可能である。更に、このようなベクターは、発現したタンパク質の分泌か細胞内への保持のどちらかを誘導し、安定した増殖のために宿主ゲノムの中に外来配列を組み込む。(例えば、Ausubel, 1995, 前出、Bitter, G. A. ら (1987) Methods Enzymol. 153:516-544、及びScorer, C. A. ら (1994) Bio/Technology 12:1-181-184.を参照)。

【0227】

植物系もDMEの発現に使用可能である。DMEをコードする配列の転写は、ウイルスプロモーター、例えば単独或いはTMV(Takamatsu, N. (1987) EM 50

BOJ 6:307-311)由来のオメガリーダー配列と組み合わせて用いられるようなCaMV由来の35S及び19Sプロモーターによって促進される。或いは、RUBISCOの小サブユニット等の植物プロモーターまたは熱ショックプロモーターを用いてもよい(例えば、Coruzzi, G. ら (1984) EMBO J. 3:1671-1680; Broglie, R. ら (1984) Science 224:838-843; および Winter, J. ら (1991) Results Probl. Cell Differ. 17:85-105を参照)これらの構成物は、直接DNA形質転換または病原体を媒介とする形質移入によって、植物細胞内に導入可能である。(『マグローヒル科学技術年鑑』(The McGraw Hill Yearbook of Science and Technology) (1992) McGraw Hill New York NY, pp.191-196等を参照。 10

【0228】

哺乳動物細胞においては、多数のウイルスベースの発現系を利用し得る。アデノウイルスが発現ベクターとして用いられる場合、後発プロモーター及び3連リーダー配列からなるアデノウイルス転写物/翻訳複合体にDMEをコードする配列を結合し得る。可欠E1またはE3領域へウイルスのゲノムを挿入し、宿主細胞でDMEを発現する感染ウイルスを得ることが可能である。(例えば、Logan, J. および T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659を参照)。更に、ラウス肉腫ウイルス(RSV)エンハンサー等の転写エンハンサーを用いて、哺乳動物宿主細胞における発現を増大させ得る。SV40またはEBVをベースにしたベクターを用いてタンパク質を高レベルで発現させることもできる。 20

【0229】

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

【0230】

哺乳動物系の組換えタンパク質の長期にわたる産生のためには、株化細胞におけるDMEの安定した発現が望ましい。例えば、発現ベクターを用いて、DMEをコードする配列を株化細胞に形質転換することが可能である。このような発現ベクターは、ウイルス起源の複製及び/または内在性の発現要素や、同じ或いは別のベクター上の選択マーカー遺伝子を含む。ベクターの導入後、選択培地に移す前に強化培地で約1~2日間細胞を増殖させることができる。選択可能マーカーの目的は選択培地への抵抗性を与えることであり、選択可能マーカーが存在することにより、導入された配列をうまく発現するような細胞の成長及び回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に適した組織培養技術を用いて増殖可能である。 30

【0231】

任意の数の選択系を用いて、形質転換細胞株を回収できる。限定するものではないがこのような選択系には、tk⁻単細胞のために用いられるヘルペスウイルスチミジンキナーゼ遺伝子と、apr⁻細胞のために用いられるアデニンホスホリボシルトランスフェラーゼ遺伝子がある(例えば、Wigler, M. 他 (1977) Cell 11:223-232; 及び Lowy, I. 他 (1980) Cell 22:817-823を参照)。また、選択の基礎として代謝拮抗物質、抗生物質或いは除草剤への耐性を用いることができる。例えばdhfrはメトトレキセートに対する耐性を与え、neoはアミノグリコシッドネオマイシン及びG-418に対する耐性を与え、alsはクロルスルフロンに対する耐性を、patはホスフィノトリシンアセチルトランスフェラーゼに対する耐性を各々与える(Wigler, M. ら (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere Garapin, F. ら (1981) J. Mol. Biol. 150:114等を参照。)この他 40 50

の選択可能な遺伝子、例えば、代謝のための細胞の必要条件を変える *trpB* 及び *hisD* は、文献に記載されている (Hartman, S. C. および R. C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85: 80478051 を参照。) 可視マーカー、例えばアントシアニン、緑色蛍光タンパク質 (GFP; Clontech)、グルクロニダーゼ及びその基質 グルクロニド、またはルシフェラーゼ及びその基質ルシフェリン等を用いてもよい。これらのマーカーを用いて、トランスフォーマントを特定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することが可能である (Rhodes, C. A. (1995) Methods Mol. Biol. 55: 121131 等を参照)。

【0232】

マーカー遺伝子発現の存在 / 不存在によって目的の遺伝子の存在が示唆されても、その遺伝子の存在及び発現の確認が必要な場合もある。例えば、DME をコードする配列がマーカー遺伝子配列の中に挿入された場合、DME をコードする配列を含む形質転換された細胞は、マーカー遺伝子機能の欠落により同定可能である。または、1つのプロモーターの制御下でマーカー遺伝子が DME をコードする配列と一列に配置することも可能である。誘導または選択に応答したマーカー遺伝子の発現は通常、タンデム遺伝子の発現も示す。

【0233】

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

【0234】

特異的なポリクローナル抗体またはモノクローナル抗体のどちらかを用いる DME の発現の検出及び計測のための免疫学的方法は、当分野で周知である。このような技法には、酵素に結合したイムノソルベントアッセイ (ELISA)、ラジオイムノアッセイ (RIA)、蛍光標識式細胞分取器 (FACS) などがある。DME 上の 2 つの非干渉エピトープに反応するモノクローナル抗体を用いた、2 部位のモノクローナルベースイムノアッセイ (two-site, monoclonal-based immunoassay) が好ましいが、競合の結合アッセイも用いることもできる。これらのアッセイ及びこれ以外のアッセイは、当分野で公知である (Hampton, R. ら (1990) Serological Methods, a Laboratory Manual, APS Press, St Paul, MN, Sect. IV、Coligan, J. E. ら (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY、Pound, J. D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ 等を参照)。

【0235】

多岐にわたる標識方法及び結合方法が、当業者に知られており、様々な核酸アッセイおよびアミノ酸アッセイにこれらの方法を用い得る。DME をコードするポリヌクレオチドに関連する配列を検出するための、標識されたハイブリダイゼーションプローブ或いは PCR プローブを生成する方法には、オリゴ標識化、ニックトランスレーション、末端標識化、または標識されたヌクレオチドを用いる PCR 増幅が含まれる。別法として、DME をコードする配列、またはその任意の断片を mRNA プローブを生成するためのベクターにクローニングすることも可能である。このようなベクターは、当分野において知られており、市販もされており、T7、T3 または SP6 等の好適な RNA ポリメラーゼ及び標識されたヌクレオチドを加えて、in vitro で RNA プローブの合成に用いることができる。このような方法は、例えば Amersham Pharmacia Biotech

10

20

30

40

50

h、Promega (Madison WI)、U.S. Biochemical等から市販されている種々のキットを用いて実行することができる。検出を容易にするために用い得る好適なレポーター分子或いは標識には、基質、補助因子、インヒビター、磁気粒子のほか、放射性核種、酵素、蛍光剤、化学発光剤、発色剤等がある。

【0236】

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

【0237】

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

【0238】

本発明の別の実施例では、DMEをコードする自然或いは変更された、または組換えの核酸配列を上記した任意の宿主系の融合タンパク質の翻訳となる異種配列に結合させる。例えば、市販の抗体によって認識できる異種部分を含むキメラDMEタンパク質が、DME活性のインヒビターに対するペプチドライブラリのスクリーニングを促進し得る。また、異種タンパク質部分及び異種ペプチド部分も、市販されている親和性基質を用いて融合タンパク質の精製を促進し得る。限定されるものではないがこのような部分には、グルタチオンSトランスフェラーゼ(GST)、マルトース結合タンパク質(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及び赤血球凝集素(HA)は、これらのエピトープ標識を特異的に認識する市販されているモノクローナル抗体及びポリクローナル抗体を用いて、融合タンパク質の免疫親和性精製を可能にする。また、DMEをコードする配列と異種タンパク質配列との間にあるタンパク質分解切断部位を融合タンパク質が含むように遺伝子操作すると、DMEが精製の後に異種部分から切断され得る。融合タンパク質の発現及び精製方法は、前出のAusubel(1995)10章に記載されている。市販されている種々のキットを用いて融合タンパク質の発現及び精製を促進することもできる。

【0239】

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DMEの合成が可能である。これらの系は、T7、T3またはSP6プロモーターと機能的に結合したタンパク質コード配列の転写及び翻訳を結合する。翻訳は、例えば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。

【0240】

本発明のDMEまたはその断片を用いて、DME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、DMEへの特異的な結合をスクリーニングすることが可能である。試験化合物の例には、抗体、オリゴヌクレオチド、タンパク質（例えば受容体）または小分子が挙げられる。

【0241】

一実施例では、このように同定された化合物は、例えばリガンドやその断片などのDMEの天然のリガンド、または天然の基質、構造的または機能的な擬態性または自然結合パートナーに密接に関連している（Coligan, J. E. 他（1991）Current Protocols in Immunology 1（2）の5章等を参照）。（Coligan, J. E. 他（1991）Current Protocols in Immunology 1（2）の5章等を参照）。同様に、化合物は、DMEが結合する天然受容体、或いは例えばリガンド結合部位などの少なくとも受容体のある断片に密接に関連し得る。何れの場合も、既知の技術を用いてこの化合物を合理的に設計することができる。一実施例では、これらの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DMEを発現する好適な細胞の作製が含まれる。好適な細胞には、哺乳動物、酵母、ショウジョウバエ、大腸菌からの細胞が含まれる。DMEを発現する細胞またはDMEを含有する細胞膜断片を試験化合物と接触させて、DMEまたは化合物の何れかの結合、刺激または阻害を分析する。

10

【0242】

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

20

【0243】

本発明のDMEまたはその断片を用いて、DMEの活性を調整する化合物をスクリーニングすることが可能である。このような化合物には、アゴニスト、アンタゴニスト、或いは部分的または逆アゴニスト等が含まれる。一実施例では、DMEが少なくとも1つの試験化合物と結合する、DME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DMEの活性を試験化合物不在下でのDMEの活性と比較する。試験化合物の存在下でのDMEの活性の変化は、DMEの活性を調整する化合物の存在を示唆する。別法では、試験化合物をDMEの活性に適した条件下でDMEを含む*in vitro*または無細胞系と結合させてアッセイを実施する。これらアッセイの何れにおいても、DMEの活性を調節する試験化合物は間接的に結合することが可能であり、試験化合物と直接接触する必要がない。少なくとも1つから複数の試験化合物をスクリーニングすることができる。

30

【0244】

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

40

50

est. 97:1999-2002; Wagner, K.U. 他 (1997) Nucleic Acids Res. 25:4323-4330)。形質転換したES細胞を同定し、例えばC57BL/6マウス系等から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を決め、これを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、可能性のある治療薬や毒性薬剤で検査することができる。

【0245】

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

10

【0246】

DMEをコードするポリヌクレオチドを用いて、ヒト疾患をモデルとした「ノックイン」ヒト化動物(ブタ)または遺伝子組換え動物(マウスまたはラット)を作製することが可能である。ノックイン技術を用いて、DME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。形質転換細胞を胚に注入し、胚を上記のように移植する。遺伝子組換え子孫または近交系について研究し、可能性のある医薬品を用いて処理し、ヒトの疾患の治療に関する情報を得る。別法では、例えばDMEを乳汁内に分泌するなどDMEを過剰発現する哺乳動物近交系は、便利なたんぱく質源となり得る (Janne, J. 他 (1998) Biotechnol. Annu. Rev. 4:55-74)。

20

【0247】

(治療)

DMEのある領域と薬剤代謝酵素のある領域との間に、例えば配列及びモチーフのからみで化学的及び構造的類似性が存在する。さらに、DMEの発現は、脳、前立腺、骨、腸、乳房などのさまざまな病変組織に密接に関連している。従って、DMEは、自己免疫/炎症の疾患、細胞増殖異常、発生または発達障害、内分泌障害、眼の疾患、代謝障害、および肝臓の疾患を含む胃腸疾患においてある役割を果たすと考えられる。DMEの発現若しくは活性亢進に関連する疾患の治療においては、DMEの発現または活性を低下させることが望ましい。また、DME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、DMEの発現または活性を亢進させることが望ましい。

30

【0248】

従って、一実施例において、DMEの発現または活性の低下に関連した疾患の治療または予防のために、患者にDMEまたはその断片や誘導体を投与することが可能である。限定するものではないが、このような疾患には自己免疫/炎症の疾患、細胞増殖異常、発生または発達障害、内分泌障害、眼の疾患、代謝障害、および肝臓の疾患を含む胃腸疾患が含まれ、自己免疫/炎症の疾患の中には、炎症及び日光性角化症、後天性免疫不全症候群(AIDS)及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時的リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、腓炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬

40

50

変、肝炎、混合型結合組織病 (MCTD)、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌が含まれ、発生または発達障害の中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、精神薄弱)、スミス マジェニス症候群 (Smith-Magenis syndrome)、脊髄形成異常症候群、遺伝性粘膜上皮異形成、遺伝性角皮症、シャルコー マリー ツース病及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、Syndenham 舞踏病 (Syndenham's chorea) 及び脳性小児麻痺などの発作障害、脊髄二分裂、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感覚神経性聴力損失が含まれ、内分泌障害の中には原発脳腫瘍及び腺腫、妊娠性梗塞、下垂体切除、動脈瘤、血管奇形、血栓症、感染症、免疫異常、頭部外傷による合併症などの病変から起こる視床下部及び下垂体の障害と、性機能低下及びシーハン症候群、尿崩症、カルマン病、ハンド シュラー クリスマン病、レトラ シヴェ病、サルコイドーシス、エンプレセラ症候群、小人症を含む下垂体低下に関連した障害と、良性線種によって発生しやすい不适当抗利尿ホルモン (ADH) 分泌症候群 (SIADH) 及び先端巨大症、巨人症を含む下垂体亢進に関連した障害と、甲状腺腫及び粘液水腫、細菌感染性急性甲状腺炎、ウイルス感染性亜急性甲状腺炎、自己免疫性甲状腺炎 (橋本病)、クレチン病を含む甲状腺機能低下症に関連した障害と、甲状腺中毒症及びその様々な型、グレース病、前脛骨粘液水腫、中毒性多結節性甲状腺腫、甲状腺癌、プランマー病を含む甲状腺機能亢進症と、Conn病 (chronic hypercalcemia) を含む副甲状腺機能亢進症と、I型及びII型糖尿病及び合併症などの膵臓疾患と、過形成及び副腎皮質の癌腫や腺腫、アルカローシスに関連した高血圧、アミロイド症、低カリウム血、クッシング病、リドル症候群、Arnold-Healy-Gordon症候群、褐色細胞腫瘍、副腎機能不全などの副腎に関連した障害と、女性の異常プロラクチン産生及び不妊症、子宮内膜症、月経周期の摂動、多嚢胞性卵巣疾患、高プロラクチン血症、選択的性腺刺激ホルモン不全 (isolated gonadotropin deficiency)、無月経、乳汁漏出症、半陰陽、多毛症及び男性化、乳癌、閉経期後の骨粗鬆症、男性のライジッヒ細胞過形成、男性更年期、生殖細胞無形成症、ライジッヒ細胞腫瘍に関連した性機能亢進、アンドロゲン受容体の欠如に関連したアンドロゲン耐性、5-還元酵素症候群、女性乳房症などの生殖腺ステロイドホルモンに関連した疾患とが含まれ、眼の疾患の中には、結膜炎、乾性角結膜炎、角膜炎、上強膜炎、虹彩炎、後部ブドウ膜炎、緑内障、一過性黒内障、虚血性視神経症、視神経炎、レーバー遺伝性視神経症、硝子体剥離、網膜剥離、白内障、黄斑変性症、中心性漿液性脈絡網膜症、色素性網膜炎、脈絡膜黒色腫、球後腫瘍、交叉腫瘍 (chiasmaltumor) が含まれ、代謝障害の中には、副腎機能不全、脳腱黄色腫症、副腎皮質過形成、クマリン耐性、嚢胞性線維症、糖尿病、脂肪性肝硬変、果糖-1,6-ジホスファターゼ欠損症、ガラクトース血症、甲状腺腫、グルカゴノーマ、糖原病、遺伝性果糖不耐症、アドレナリン過剰症、腎臓不全症、上皮小体亢進症、副甲状腺低下症、高コレステロール血症、甲状腺亢進症、低血糖症、甲状腺低下症、高脂血症、脂質ミオパシー (lipid myopathies)、脂肪異栄養症、リソソーム蓄積症、メンケス症候群、後角症候群 (occipital horn syndrome)、マンノシドーシス、ノイラミニダーゼ欠損症、肥満症、ペントース-フェニルケトン尿症、プソイドビタミンD欠損症が含まれ、胃腸疾患の中には、嚥下障害、消化性食道炎、食道痙攣、食道狭窄、食道癌、消化不良、消化障害、胃炎、胃癌、食欲不振、悪心、嘔吐、胃不全麻痺、洞または幽門の浮腫、腹部アンギナ、胸焼け、胃腸炎、イレウス、腸管感染、消化性潰瘍、胆石症、胆嚢炎、胆汁うっ滞、膵臓炎、膵臓癌、胆道疾患、肝炎、高ビリルビン血症、遺伝性高ビリルビン血症、硬変症、肝臓の受動性うっ血、ヘパトーム、感染性大腸炎、潰瘍性大腸炎、潰瘍性直腸炎、クローン病、ホイ

10

20

30

40

50

ツプル病、マロリー ヴァイス症候群、結腸癌、結腸閉塞、過敏性腸症候群、短小腸症候群、下痢、便秘、胃腸出血、及び後天性免疫不全症候群（A I D S）腸症、黄疸、肝性脳症、肝腎症候群、肝炎、肝脂肪症、血色素症、ウィルソン病、 - 1 - アンチトリプシン欠損症、ライ症候群、原発性硬化性胆管炎、肝梗塞、門脈循環閉塞及び血栓、小葉中心壊死、肝臓紫斑病、肝静脈血栓、肝静脈閉塞症、子癇前症、子癇、妊娠性急性肝脂肪、妊娠性肝臓内胆汁うっ滞と、結節性再生及び腺腫、癌腫を含む肝癌とが含まれる。

【 0 2 4 9 】

別の実施例では、限定するものではないが上に列記した疾患を含む D M E 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、D M E またはその断片や誘導体を発現し得るベクターを患者に投与することも可能である。

10

【 0 2 5 0 】

更に別の実施例では、限定するものではないが上に列記した疾患を含む D M E 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 D M E を含む組成物を好適な医薬用担体と共に患者に投与することも可能である。

【 0 2 5 1 】

更に別の実施例では、限定するものではないが上に列記した疾患を含む D M E 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、D M E の活性を調節するアゴニストを患者に投与することも可能である。

【 0 2 5 2 】

更なる実施例では、D M E の発現または活性の増大に関連した疾患の治療または予防のために、患者に D M E のアンタゴニストを投与することが可能である。限定するものではないが、このような疾患の例には、上記した自己免疫 / 炎症の疾患、細胞増殖異常、発生または発達障害、内分泌障害、眼の疾患、代謝障害、および肝臓の疾患を含む胃腸疾患が含まれる。一実施態様では、D M E 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 D M E を発現する細胞または組織に薬剤を運ぶターゲティング或いは運搬機構として間接的に用いられ得る。

20

【 0 2 5 3 】

別の実施例では、限定するものではないが上に列記した疾患を含む D M E 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、D M E をコードするポリヌクレオチドの相補配列を発現するベクターを患者に投与することも可能である。

30

【 0 2 5 4 】

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

【 0 2 5 5 】

D M E のアンタゴニストは、当分野で一般的な方法を用いて製造することが可能である。詳しくは、精製された D M E を用いて抗体を作ったり、治療薬のライブラリをスクリーニングして D M E と特異的に結合するものを同定することが可能である。D M E の抗体も、当分野で一般的な方法を用いて製造することが可能である。このような抗体には、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、一本鎖、F a b フラグメント、及び F a b 発現ライブラリによって作られたフラグメントが含まれる。但し、これらに限定されるものではない。中和抗体（即ち二量体の形成を阻害する抗体）は通常、治療用に好適である。

40

【 0 2 5 6 】

抗体の産生のためには、ヤギ、ウサギ、ラット、マウス、ヒト及びその他のものを含む種々の宿主が、D M E または任意の断片、または免疫原性の特性を備えるそのオリゴペプチドの注入によって免疫化され得る。宿主の種に応じて、種々のアジュバントを用いて免疫

50

応答を高めることもできる。限定するものではないがこのようなアジュバントには、フロイントアジュバントと、水酸化アルミニウム等のミネラルゲルアジュバントと、リゾレシチン、ブルニックポリオール、ポリアニオン、ペプチド、油性乳剤、スカシガイのヘモシニアン、ジニトロフェノール等の界面活性剤とがある。ヒトに用いられるアジュバントの中では、BCG(カルメット ゲラン桿菌)及びコリネバクテリウム パルヴムが特に好ましい。

【0257】

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

10

【0258】

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

20

【0259】

更に、「キメラ抗体」作製のために発達したヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの技術が、好適な抗原特異性及び生物学的活性を備える分子を得るために用いられる(例えば、Morrisson, S. L. 他 (1984) Proc. Natl. Acad. Sci. 81: 6851-6855; Neuberger, M. S. 他 (1984) Nature 312: 604-608; Takeda, S. ら (1985) Nature 314: 452, 454を参照。)別法では、当分野で周知の方法を用いて、一本鎖抗体の産生のための記載された技術を適用して、DME特異性一本鎖抗体を生成する。関連特異性を有するガイディオタイプ組成が異なるような抗体を、ランダムな組合せの免疫グロブリンライブラリからチェーンシャッフリングによって産生することもできる(Burton D. R. (1991) Proc. Natl. Acad. Sci. USA 88: 10134-10137等を参照)。

30

【0260】

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

40

【0261】

DMEに対する特異的な結合部位を含む抗体断片も得ることができる。例えば、限定するものではないが、このような断片には、抗体分子のペプシン消化によって作製されるF(ab')₂断片と、F(ab')₂断片のジスルフィド架橋を還元することによって作製されるFab断片とがある。或いは、Fab発現ライブラリを作製することによって、モノクローナルFab断片を所望の特異性と迅速且つ容易に同定することが可能となる(Huse, W. D. 他 (1989) Science 256: 1275-1281等を参照)。

【0262】

50

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

【0263】

ラジオイムノアッセイ技術と共にScatchard分析などの様々な方法を用いて、DMEに対する抗体の親和性を評価する。親和性を結合定数 K_a で表すが、この K_a は、平衡状態の下でDME抗体複合体のモル濃度を遊離抗体と遊離抗原のモル濃度で除して得られる値である。多数のDMEエピトープに対して親和性が不均一なポリクローナル抗体医薬の K_a は、DMEに対する抗体の平均親和性または結合活性を表す。特定のDMEエピトープに単一特異的なモノクローナル抗体医薬の K_a は、親和性の真の測定値を表す。 K_a 値が $10^9 \sim 10^{12}$ L/molの高親和性抗体医薬は、DME抗体複合体が激しい操作に耐えなければならないイムノアッセイに用いるのが好ましい。 K_a 値が $10^6 \sim 10^7$ L/molの低親和性抗体医薬は、DMEが抗体から最終的に活性化状態で解離する必要がある免疫精製(immunopurification)及び類似の処理に用いるのが好ましい。(Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J. E. および Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY)。

【0264】

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

【0265】

本発明の別の実施例では、DMEをコードするポリヌクレオチド、またはその任意の断片や相補配列が、治療目的で使用することができる。ある実施態様では、DMEをコードする遺伝子のコーディング領域や調節領域に相補的な配列やアンチセンス分子(DNA及びRNA、修飾ヌクレオチド)を設計して遺伝子発現を変更することができる。このような技術は当分野では周知であり、センスまたはアンチセンスオリゴヌクレオチドまたは大きな断片が、DMEをコードする配列の制御領域から、またはコード領域に沿ったさまざまな位置から設計可能である。(Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totowa NJを参照)。

【0266】

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

10

20

30

40

50

) Blood 76:271、前出の Ausubel、Uckert, W. および W. Walther (1994) Pharmacol. Ther. 63(3):323-347等を参照)。その他の遺伝送達機構には、リポソーム系、人工的なウイルスエンベロープ及び当分野で公知のその他の系が含まれる(Rossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J. ら (1998) J. Pharm. Sci. 87(11):1308-1315、Morris, M.C. ら (1997) Nucleic Acids Res. 25(14):2730-2736. 等を参照)。

【0267】

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

【0268】

本発明の更なる実施例では、DMEの欠損による疾患や異常症は、DME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DME欠損細胞に導入することによって治療する。in vivo 或いは ex vitro の細胞に用いる機械的導入技術には、(i) 個々の細胞内への直接的なDNA微量注射法、(ii) 遺伝子銃、(iii) リポソームを介した形質移入、(iv) 受容体を介した遺伝子導入、及び(v) DNAトランスポソンの使用(Morgan, R.A. および W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217、Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. および H. Recipon (1998) Curr. Opin. Biotechnol. 9:445-450)がある。

【0269】

DMEの発現に影響を及ぼし得る発現ベクターには、限定するものではないが、PCDNA3.1、EPITAG、PRCCMV2、PREP、PVAX、PCR2-TOPOTAベクター(Invitrogen, Carlsbad CA)、PCMV-SCR

IPT、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。DMEを発現させるために、(i)恒常的に活性なプロモーター(例えば、サイトメガロウイルス(CMV)、ラウス肉腫ウイルス(RSV)、SV40ウイルス、チミジンキナーゼ(TK)、若しくは -アクチン遺伝子等)、(ii)誘導性プロモーター(例えば、市販されているT-REXプラスミド(Invitrogen)に含まれている、テトラサイクリン調節性プロモーター(Gossen, M. 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及びPINDに含まれている: Invitrogen)、FK506/ラパマイシン誘導性プロモーター、またはRU486/ミフェプリストーン誘導性プロモーター(Rossi, F.M.V. および H. M. Blau, 前出)、または(iii)正常な個体に由来するDMEをコードする内在性遺伝子の天然のプロモーター若しくは組織特異的プロモーターを用いることが可能である。

【0270】

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

【0271】

本発明の別の実施例では、DMEの発現に関連する遺伝子欠損によって起こる疾患や異常症は、(i)レトロウイルス末端反復配列(LTR)プロモーター若しくは独立したプロモーターのコントロール下でDMEをコードするポリヌクレオチドと、(ii)好適なRNAパッケージングシグナルと、(iii)追加のレトロウイルス・シス作用性RNA配列及び効率的なベクターの増殖に必要なコーディング配列を伴うRev応答性エレメント(RRE)とからなるレトロウイルスベクターを作製して治療することができる。レトロウイルスベクター(例えばPFB及びPFBNEO)はStratagene社から市販されており、刊行データ(Riviere, I. ら (1995) Proc. Natl. Acad. Sci. U.S.A. 92:6733-6737)に基づいている。上記データを引用することをもって本明細書の一部とする。ベクターは、好適なベクター産生細胞系(VPCL)において増殖され、VPCLは、標的細胞上の受容体に対する親和性を有するエンベロープ遺伝子またはVSVg等の汎親和性エンベロープタンパク質を発現する(Armentano, D. ら (1987) J. Virol. 61:1647-1650、Bender, M.A. ら (1987) J. Virol. 61:1639-1646、Adam, M.A. および A.D. Miller (1988) J. Virol. 62:3802-3806、Dull, T. ら (1998) J. Virol. 72:8463-8471、Zufferey, R. ら (1998) J. Virol. 72:9873-9880)。RIGGに付与された米国特許第5,910,434号(「Method for obtaining retrovirus packaging cell lines producing high t ransducing efficiency retroviral supernatant」)において、レトロウイルスパッケージング細胞系を得るための方法が開示され

ており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、細胞集団（例えばCD4⁺T細胞）の形質導入、及び形質導入した細胞の患者への戻しは、遺伝子治療の分野では当業者に公知の方法であり、多数の文献に記載されている（Ranga, U. ら（1997）*J. Virol.* 71:7020-7029、Bauer, G. ら（1997）*Blood* 89:2259-2267、Bonyhadi, M. L.（1997）*J. Virol.* 71:4707-4716、Ranga, U. ら（1998）*Proc. Natl. Acad. Sci. U.S.A.* 95:1201-1206、Su, L.（1997）*Blood* 89:2283-2290）。

【0272】

別法では、アデノウイルス系遺伝子治療の送達系を用いて、DMEの発現に関連する1或いは複数の遺伝子異常を有する細胞にDMEをコードするポリヌクレオチドを送達する。アデノウイルス系ベクターの作製及びパッケージングについては、当業者に公知である。

複製欠損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を臍臓の無損傷の臍島内に導入するために可変性であることが証明された（Csete, M. E. ら（1995）*Transplantation* 27:263-268）。使用できる可能性のあるアデノウイルスベクターは、Armentanoに付与された米国特許第5,707,618号（"Adenovirus vectors for gene therapy"）に記載されており、引用することをもって本明細書の一部とする。アデノウイルスベクターについては、Antinozzi, P. A. ら（1999）*Annu. Rev. Nutr.* 19:511-544 及び Verma, I. M. および N. Somia（1997）*Nature* 18:389:239-242 も参照されたい。両文献は、引用することをもって本明細書の一部とする。

【0273】

別法では、ヘルペス系遺伝子治療の送達系を用いて、DMEの発現に関連する1或いは複数の遺伝子異常を有する標的細胞にDMEをコードするポリヌクレオチドを送達する。単純疱疹ウイルス（HSV）系のベクターは、HSV親和性の中枢神経細胞にDMEを導入する際に特に重要である。ヘルペス系ベクターの作製及びパッケージングは、当業者に公知である。複製適格性単純ヘルペスウイルス（HSV）I型系のベクターは、レポーター遺伝子を霊長類の眼に送達するために用いられてきた（Liu, X. ら（1999）*Exp. Eye Res.* 169:385-395）。HSV-1ウイルスベクターの作製についても、DeLucaに付与された米国特許第5,804,413号（"Herpes simplex virus swains for gene transfer"）に開示されており、該特許の引用をもって本明細書の一部とする。米国特許第5,804,413号には、ヒト遺伝子治療を含む目的のために好適なプロモーターの制御下において細胞に導入される少なくとも1つの外在性遺伝子を有するゲノムを含む組換えHSV d92についての記載がある。上記特許はまた、ICP4、ICP27及びICP22のために除去される組換えHSV系統の作製及び使用について開示している。HSVベクターについては、Goins, W. F. ら（1999）*J. Virol.* 73:519-532 及び Xu, H. ら（1994）*Dev. Biol.* 163:152-161 も参照されたい。両文献は、引用をもって本明細書の一部とする。クロン化ヘルペスウイルス配列の操作、巨大ヘルペスウイルスのゲノムの異なった部分を含む多数のプラスミドを形質移入した後の組換えウイルスの産生、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は、当業者に公知の技術である。

【0274】

別法では、ウイルス（正の一本鎖RNAウイルス）ベクターを用いてDMEをコードするポリヌクレオチドを標的細胞に送達する。プロトタイプのウイルスであるセムリキ森林熱ウイルス（Semliki Forest Virus, SFV）の生物学的研究が広範に行われており、遺伝子導入ベクターがSFVゲノムに基づいていることが分かった（Garoff, H. および K. - J. Li（1998）*Cun. Opin. B*

10

20

30

40

50

i o t e c h . 9 : 4 6 4 - 4 6 9) 。 ウイルスRNAの複製中に、通常はウイルスのキャプシッドタンパク質をコードするサブゲノムRNAが作り出される。このサブゲノムRNAは、完全長のゲノムRNAより高いレベルに複製されるため、酵素活性(例えばプロテアーゼ及びポリメラーゼ)を有するウイルスタンパク質に比べてキャプシッドタンパク質が過剰産生される。同様に、DMEをコードする配列を ウイルスゲノムのカプシドをコードする領域に導入することによって、ベクター導入細胞において多数のDMEをコードするRNAが産生され、高いレベルでDMEが合成される。通常は ウイルスの感染が数日以内での細胞溶解に関係する一方で、シンドビスウイルス(SIN)の変異体を有するハムスター正常腎臓細胞(BHK-21)の持続的な感染を確立する能力は、 ウイルスの溶解複製を遺伝子治療に適用できるように好適に変更可能であることを示唆している(Dryga, S. A. ら。(1997) *Virology* 228 : 74 - 83)。様々な宿主に ウイルスを導入できることから、様々なタイプの細胞にDMEを導入することができる。或る集団におけるサブセットの細胞の特定形質導入は、形質導入前に細胞の選別を必要とし得る。 ウイルスの感染性cDNAクローンの処置方法、 ウイルスのcDNA及びRNAの形質移入方法及び ウイルスの感染方法は、当業者に公知である。

【0275】

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

【0276】

リボザイムは酵素性RNA分子であり、RNAの特異的切断を触媒するためにリボザイムを用いることもできる。リボザイム作用のメカニズムは、相補的標的RNAへのリボザイム分子の配列特異性ハイブリダイゼーションその後起こる内ヌクレオチド鎖切断に参与している。例えば、DMEをコードする配列のヌクレオチド鎖切断を、特異的且つ効果的に触媒する組換え型のハンマーヘッド型リボザイム分子が含まれる。

【0277】

任意のRNA標的内の特異的リボザイム切断部位を、GUA、GUU、GUC配列を含めたリボザイム切断部位に対する標的分子をスキャンすることによって先ず同定する。一度同定すると、切断部位を含む標的遺伝子の領域に対応する15~20リボヌクレオチドの短いRNA配列が、そのオリゴヌクレオチドを機能不全にするような2次構造の特徴をもっていないかを、評価することが可能になる。候補標的の適合性の評価も、リボヌクレアーゼ保護アッセイを用いて相補的オリゴヌクレオチドとのハイブリダイゼーションの実施容易性をテストすることによって行うことができる。

【0278】

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

ことができる。

【0279】

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

10

【0280】

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

【0281】

特異ポリヌクレオチドの変異発現における有効性に対して、少なくとも1個から複数個の試験化合物をスクリーニングし得る。試験化合物は、当分野で通常知られている任意の方法により得られる。このような方法には、ポリヌクレオチドの発現を変異させる場合と、既存の、市販のまたは専売の、天然または非天然の化合物ライブラリから選択する場合と、標的ポリヌクレオチドの化学的及び/または構造的特性に基づく化合物を合理的にデザインする場合と、組合せ的にまたは無作為に生成した化合物のライブラリから選択する場合に有効であることが知られているような化合物の化学修飾がある。DMEをコードするポリヌクレオチドを含むサンプルは、このようにして得られた試験化合物の少なくとも1つの試験化合物に曝露する。サンプルには例えば、無傷細胞、透過化処理した細胞、*in vitro*無細胞系または再構成生化学系があり得る。DMEをコードするポリヌクレオチドの発現における変化は、当分野で周知の任意の方法でアッセイする。通常、DMEをコードするポリヌクレオチドの配列に相補的なヌクレオチド配列を有するプローブを用いたハイブリダイゼーションにより、特定のヌクレオチドの発現を検出する。ハイブリダイゼーション量を定量し、それによって1つ以上の試験化合物に曝露される及び曝露されないポリヌクレオチドの発現の比較に対する基礎を形成し得る。試験化合物に曝露されるポリヌクレオチドの発現における変化の検出は、ポリヌクレオチドの発現を変異する際に試験化合物が有効であることを示している。特異ポリヌクレオチドの変異発現に有効な化合物に対して、例えば*Schizosaccharomyces pombe*遺伝子発現系(Atkins, D. ら(1999) 米国特許第5,932,435号、Arndt, G.M. ら(2000) *Nucleic Acids Res.* 28:E15)またはHeLa細胞等のヒト細胞系(Clarke, M.L. ら(2000) *Biochem. Biophys. Res. Commun.* 268:8-13)を用いてスクリーニングを実行する。本発明の特定の実施例は、特異的ポリヌクレオチド配列に対するアンチセンス活性のためのオリゴヌクレオチド(デオキシリボヌクレオチド、リボヌクレオチド、ペプチド核酸、修飾オリゴヌクレオチド)の組み合わせライブラリをスクリーニングすることに関与している(Bruice, T.W. ら(1997) 米国特

30

40

50

許第5,686,242号、Bruce, T.W. ら (2000) 米国特許第6,022,691号)。

【0282】

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

10

【0283】

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

【0284】

本発明の追加実施例は、通常薬剤として許容できる賦形剤で処方される活性成分を有する成分の投与に関連する。賦形剤には例えば、糖、でんぷん、セルロース、ゴム及びタンパク質がある。様々な処方通常知られており、詳細は Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA) の最新版に記載されている。このような組成物は、DME、DMEの抗体、擬態、アゴニスト、アンタゴニスト、またはDMEのインヒビターなどからなる。

20

【0285】

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

【0286】

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

30

【0287】

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

【0288】

DMEまたはその断片を含む高分子を直接細胞内に輸送するべく、特殊な形態に組成物が調製されるのが好ましい。例えば、細胞不透過性高分子を含むリポソーム製剤は、細胞融合及び高分子の細胞内送達を促進し得る。別法では、DMEまたはその断片を HIV Tat-1 タンパク質の陽イオンN末端部に結合することもできる。このようにして生成された融合タンパク質は、マウスモデル系の脳を含む全ての組織の細胞に形質導入することがわかっている (Schwarze, S.R. ら (1999) Science 285:1569-1572)。

40

【0289】

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

50

【0290】

医学的に効果的な薬用量は、症状や容態を回復させる、たとえばDMEまたはその断片、DMEの抗体、DMEのアゴニストまたはアンタゴニスト、インヒビターなどの活性処方成分の量に関連する。治療有効度及び毒性は、細胞培養または動物実験における標準的な薬剤手法によって、例えばED₅₀（集団の50%の医薬的有効量）またはLD₅₀（集団の50%の致死量）を測定するなどして決定することができる。毒性効果の薬用効果に対する投与量の比は、治療指数であり、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成分が望ましい。細胞培養アッセイ及び動物実験から得られたデータは、ヒトに用いるための投与量の範囲を調剤するのに用いられる。このような組成物が含まれる投与量は、毒性を殆ど或いは全く含まず、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、患者の感受性及び投与の経路によって、投与量はこの範囲内で様々に変わる。

10

【0291】

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

20

【0292】

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

【0293】

(診断)

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

30

【0294】

DMEを測定するためのELISA、RIA、及びFACSを含む種々のプロトコルは、当分野では周知であり、変わった或いは異常なレベルのDMEの発現を診断する元となるものを提供する。正常或いは標準的なDMEの発現の値は、複合体の形成に適した条件下、正常な哺乳動物、例えばヒトなどの被験者から採取した体液または細胞とDMEに対する抗体とを結合させることによって決定する。標準複合体形成量は、種々の方法、例えば測光法で定量できる。被験者のDMEの発現の量、対照及び疾患、生検組織からのサンプルが基準値と比較される。基準値と被験者との偏差が疾患を診断するパラメータとなる。

40

【0295】

本発明の別の実施例によれば、DMEをコードするポリヌクレオチドを診断のために用いることもできる。用いることができるポリヌクレオチドには、オリゴヌクレオチド配列、

50

相補的RNA及びDNA分子、そしてPNAが含まれる。このポリヌクレオチドを用いて、疾患と相関し得るDMEを発現する生検組織における遺伝子の発現を検出し定量する。この診断アッセイを用いて、DMEの存在の有無、更に過剰な発現を調べ、治療中のDME値の調節を監視する。

【0296】

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

10

【0297】

プローブはまた、関連する配列の検出に利用され、DMEをコードする任意の配列と少なくとも50%の配列同一性を有し得る。目的の本発明のハイブリダイゼーションプローブには、DNAあるいはRNAが可能であり、SEQ ID NO: 20-38の配列、或いはDME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。

【0298】

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

20

【0299】

DMEをコードするポリヌクレオチド配列を用いて、DMEの発現に関連する疾患を診断することが可能である。限定するものではないが、このような疾患には自己免疫/炎症の疾患、細胞増殖異常、発生または発達障害、内分泌障害、眼の疾患、代謝障害、および肝臓の疾患を含む胃腸疾患が含まれ、自己免疫/炎症の疾患の中には、炎症及び日光性角化症、後天性免疫不全症候群(AIDS)及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫性甲状腺炎、自己免疫性多腺性内分泌カンジダ性外胚葉ジストロフィー(APECED)、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、赤芽球症、結節性紅斑、萎縮性胃炎、糸球体腎炎、グッドパスチャー症候群、痛風、グレーブス病、橋本甲状腺炎、過好酸球増加症、過敏性大腸症候群、多発性硬化症、重症筋無力症、心筋または心膜炎、骨関節炎、骨粗しょう症、腓炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、細胞増殖異常の中には日光性角化症及びアテローム性動脈硬化、滑液包炎、硬変、肝炎、混合型結合組織病(MCTD)、骨髄線維症、発作性夜間ヘモグロビン尿症、真性多血症、乾癬、原発性血小板血症、並びに腺癌及び白血病、リンパ腫、黒色腫、骨髄腫、肉腫、及び奇形癌、具体的には、副腎、膀胱、骨、骨髄、脳、乳房、頸部、胆嚢、神経節、消化管、心臓、腎臓、肝臓、肺、筋肉、卵巣、膵臓、副甲状腺、陰茎、前立腺、唾液腺、皮膚、脾臓、精巣、胸腺、甲状腺、子宮の癌が含まれ

30

40

50

、発生または発達障害の中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、精神薄弱）、スミス マジェニス症候群（Smith-Magenis syndrome）、脊髄形成異常症候群、遺伝性粘膜上皮異形成、遺伝性角皮症、シャルコー マリー ツース病及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、Syndenham 舞蹈病（Syndenham's chorea）及び脳性小児麻痺などの発作障害、脊髄二分裂、無脳症、頭蓋脊椎披裂、先天性緑内障、白内障、感覚神経性聴力損失が含まれ、内分泌障害の中には原発脳腫瘍及び腺腫、妊娠性梗塞、下垂体切除、動脈瘤、血管奇形、血栓症、感染症、免疫異常、頭部外傷による合併症などの病変から起こる視床下部及び下垂体の障害と、性機能低下及びシーハン症候群、尿崩症、カルマン病、ハンド シュラー クリスマン病、レトラ シヴェ病、サルコイドーシス、エンブティセラ症候群、小人症を含む下垂体低下に関連した障害と、良性線種によって発生しやすい不适当抗利尿ホルモン（ADH）分泌症候群（SIADH）及び先端巨大症、巨人症を含む下垂体亢進に関連した障害と、甲状腺腫及び粘液水腫、細菌感染性急性甲状腺炎、ウイルス感染性亜急性甲状腺炎、自己免疫性甲状腺炎（橋本病）、クレチン病を含む甲状腺機能低下症に関連した障害と、甲状腺中毒症及びその様々な型、グレーブス病、前脛骨粘液水腫、中毒性多結節性甲状腺腫、甲状腺癌、プランマー病を含む甲状腺機能亢進症と、Conn病（chronic hypercalcemia）を含む副甲状腺機能亢進症と、I型及びII型糖尿病及び合併症などの膵臓疾患と、過形成及び副腎皮質の癌腫や腺腫、アルカローシスに関連した高血圧、アミロイド症、低カリウム血、クッシング病、リドル症候群、Arnold-Healy-Gordon症候群、褐色細胞腫瘍、副腎機能不全などの副腎に関連した障害と、女性の異常プロラクチン産生及び不妊症、子宮内膜症、月経周期の摂動、多嚢胞性卵巣疾患、高プロラクチン血症、選択的性腺刺激ホルモン不全（isolated gonadotropin deficiency）、無月経、乳汁漏出症、半陰陽、多毛症及び男性化、乳癌、閉経期後の骨粗鬆症、男性のライジッヒ細胞過形成、男性更年期、生殖細胞無形成症、ライジッヒ細胞腫瘍に関連した性機能亢進、アンドロゲン受容体の欠如に関連したアンドロゲン耐性、5-還元酵素症候群、女性乳房症などの生殖腺ステロイドホルモンに関連した疾患とが含まれ、眼の疾患の中には、結膜炎、乾性角結膜炎、角膜炎、上強膜炎、虹彩炎、後部ブドウ膜炎、緑内障、一過性黒内障、虚血性視神経症、視神経炎、レーバー遺伝性視神経症、硝子体剥離、網膜剥離、白内障、黄斑変性症、中心性漿液性脈絡網膜症、色素性網膜炎、脈絡膜黒色腫、球後腫瘍、交叉腫瘍（chiasmaltumor）が含まれ、代謝障害の中には、副腎機能不全、脳腱黄色腫症、副腎皮質過形成、クマリン耐性、嚢胞性線維症、糖尿病、脂肪性肝硬変、果糖-1,6-ジホスファターゼ欠損症、ガラクトース血症、甲状腺腫、グルカゴノーマ、糖原病、遺伝性果糖不耐症、アドレナリン過剰症、腎臓不全症、上皮小体亢進症、副甲状腺低下症、高コレステロール血症、甲状腺亢進症、低血糖症、甲状腺低下症、高脂血症、脂質ミオパシー（lipid myopathies）、脂肪異栄養症、リソソーム蓄積症、メンケス症候群、後角症候群（occipital horn syndrome）、マンノシドーシス、ノイラミニダーゼ欠損症、肥満症、ペントース-フェニルケトン尿症、プソイドビタミンD欠損症が含まれ、胃腸疾患の中には、嚥下障害、消化性食道炎、食道痙攣、食道狭窄、食道癌、消化不良、消化障害、胃炎、胃癌、食欲不振、悪心、嘔吐、胃不全麻痺、洞または幽門の浮腫、腹部アングナ、胸焼け、胃腸炎、イレウス、腸管感染、消化性潰瘍、胆石症、胆嚢炎、胆汁うっ滞、膵臓炎、膵臓癌、胆道疾患、肝炎、高ビリルビン血症、遺伝性高ビリルビン血症、硬変症、肝臓の受動性うっ血、ヘパトーム、感染性大腸炎、潰瘍性大腸炎、潰瘍性直腸炎、クローン病、ホイップル病、マロリー ヴァイス症候群、結腸癌、結腸閉塞、過敏性腸症候群、短小腸症候群、下痢、便秘、胃腸出血、及び後天性免疫不全症候群（AIDS）腸症、黄疸、肝性脳症、肝腎症候群、肝炎、肝脂肪症、血色素症、ウィルソン病、-1-アンチトリプシン欠損症、ライ症候群、原発性硬化性胆管炎、肝梗塞、門脈循環閉塞及び血栓、小葉中心壊死、肝臓紫斑病、肝静脈血栓、肝静脈閉塞症、子癇前症、子癇

10

20

30

40

50

、妊娠性急性肝脂肪、妊娠性肝臓内胆汁うっ滞と、結節性再生及び腺腫、癌腫を含む肝癌とが含まれる。MADをコードするポリヌクレオチド配列は、サザン法やノーザン法、ドットプロット法、或いはその他の膜系の技術、PCR法、ディップスティック (dipstick)、ピン (pin)、ELISA式アッセイ、及び変異DMEの発現を検出するために患者から採取した体液或いは組織を利用するマイクロアレイに使用することが可能である。このような定性方法または定量方法は、当分野で公知である。

【0300】

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

10

【0301】

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

20

【0302】

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

30

【0303】

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

【0304】

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

40

【0305】

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

50

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

【0306】

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

【0307】

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

【0308】

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

【0309】

或る実施例は、或る組織または細胞タイプの転写イメージを生成するような本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞タイプにより遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所定の条件下で所定の時間に発現した遺伝子の数及び相対存在量を定量することにより分析される(Seilliamer 他、米国特許第5,840,484号の"Comparative Gene Transcript Analysis"を参照。この特許に言及することを以って本明細書の一部とする)。従って、特定の組織または細胞タイプの転写または逆転写全体に本発明のポリヌクレオチドまたはその補体をハイブリダイズすることにより、転写イメ

ージを生成し得る。或る実施例では、本発明のポリヌクレオチドまたはその補体がマイクロアレイ上のエレメントのサブセットを複数含むような高処理フォーマットでハイブリダイゼーションを発生させる。結果として得られる転写イメージは、遺伝子活性のプロフィールを提供し得る。

【0310】

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

【0311】

本発明のポリヌクレオチドの発現のプロフィールを作製する転写イメージはまた、工業的または天然の環境化合物の毒性試験のみならず、in vitro モデル系及び薬剤の前臨床評価に関連して使用し得る。全ての化合物は、作用及び毒性のメカニズムを暗示する、しばしば分子フィンガープリントまたは毒性シグネチャと称されるような特徴的な遺伝子発現パターンを惹起する (Nuwaysir, E. F. ら. (1999) Mol. Carcinog. 24: 153-159, Steiner, S. および N. L. Anderson (2000) Toxicol. Lett. 112-113: 467-471、該文献は特に引用することを以って本明細書の一部となす)。試験化合物が、既知の毒性を有する化合物のシグネチャと同一のシグネチャを有する場合には、毒性特性を共有している可能性がある。フィンガープリントまたはシグネチャは、多数の遺伝子及び遺伝子ファミリーからの発現情報を含んでいる場合には、最も有用且つ正確である。理想的には、発現のゲノム全域にわたる測定が最高品質のシグネチャを提供する。たとえ、発現が任意の試験された化合物によって変化しない遺伝子があったとしても、それらの発現レベルを残りの発現データを標準化するために使用できるため、それらの遺伝子は重要である。規準化手順は、異なる化合物で処理した後の発現データの比較に有用である。毒性シグネチャの要素に遺伝子機能を割り当てることが毒性メカニズムの解釈に役立つが、毒性の予測につながるシグネチャの統計的に一致させるのに遺伝子機能の知識は必要とされない (例えば2000年2月29日にNational Institute of Environmental Health Sciencesより発行されたPress Release 00-02を参照されたい。これについては <http://www.niehs.nih.gov/oc/news/toxchip.htm> で入手可能である)。従って、中毒学的スクリーニングの際に毒性シグネチャを用いて、全ての発現した遺伝子配列を含めることは重要且つ望ましいことである。

【0312】

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

【0313】

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

10

20

30

40

50

ブルーまたはシルバーまたは蛍光染色などの物質を用いてゲルで染色することにより、分散した、独特な位置にある点としてゲル中で可視化される。各タンパク質スポットの光学密度は、通常サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物または治療薬で処理または非処理のいずれかの生物学的サンプルから得られるタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を同定する。スポット内のタンパク質は、例えば化学的または酵素的切断とそれに続く質量分析を用いる標準的な方法を用いて部分的にシーケンシングする。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が得られる。

10

【0314】

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

20

【0315】

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

30

【0316】

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

【0317】

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

40

【0318】

マイクロアレイは、本技術分野でよく知られている方法を用いて調製し、使用し、そして分析する(Brennan, T.M. ら (1995) の米国特許第5,474,796号、Schena, M. ら (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619、Baldeschweiler らの (1995) PCT出願第WO95/251116号、Shalon, D. らの (199

50

5) PCT出願第WO95/35505号、Heller, R.A. ら (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155、Heller, M.J. らの (1997) 米国特許第5,605,662号等を参照)。様々なタイプのマイクロアレイが公知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, 編集。(1999) Oxford University Press, Londonに記載されている。該文献は、特に引用することを以って本明細書の一部となす。

【0319】

本発明の別の実施例ではまた、DMEをコードする核酸配列を用いて、天然のゲノム配列をマッピングするのに有用なハイブリダイゼーションプローブを作製することが可能である。コード配列または非コード配列のいずれかを用いることができ、或る例では、コード配列全体で非コード配列が好ましい。例えば、多重遺伝子ファミリーのメンバー内でのコード配列の保存により、染色体マッピング中に望ましくないクロスハイブリダイゼーションが生じる可能性がある。核酸配列は、特定の染色体、染色体の特定領域または人工形成の染色体、例えば、ヒト人工染色体(HAC)、酵母人工染色体(YAC)、細菌人工染色体(BAC)、細菌P1産物、或いは単一染色体cDNAライブラリに対してマッピングされる(例えば、Harrington, J.J. 他(1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127134; および Trask, B.J. (1991) Trends Genet. 7:149154を参照。)一度マッピングすると、本発明の核酸配列を用いて例えば病状の遺伝を特定の染色体領域の遺伝または制限酵素断片長多型(RFLP)と関連させるような遺伝子連鎖地図を発生させ得る。(例えば、Lander, E.S. および D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357を参照。)

蛍光 in situ ハイブリッド形成法(FISH)は、他の物理的及び遺伝地図データと関連し得る(前出のHeinz-Ulrich, ら(1995) in Meyers, 965-968ページ、等を参照) 遺伝地図データの例は、種々の科学雑誌あるいはOnline Mendelian Inheritance in Man(OMIM)のウェブサイトに見ることができる。物理的な染色体地図上のDMEをコードする遺伝子の位置と特定の疾患との相関性、或いは特定の疾患に対する素因が、このような疾患と関連するDNA領域の決定に役立つため、更なる位置を決定するクローニングが行われる。

【0320】

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

【0321】

本発明の別の実施例では、DME、その触媒作用断片或いは免疫原断片またはそのオリゴペプチドを、種々の任意の薬剤スクリーニング技術における化合物のライブラリのスクリーニングに用いることができる。薬剤スクリーニングに用いる断片は、溶液中に遊離しているか、固体支持物に固定されるか、細胞表面上に保持されるか、細胞内に位置することになる。DMEと検査する薬剤との結合による複合体の形成を測定してもよい。

【0322】

10

20

30

40

50

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

【0323】

別の実施例では、DMEと結合可能な中和抗体がDMEと結合するため試験用化合物と特に競合する、競合的薬剤スクリーニングアッセイを用いることができる。この方法では、抗体が、DMEと1つ以上の抗原決定因子を共有するどのペプチドの存在も検出する。

【0324】

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

【0325】

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

【0326】

本明細書において開示した全ての特許、特許出願及び刊行物、特に米国特許第60/223,055号、第60/224,728号、第226,440号、第228,067号、第230,063号、第232,244及び第234,269は、言及することをもって本明細書の一部となす。

【0327】

(実施例)

1 cDNAライブラリの作製

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

【0328】

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

【0329】

場合によってはStratagene社へのRNA提供を行い、対応するcDNAライブラリをStratagene社が作製することもあった。そうでない場合は、UNIZAPベクターシステム (Stratagene) またはSUPERSCRIPTプラスミド

システム (Life Technologies) を用いて本技術分野で公知の推奨方法または類似の方法で cDNA を合成し、cDNA ライブラリを作製した (前出の Ausubel, 1997, unit 5.1-6.6 等を参照)。逆転写は、オリゴ d(T) またはランダムプライマーを用いて開始した。合成オリゴヌクレオチドアダプターを二本鎖 cDNA に連結反応させ、好適な制限酵素または酵素で cDNA を消化した。殆どのライブラリに対して、cDNA のサイズ (300 ~ 1000 bp) 選択は、SEPHACRYL S1000、SEPHAROSE CL2B または SEPHAROSE CL4B カラムクロマトグラフィー (Amersham Pharmacia Biotech)、或いは調製用アガロースゲル電気泳動法を用いて行った。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でライゲーションされた。好適なプラスミドは、例えば PBLUESCRIPT プラスミド (Stratagene)、PSPORT1 プラスミド (Life Technologies)、PCDNA2.1 プラスミド (Invitrogen, Carlsbad CA)、PBK-CMV プラスミド (Stratagene)、PCR2-TOPO TA (Invitrogen)、pINCY (Incyte Pharmaceuticals, Palo Alto CA) 等およびその誘導体である。組換えプラスミドは、Stratagene 社の XL1-Blue、XL1-Blue MRF または SOLR、或いは Life Technologies 社の DH5、DH10B または ELECTROMAX DH10B を含む大腸菌細胞に形質転換した。

【0330】

2 cDNA クローンの単離

UNIZAP ベクターシステム (Stratagene) を用いた *in vivo* 切除によって、或いは細胞溶解によって、実施例 1 のようにして得たプラスミドを宿主細胞から回収した。Magic または WIZARD Minipreps DNA 精製システム (Promega)、AGTC Miniprep 精製キット (Edge Biosystems, Gaithersburg MD)、QIAGEN 社の QIAWELL 8 Plasmid、QIAWELL 8 Plus Plasmid 及び QIAWELL 8 Ultra Plasmid 精製システム、R.E.A.L. Prep 96 プラスミドキットの中から少なくとも 1 つを用いて、プラスミドを精製した。沈殿させた後、0.1 ml の蒸留水に再懸濁して、凍結乾燥して或いは凍結乾燥せずに、4 で保管した。

【0331】

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

【0332】

3 シークエンシング及び分析

実施例 2 に記載したようにプラスミドから回収した Incyte cDNA を、以下に示すようにシークエンシングした。cDNA のシークエンス反応は、標準的方法或いは高処理装置、例えば ABI CATALYST 800 サーマルサイクラー (Applied Biosystems) または PTC-200 サーマルサイクラー (MJ Research) を HYDRA マイクロディスペンサー (Robbins Scientific) または MICROLAB 2200 (Hamilton) 液体転移システムと併用して処理した。cDNA のシークエンス反応は、Amersham Pharmacia Biotech 社が提供する試薬または ABI シークエンシングキット、例えば ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems) に与えられた

試薬を用いて準備した。cDNAのシーケンス反応の電気泳動的分離及び標識したポリヌクレオチドの検出には、MEGABACE 1000 DNAシーケンシングシステム (Molecular Dynamics) が、標準ABIプロトコル及び塩基対呼び出しソフトウェアを用いるABI PRISM 373または377シーケンシングシステム (Applied Biosystems) が、或いはその他の本技術分野でよく知られている配列解析システムを用いた。cDNA配列内のリーディングフレームは、標準的方法 (前出のAusubel, 1997, unit 7.7に概説) を用いて決定した。cDNA配列の幾つかを選択して、実施例8に記載した方法で配列を伸長させた。

【0333】

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

【0334】

Incyte cDNA及び完全長配列の分析及びアセンブリに利用したツール、プログラム及びアルゴリズムの概略と、適用可能な説明、引用文献、閾値パラメータを表7に示す。用いたツール、プログラム及びアルゴリズムを表7の列1に、それらの簡単な説明を列2に示す。列3は好適な引用文献であり、全ての文献はそっくりそのまま引用を以って本明細書の一部となす。適用可能な場合には、列4は2つの配列が一致する強さを評価するために用いたスコア、確率値その他のパラメータを示す(スコアが高ければ高いほどまたは確率値が低ければ低いほど、2配列間の相同性が高くなる)。

【0335】

完全長のポリヌクレオチド配列およびポリペプチド配列の組み立て及び分析に用いる上記

のプログラムは、SEQ ID NO: 20 - 38のポリヌクレオチド配列断片の同定にも利用できる。ハイブリダイゼーション及び増幅技術に有用である約20～約4000ヌクレオチドの断片を表4の列4に示した。

【0336】

4 ゲノムDNAからのコード配列の同定及び編集

推定薬剤代謝酵素は、公共のゲノム配列データベース（例えば、gbpriやgbhtg）においてGenscan遺伝子同定プログラムを実行して初めに同定された。Genscanは、様々な生物からゲノムDNA配列を分析する汎用遺伝子同定プログラムである（Burge, C. および S. Karlin (1997) J. Mol. Biol. 268: 78-94 及び Burge, C. 及び S. Karlin (1998) Cuff. Opin. Struct. Biol. 8: 346-354 参照）。プログラムは予測エキソンを連結し、メチオニンから停止コドンに及ぶ構築されたcDNA配列を形成する。Genscanの出力は、ポリヌクレオチド及びポリペプチド配列のFASTAデータベースである。Genscanが一度に分析する配列の最大範囲は、30kbに設定した。これらのGenscan推定cDNA配列の内、どの配列が薬剤代謝酵素をコードするかを決定するために、コードされたポリペプチドをPFAMモデルにおいて薬剤代謝酵素について問合せて分析した（7tm__1、7tm__2、7tm__3、および7tm__4）。潜在的な薬剤代謝酵素が、薬剤代謝酵素としてアノテーションが付けられたインサイトcDNA配列に対する相同性を基に同定された。こうして選択されたGenscan予測配列は、次にBLAST分析により公共データベースgbpri及びgbhtgと比較した。必要であれば、genpeptからヒットしたトップのBLASTと比較することによりGenscan予測配列を編集し、余分なまたは取り除かれたエキソンなどのGenscanにより予測された配列のエラーを修正する。BLAST分析はまた、任意のIncyte cDNAまたはGenscan予測配列の公共cDNA適用範囲の発見に用いられるので、転写の証拠を提供する。Incyte cDNA適用範囲が利用できる場合には、この情報を用いてGenscan予測配列を修正または確認した。完全長ポリヌクレオチド配列は、実施例3に記載された構築プロセスを用いて、Incyte cDNA配列及び/または公共のcDNA配列でGenscan予測コード配列を構築することにより得た。或いは、完全長ポリヌクレオチド配列は編集または非編集のGenscan予測コード配列に完全に由来する。

【0337】

5 cDNA配列データを使ったゲノム配列データの構築

ステッチ配列 (Stitched Sequence)

部分cDNA配列は、実施例4に記載のGenscan遺伝子同定プログラムにより予測されたエキソンを用いて伸長させた。実施例3に記載されたように構築された部分cDNAは、ゲノムDNAにマッピングし、関連するcDNA及び1つ若しくは複数のゲノム配列から予測されたGenscanエキソンを含むクラスタに分解した。cDNA及びゲノム情報を統合するべくグラフ理論及び動的プログラミングに基づくアルゴリズムを用いて各クラスタを分析し、引き続いて確認、編集または伸長して完全長配列を産出するような潜在的スプライス変異体を生成した。間隔全体の長さがクラスタ中の2以上の配列に存在するような配列を同定し、そのように同定された間隔は推移により等しいと考えられた。例えば、1つのcDNA及び2つのゲノム配列に間隔が存在する場合、3つの間隔は全て等しいと考えられる。このプロセスは、無関係であるが連続したゲノム配列をcDNA配列により結び合わせて架橋し得る。このようにして同定された区間を、親配列 (parent sequence) に沿って現われるようにステッチアルゴリズムで縫い合わせ、可能な最も長い配列および変異配列を作製する。1種類の親配列に沿って発生した間隔と間隔との連鎖 (cDNA - cDNA または ゲノム配列 - ゲノム配列) は、親の種類を変える連鎖 (cDNA - ゲノム配列) に優先した。結果として得られるステッチ配列は、BLAST分析により公共データベースgenpept及びgbpriに翻訳されて比較された。Genscanにより予測された不正確なエキソンは、genpeptからヒットし

たトップのBLASTと比較することにより修正した。必要な場合には、追加cDNA配列を用いるかゲノムDNAの検査により配列を更に伸長させた。

【0338】

ストレッチ配列 (Stretched Sequence)

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

10

【0339】

6 DME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: 20-38を構築するために用いた配列を、BLAST及びSmith-Watermanアルゴリズムを用いて、Incyte LIFESEQデータベース及び公共のドメインデータベースの配列と比較した。SEQ ID NO: 20-38と一致するこれらのデータベースの配列を、Phrap(表7)などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。スタンフォード・ヒトゲノムセンター(SHGC)、ホワイトヘッド・ゲノム研究所(WIGR)、Genethon等の公的な情報源から入手可能な放射線ハイブリッド及び遺伝地図データを用いて、クラスタ化された配列が前もってマッピングされたかを決定した。マッピングされた配列がクラスタに含まれている場合は、個々の配列番号を含めてそのクラスタの全配列が地図上の位置に割り当てられた。

20

【0340】

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

30

40

【0341】

7 ポリヌクレオチド発現の分析

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

【0342】

BLASTに適用する類似のコンピュータ技術を用いて、GenBankやLifeSeq(Incyte Genomics)等のcDNAデータベースにおいて同一または関

50

連分子を検索する。ノーザン分析は、多数膜系ハイブリダイゼーションよりも非常に速い。更に、特定の同一を厳密な或いは相同的なものとして分類するか否かを決定するため、コンピュータ検索の感度を変更することができる。検索の基準は積スコアであり、次式で定義される。

【数 1】

(BLAST スコア×配列一致率)

5 × (長さ(配列 1), 長さ(配列 2)) の最小値

【0343】

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

10

20

【0344】

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

30

【0345】

8 DMEをコードするポリヌクレオチドの伸長
完全長のポリヌクレオチド配列もまた、完全長分子の適切な断片から設計したオリゴヌクレオチドプライマーを用いて該断片を伸長させて生成した。一方のプライマーは既知の断片の5'伸長を開始するべく合成し、他方のプライマーは既知の断片の3'伸長を開始するべく合成した。開始プライマーは、長さが約22～30ヌクレオチド、GC含有率が約50%以上となり、約68～72の温度で標的配列にアニーリングするように、OLIGO 4.06ソフトウェア(National Biosciences)或いは別の適切なプログラムを用いて、cDNAから設計した。ヘアピン構造及びプライマー-プライマー二量体を生ずるようなヌクレオチドの伸長は全て回避した。

40

【0346】

配列を伸長するために、選択されたヒトcDNAライブラリを用いた。2段階以上の伸長

50

が必要または望ましい場合には、付加的プライマー或いはプライマーのネステッドセットを設計した。

【0347】

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

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

【0348】

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

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

【0349】

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

【0350】

伸長させたヌクレオチドは、脱塩及び濃縮して384穴プレートに移し、CviJIコラウイルスエンドヌクレアーゼ(Molecular Biology Research, Madison WI)を用いて消化し、pUC 18ベクター(Amersham Pharmacia Biotech)への再連結反応前に音波処理またはせん断した。シヨットガン・シークエンシングのために、消化したヌクレオチドを低濃度(0.6~0.8%)のアガロースゲル上で分離し、断片を切除し、寒天をAgar ACE(Promega)で消化した。伸長させたクローンをT4リガーゼ(New England Biolabs, Beverly MA)を用いてpUC 18ベクター(Amersham Pharmacia Biotech)に再連結し、Pfu DNAポリメラーゼ(Stratagene)で処理して制限部位の張出部(overhang)を満たし、大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りとってLB/2Xカルベニシリン培養液の384ウェルプレートに37

で一晩培養した。

【0351】

細胞を溶解して、Taq DNAポリメラーゼ (Amersham Pharmacia Biotech) 及び Pfu DNAポリメラーゼ (Stratagene) を用いて以下の手順でDNAをPCR増幅した。

ステップ1 94 で3分間
 ステップ2 94 で15秒間
 ステップ3 60 で1分間
 ステップ4 72 で2分間
 ステップ5 ステップ2、3、及び4を29回繰り返す
 ステップ6 72 で5分間
 ステップ7 4 で保存

10

【0352】

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

20

【0353】

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

【0354】

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: 20 - 38 から導き出されたハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記載するが、より大きなヌクレオチド断片に対しても事実上同一の手順が用いられる。オリゴヌクレオチドを、OLIGO4.06ソフトウェア (National Bioscience) のような最新式のソフトウェアを用いてデザインし、50 pmolの各オリゴマーと、250 µCiの [³²P] アデノシン三リン酸 (Amersham Pharmacia Biotech)) 及びT4ポリヌクレオチドキナーゼ (DuPont NEN, Boston MA) とを組み合わせて用いることにより標識する。標識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 ビードカラム (Amersham Pharmacia Biotech) を用いて十分に精製する。Ase I、Bgl II、Eco RI、Pst I、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10⁷ 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。

30

40

【0355】

各消化物から得たDNAは、0.7%アガロースゲル上で分画してナイロン膜 (Nytran Plus, Schleicher & Schuell, Durham NH) に移す。ハイブリダイゼーションは、40 で16時間行う。非特異的シグナルを除去するため、例えば0.1×クエン酸ナトリウム食塩水及び0.5%ドデシル硫酸ナトリウムに一致する条件下で、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。

50

【0356】

10 マイクロアレイ

マイクロアレイの表面上でアレイエレメントの連鎖または合成は、フォトリソグラフィ、圧電印刷（インクジェット印刷、前出のBaldeschweiler等を参照）、機械的マイクロスポッティング技術及びこれらから派生したものをを用いて達成することが可能である。上記各技術において基質は、均一且つ非多孔性の固体とするべきである（Sचना (1999) 前出）。推奨する基質には、シリコン、シリカ、スライドガラス、ガラスチップ及びシリコンウエハがある。或いは、ドットプロット法またはスロットプロット法に類似のアレイを利用して、熱的、紫外線的、化学的または機械的結合手順を用いて基質の表面にエレメントを配置及び結合させてもよい。通常のアレイは、手作業で、または利用可能な方法や機械を用いて作製でき、任意の適正数のエレメントを有し得る（Sचना, M. ら (1995) Science 270: 467 - 470、Shallon, D. ら (1996) Genome Res. 6: 639 - 645、Mars hall, A. および J. Hodgson (1998) Nat. Biotechnol. 16: 27 - 31. を参照）。

10

【0357】

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

20

【0358】

組織または細胞サンプルの調製

グアニジウムチオシアネート法を用いて組織サンプルから全RNAを単離し、オリゴ（dT）セルロース法を用いてポリ（A）⁺ RNAを精製する。各ポリ（A）⁺ RNAサンプルを、MMLV逆転写酵素、0.05 pg/μlのオリゴ（dT）プライマー（21mer）、1×第一鎖合成バッファー、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）を用いてポリ（A）⁺ RNA含有の25体積ml内で行う。特異的対照ポリ（A）⁺ RNAは、非コード酵母ゲノムDNAから*in vitro* 転写により合成する。各反応サンプル（1つはCy3、もう1つはCy5標識）は、2.5 mlの0.5 M水酸化ナトリウムで処理し、85 °Cで20分間インキュベートし、反応を停止させてRNAを分解させる。サンプルは、2つの連続するCHROMA SPIN 30ゲル濾過スピナラム（CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA）を用いて精製する。結合後、2つの反応サンプルは、1 mlのグリコーゲン（1 mg/ml）を用いて析出させたエタノール、60 mlの酢酸ナトリウム及び300 mlの100%エタノールである。サンプルは次に、Speed VAC（Savant Instruments Inc., Holbrook NY）を用いて乾燥して仕上げ、14 μlの5×SSC/0.2% SDS中で再懸濁する。

30

40

【0359】

50

マイクロアレイの調製

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

【0360】

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、処理中及び処理後に 0.1% の SDS 及びアセトン中で超音波をかけ、蒸留水で非常に良く洗って洗浄する。スライドガラスは、4% フッ化水素酸 (VWR Scientific Products Corporation (VWR), West Chester PA) 中でエッチングし、蒸留水で広範囲にわたって洗浄し、95% エタノール中で 0.05% アミノプロピルシラン (Sigma) を用いてコーティングする。コーティングしたスライドガラスは、110 のオープンで硬化させる。

10

【0361】

米国特許第 5,807,522 号で説明されている方法を用いて、コーティングしたガラス基板にアレイエレメントを付加する。該特許は、引用を以って本明細書の一部となす。平均濃度が 100 ng/µl のアレイエレメント DNA 1 µl を高速機械装置により開放型キャピラリープリンティングエレメント (open capillary printing element) に充填する。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加える。

20

【0362】

マイクロアレイには、STRATALINKER UV 架橋剤 (Stratagene) を用いて UV 架橋する。マイクロアレイは、室温において 0.2% SDS で 1 度洗浄し、蒸留水で 3 度洗浄する。リン酸緩衝生理食塩水 (PBS) (Tropix, Inc., Bedford MA) 中の 0.2% カゼイン中において 60 で 30 分間マイクロアレイをインキュベートした後、前に行ったように 0.2% SDS 及び蒸留水で洗浄することにより、非特異結合部位をブロックする。

30

【0363】

ハイブリダイゼーション

ハイブリダイゼーション反応は、5 × SSC, 0.2% SDS ハイブリダイゼーション緩衝液中の Cy3 及び Cy5 標識した cDNA 合成生成物を各 0.2 µg 含む 9 µl のサンプル混合体を有する。サンプル混合体は、65 まで 5 分間加熱し、マイクロアレイ表面上で等分して 1.8 cm² のカバーガラスで覆う。アレイは、顕微鏡スライドより僅かに大きい空洞を有する防水チェンバーに移す。チェンバーのコーナーに 140 µl の 5 × SSC を加えることにより、チェンバー内部を湿度 100% に保持する。アレイを含むチェンバーは、60 で約 6.5 時間インキュベートする。アレイは、第 1 洗浄緩衝液中 (1 × SSC, 0.1% SDS) において 45 で 10 分間洗浄し、第 2 洗浄緩衝液中 (0.1 × SSC) において 45 で 10 分間各々 3 度洗浄して乾燥させる。

40

【0364】

検出

レポーター標識ハイブリダイゼーション複合体は、Cy3 の励起のためには 488 nm、Cy5 の励起のためには 632 nm でスペクトル線を発生し得る Innova 70 混合ガス 10 W レーザ (Coherent, Inc., Santa Clara CA) を備えた顕微鏡で検出する。20 × 顕微鏡対物レンズ (Nikon, Inc., Melville NY) を用いて、アレイ上に励起レーザ光の焦点を当てる。アレイを含むスライドを顕微鏡のコンピュータ制御の X-Y ステージに置き、対物レンズを通過してラスタースキャンする。本実施例で用いた 1.8 cm × 1.8 cm のアレイは、20 µm の解像

50

度でスキャンした。

【0365】

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

10

【0366】

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

【0367】

光電子増倍管の出力は、IBMコンパチブルPCコンピュータにインストールされた12ビットRTI-835Hアナログ-デジタル(A/D)変換ボード(Analog Devices, Inc., Norwood MA)を用いてデジタル化される。デジタル化されたデータは、青色(低シグナル)から赤色(高シグナル)までの擬似カラー範囲へのリニア20色変換を用いてシグナル強度がマッピングされたようなイメージとして表示される。データは、定量的にも分析される。2つの異なる蛍光色素を同時に励起及び測定する場合には、各蛍光色素の発光スペクトルを用いて、データは先ず蛍光色素間の光学的クロストーク(発光スペクトルの重なり起因する)を補正する。

20

【0368】

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

30

【0369】

1.1 相補的ポリヌクレオチド

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

40

【0370】

1.2 DMEの発現

DMEの発現及び精製は、細菌若しくはウイルスを基にした発現系を用いて行うことができる。細菌でDMEが発現するために、抗生物質耐性及びcDNAの転写レベルを高める誘導性のプロモーターを含む好適なベクターにcDNAをサブクローニングする。このようなプロモーターには、lacオペレーター調節エレメントに関連するT5またはT7バ

50

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

【0371】

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

【0372】

1.3 機能的アッセイ

DMEの機能は、哺乳動物細胞培養系において生理学的に高められたレベルでのDMEをコードする配列の発現によって評価する。cDNAを、cDNAを高いレベルで発現する強いプロモーターを含む哺乳動物発現ベクターにサブクローニングする。選り抜きのベクターには、pCMVSPORTプラスミド(Life Technologies)及びpCR3.1プラスミド(Invitrogen, Carlsbad CA)が含まれ、どちらもサイトメガロウイルスプロモーターを有する。リポソーム製剤或いは電気穿孔法を用いて、5~10µgの組換えベクターをヒト細胞株、例えば内皮由来または造血由来の細胞株に一時的に形質移入する。更に、標識タンパク質をコードする配列を含む1~2µgのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入細胞と非形質移入細胞を区別する手段が与えられる。また、標識タンパク質の発現によって、cDNAの組換えベクターからの発現を正確に予想できる。標識タンパク質は、例えば緑色蛍光タンパク質(GFP; Clontech)、CD64またはCD64-GFP融合タンパク質から選択できる。自動化された、レーザー光学に基づく技術であるフローサイトメトリー(FCM)を用いて、GFPまたはCD64-GFPを発現する形質移入された細胞を同定し、その細胞のアポトーシス状態や他の細胞特性を評価する。FCMは、細胞死に先行するか或いは同時に発生する現象を診断する蛍光分子の取込を検出して計量する。このような現象として挙げられるのは、プロピジウムヨウ化物によるDNA染色によって計測される核DNA内容物の変化、前方光散乱と90°側方光散乱によって計測

される細胞サイズと顆粒状性の変化、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及び細胞内におけるタンパク質の発現の変化、及び蛍光複合アネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変化とがある。フローサイトメトリー法については、Ormerod, M. G. (1994) Flow Cytometry Oxford, New York, NY. に記述がある。

【0373】

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

10

【0374】

1.4 DMEに特異的な抗体の作製

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

20

【0375】

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

【0376】

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

30

【0377】

1.5 特異的抗体を用いる天然DMEの精製

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

40

【0378】

DMEを含む培養液をイムノアフィニティークラムに通し、DME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DMEとの結合を切るような条件で(例えば、pH2~3のバッファー、或いは高濃度の尿素またはチオシアン酸塩イオンのようなカオトロピ

50

ックイオンで)溶出させ、DMEを回収する。

【0379】

16 DMEと相互作用する分子の同定

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

10

【0380】

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

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

20

【0381】

17 DME活性の実証

DMEのチトクロームP450はアニリンの4-ヒドロキシル化を用いて測定する。アニリンはこの酵素によって4-アミノフェノールに変換され、630nmで最大吸光度を有する(Gibson および Skett, 前出)。このアッセイは便利な測定ではあるが、2- および3-位置でも発生する全ヒドロキシル化が過小評価される。アッセイは37で行い、反応緩衝液には一定分量の酵素および適量のアニリン(約2mM)が含まれる。この反応の場合、緩衝液はNADPH またはNADPH補酵素産生システムを含める必要がある。この緩衝液の一つの処方には85mM トリス(pH 7.4)、15mM 塩化マグネシウム、50mM ニコチン酸アミド、40mg イソクエン酸三ナトリウムおよび2ユニットのイソクエン酸デヒドロゲナーゼが含まれており、アッセイの直前に10mlの反応貯蔵緩衝液に8mgの NADP^+ を加える。反応は光学キュベットで行い、630nmの吸光度を測定する。このアッセイでは、吸光度の増加率が酵素活性に比例する。標準曲線は4-アミノフェノールの既知濃度を用いて作成できる。

30

【0382】

DMEの1,25-ジヒドロキシビタミンD₂₄-ヒドロキシラーゼ活性を、DMEを発現する遺伝子組換えラットにおける ^3H 標識された1,25-ジヒドロキシビタミンD(1,25(OH)₂D)の24,25-ジヒドロキシビタミンD(24,25(OH)₂D)への変換をモニタリングして決定する。エタノール(或いはコントロールとしてエタノールのみ)に溶解した1μgの1,25(OH)₂Dを、DMEの破壊変異体を発現する若しくはDMEを発現していないという点を除けばコントロールラットと同一のDMEを発現する約6週目のオス遺伝子組換えラットに静脈内投与する。ラットを8時間後に断頭により屠殺し、腎臓を手早く取り除き、洗浄し、9倍量の氷冷バッファ(15mM Tris-acetate(pH 7.4)、0.19Mスクロース、2mM酢酸マグネシウム、および5mMコハク酸ナトリウム)においてホモジナイズする。次に、各ホモジネートの所定量(例えば3ml)を、約3.5GBq/mmolの特定の活性を有する0.25nM 1,25(OH)₂[1- ^3H]Dにおいて、常に振盪しながら酸素存在下、37で、15分間インキュベートする(Bligh, E. G. およびDyer, W. J. (1959) Can. J. Biochem. Physiol. 37

40

50

: 911 - 917)。次にクロロホルム相を、1 ml/minの流速でn - ヘキサン/クロロホルム/メタノール (10 : 2.5 : 1.5) の溶液を含むFINEPAK SILカラム (JASCO, Tokyo, Japan) を用いてHPLCで分析する。別法では、クロロホルム相は、1 ml/minの流速でアセトニトリルバッファ系 (40 ~ 100, 水中で30分間) でJ SPHERE ODS - AMカラム (YMC Co. Ltd., Kyoto, Japan) を用いて逆相HPLCで分析する。溶出液を30秒 (或いは30秒未満) 毎に集め、その容量における³Hの存在量をシンチレーションカウンタで測定する。コントロールサンプル (すなわち、1, 25 - ジヒドロキシビタミンドまたは24, 25 - ジヒドロキシビタミンド (24, 25 (OH)₂D) を含むサンプル) のクロマトグラムと反応生成物のクロマトグラムとを比較し、基質 (24, 25 (OH)₂D) および生成物 (24, 25 (OH)₂ [1 - ³H]D) の相対的な移動度を決定し、その値が収集した容量における存在量と相関性を有する。コントロールラットにおいて生成された24, 25 (OH)₂ [1 - ³H]Dの量を、DMEを発現する遺伝子組換えラットの24, 25 (OH)₂ [1 - ³H]Dの量から差し引く。遺伝子組換えラットとコントロールラットとの24, 25 (OH)₂ [1 - ³H]Dの生成物の差が、そのサンプルに存在するDMEの25 - ヒドラーゼ活性のレベルに比例する。基質および生成物の同定は、質量分光法によって行う (Miyamoto, Y. 他 (1997) J. Biol. Chem. 272 : 14115 - 14119)。

10

【0383】

DMEのフラビン含有モノオキシゲナーゼ活性を代謝産物のクロマトグラフ分析によって測定する。例えば、Ring, B. J. 他 (1999 ; Drug Metab. Dis. 27 : 1099 - 1103) は、0.1 Mリン酸ナトリウムバッファ (pH 7.4 または 8.3) および1 mM NADPHにおいてFMOを37 でインキュベートし、有機溶剤で反応を停止させ、HPLCにより生成物の形成を決定する。別法では、活性はクラーク型電極を用いて炭素の取り込み量をモニタリングして測定する。例えば、Ziegler, D. M. およびPoulsen, L. L. (1978 ; Methods Enzymol. 52 : 142 - 151) は、基質メチマゾールを含むNADPH補酵素産生システム (上記したものに類似) において37 でこの酵素をインキュベートした。

20

【0384】

酵素の取り込み割合が酵素活性に比例する。DMEのUDPグルコニルトランスフェラーゼ活性を、遊離アミノ基の比色測定を用いて測定する (GibsonおよびSkett、前出)。2 - アミノフェノールなどのアミン含有基質は、必要な補助因子を含む反応バッファ (40 mM Tris pH 8.0, 7.5 mM MgCl₂, 0.025 Triton X - 100, 1 mM アスコルビン酸, 0.75 mM UDP - グルクロン酸) において酵素のアリコットと共に37 でインキュベートする。十分な時間が経過した後、0.1 Mリン酸バッファ pH 2.7 に氷冷した20%トリクロ酢酸を加えて反応を停止させ、氷上でインキュベートし、遠心分離器にかけて上澄みを除去する。反応しなかった全ての2 - アミノフェノールをこのステップで破壊する。次に、新しく準備した亜硝酸ナトリウムを加え、グルクロン酸抱合体であるジアゾニウム塩が形成されるようにする。過剰な亜硝酸は、十分なスルファミン酸アンモニウムを加えて除去し、ジアゾニウム塩を芳香族アミン (例えば、N - ナフチルエチレンジアミン) と反応させて、分光光度計でアッセイ可能な (例えば540 nmにおいて) 着色アゾ化合物を生成する。2 - アミノフェノールグルクロニドと特性が類似した発色団を形成するアニリンの既知の濃度を用いて標準的な曲線を作成することができる。

30

40

【0385】

DMEのスルホトランスフェラーゼ活性を、[³⁵S]PAPSからフェノールなどのモデル基質内への³⁵Sの取り込みを用いて測定する (Folds, A. およびMeek, J. L. (1973) Biochim. Biophys. Acta 327 : 365 - 374)。酵素のアリコットを1 mLの10 mMリン酸バッファ、pH 6.4、

50

50 μ M フェノール、および 0.4 ~ 4.0 μ M [³⁵S] PAPS で 37 °C でインキュベートする。放射標識の 5 ~ 20 が基質に転移するのに十分な時間が経過した後、0.2 mL の 0.1 M 酢酸バリウムを加えタンパク質およびリン酸バッファを沈降させる。次に 0.2 mL の 0.1 M Ba(OH)₂ を加え、その後 0.2 mL の ZnSO₄ を加える。遠心分離して上澄みを除いて、タンパク質および未反応の [³⁵S] PAPS を除去する。上澄みの放射活性をシンチレーションカウンタで測定する。酵素活性は、反応生成物における放射活性のモル数から決定する。

【0386】

DME のグルタチオン S - トランスフェラーゼ活性を、340 nm における最大吸光率を有するモデル基質を用いて測定する。このモデル基質は、グルタチオンと反応して 2, 4 - ジニトロフェニルグルタチオンを生成する 2, 4 - ジニトロ - 1 - クロロベンゼンなどである。GST が様々な基質特異性を有することは重要であり、モデル基質は目的の GST の基質選択制に基づいて選択しなければならない。アッセイは室温で行われ、酵素のアリコットを好適な反応バッファ（例えば 1 mM グルタチオン、1 mM ジニトロクロロベンゼン、90 mM リン酸カリウムバッファ、pH 6.5）に含める。反応は蛍光キュベットで行い、340 nm における吸光度を測定する。吸光度の上昇の割合がアッセイにおける酵素活性に比例する。

10

【0387】

DME の N - アシルトランスフェラーゼ活性を、放射標識したアミノ酸基質を用い、抱合体の中に取り込まれた放射標識を測定して決定する。酵素を、無標識のアシル - CoA 化合物および放射線標識したアミノ酸を含む反応バッファにおいてインキュベートし、放射線標識アシル抱合体を n - ブタノールまたはその他の好適な有機溶剤の中に抽出して反応しなかったアミノ酸から分離する。例えば、Johnson, M. R. 他 (1990; J. Biol. Chem. 266: 10227 - 10233) が、この酵素をコリル CoA (cholylyl - CoA) および ³H グリシンまたは ³H タウリンと共にインキュベートしてトリチウム化コール酸抱合体を n - ブタノールの中に抽出し、抽出した生成物の中の放射活性をシンチレーションカウンタで測定して、胆汁酸 CoA : アミノ酸 N - アシルトランスフェラーゼ活性を決定した。別法では、N - アシルトランスフェラーゼ活性を、分光光度計で後述する還元型 CoA (CoASH) を測定して決定する。

20

【0388】

DME の N - アセルトランスフェラーゼ活性を、[¹⁴C] アセチル CoA から基質分子への放射標識の転移を利用して測定する（例えば、Deguchi, T. (1975) J. Neurochem. 24: 1083 - 5 を参照）。別法では、CoASH と DTNB (5, 5' - ジチオビス(2 - ニトロ安息香酸)；エルマン試薬) の反応に基づいた新しい分光光度アッセイが用いられ得る。N - アセルトランスフェラーゼの触媒によるアセチル基の基質への転移の間に、遊離チオール含有 CoASH が生成される。CoASH は、412 nm における DTNB 複合体の吸光度を用いて決定される (De Angelis, J. 他 (1997) J. Biol. Chem. 273: 3045 - 3050)。酵素活性は基質の中に取り込まれた放射活性の割合、または分光光度アッセイにおける吸収率の上昇の割合に比例する。

30

40

【0389】

DME のタンパク質アルギニンメチルトランスフェラーゼ活性を 37 °C で様々な時間にわたり測定する。S アデノシル L [メチル³H] メチオニン ([³H] AdoMet; 特異的活性 = 75 Ci/mmol; NEN Life Science Products) はメチル供与体基質として用いられる。有用なメチル受容基質にはグルタチオン S - トランスフェラーゼフィブリラリングリシン - アルギニンドメイン融合タンパク質 (GST - GAR)、不均一核内リボタンパク質 (hnRNP) または、アデノシンジアルデヒドで処理した細胞からの可溶化液中に存在する低メチル化タンパク質が含まれる。メチル化反応は SDS PAGE サンプルバッファを加えると停止する。反応の産物は SDS PAGE によって分離され、蛍光間接撮影法によって視覚化される。³H 標識メチ

50

ル供与体基質の存在は、DMEのタンパク質アルギニンメチルトランスフェラーゼ活性の標示となる (Tang, J. ら. (2000) J. Biol. Chem. 275 : 77237730 および Tang, J. ら (2000) J. Biol. Chem. 275 : 1986619876)。

【0390】

DMEのアルド/ケト還元酵素活性を、NADPHが消費される時の340nmにおける吸光率の低下を測定する。標準的な反応混合液は、135mMのリン酸ナトリウムバッファ(酵素によりpH6.2~7.2の範囲)、0.2mM NADPH、0.3M硫酸リチウム、0.5~2.5µg酵素、および好適なレベルの基質を含む。この混合液を30でインキュベートし、反応を分光光度計で連続的に測定する。酵素活性は、酵素1µg当たり消費されるNADPHのモル数として計算される。 10

【0391】

DMEのアルコールデヒドロゲナーゼ活性を、NAD⁺がNADHに還元される時の340nmにおける吸光率の増大を利用して測定する。標準的な反応混合液は、50mMのリン酸ナトリウム(pH7.5)、および0.25mM EDTAである。反応液を25でインキュベートし、分光光度計でモニタリングする。酵素活性は、酵素1µg当たり生成されたNADHのモル数として計算する。

【0392】

DMEの活性を、基質として4-酢酸メチルウンベルフェリル(4-methylumbelliferyl acetate)を用いて測定する。酵素反応は、約10µlのDME含有サンプルを、0.5mM 4-酢酸メチルウンベルフェリルを含む1mlの反応バッファ(90mM KH₂PO₄, 40mM KCl, pH7.3)に加えて開始させる。4-メチルウンベルフェロンの生成を分光光度計($\epsilon_{350} = 12.2 \text{ mM}^{-1} \cdot \text{cm}^{-1}$)で1分30秒モニタリングする。酵素活性は、1分間に1mgのタンパク質で生成される生成物のµモルとして表すことができ、サンプルにおけるDME活性に比例する (Evgenia, V. 他 (1997) J. Biol. Chem. 272 : 14769 - 14775)。 20

【0393】

別法では、DMEのコカインベンゾイルエステルヒドロラーゼ(cocaine benzoyl ester hydrolase)活性を、1mM ベンズアミジン(benzamidine)、1mM EDTA、および1mM ジチオスレイトールを含む反応バッファ(50mM NaH₂PO₄, pH7.4)中で約0.1mlの酵素および3.3mM コカインを37でインキュベートして測定する。全量で0.4mlの反応液を1時間インキュベートし、等量の5%トリクロロ酢酸で反応を終了させる。0.1mlの内部標準3,4-ジメチル安息香酸(10µg/ml)を加える。沈降したタンパク質を10分間12,000×gで遠心分離して分離する。上澄みを新しいチューブに移して、0.4mlの塩化メチレンで2回抽出する。2つの抽出物を混ぜ、窒素を流して乾燥させる。この残滓を、100ml当たり8µlのジエチルアミンを含む14アセトニトリル、250mM KH₂PO₄、pH4.0に再懸濁してから、C18逆相HPLCカラム上に注入して分離する。このカラム溶出液を235nmでモニタリングする。DME活性は、内部標準に対する分析物のピーク面積の割合を比較して定量する。標準曲線は、トリクロロ酢酸処理タンパク質基質内に準備した安息香酸標準で作成する (Evgenia, V. 他 (1997) J. Biol. Chem. 272 : 14769 - 14775)。 30 40

【0394】

別法では、水溶性基質のパラニトロフェニル酪酸に対するDMEカルボキシルエステラーゼ活性を当分野で周知の分光光度計法で測定する。この方法では、DME含有サンプルを、6mMのタウロコール酸塩の存在下で0.5M トリスHCl(pH7.4または8.0)または酢酸ナトリウム(pH5.0)で希釈する。このアッセイは、新たに準備したパラニトロフェニル酪酸溶液(酢酸ナトリウムに100µg/ml, pH5.0) 50

を加えて開始する。次にカルボキシルエステラーゼ活性をモニタリングし、405 nm にセットした分光光度計で基質のコントロール自己加水分解と比較する (Wan, L. 他 (2000) J. Biol. Chem. 275:10041-10046)。

【0395】

DMEのヘパラン硫酸6-スルホトランスフェラーゼ活性を、DMEを含むサンプルを、2.5 μmol イミダゾールHC1 (pH 6.8)、3.75 μgの塩化プロタミン (protamine chloride)、25 nmolの完全に脱硫酸化されN-再硫酸化されたヘパリン(ヘキソサミンとして)、および50 pmol (約 5×10^5 cpm)の $[^3\ ^5\text{S}]$ アデノシン3'-リン酸5'-ホスホ硫酸(PAPS)と共に最終反応容量50 μlで37 °Cで20分間インキュベートし、*in vitro*で測定する。この反応は、熱湯に反応チューブを1分間浸漬して停止させる。0.1 μmolのコンドロイチン硫酸A(グルクロン酸として)を、キャリアとして反応混合液に加える。 $^3\ ^5\text{S}$ 標識ポリ多糖を、1.3 酢酸カリウムを含む冷却した三倍量のエタノールで沈殿させ、脱塩カラムを用いてゲルクロマトグラフィによって取り込まれなかった $[^3\ ^5\text{S}]$ PAPSおよびその分解生成物から完全に分離する。一単位の酵素活性を1分間に1 pmolの硫酸を移送するのに必要な量と定義し、沈殿した多糖の中に取り込まれた $[^3\ ^5\text{S}]$ PAPSの量によって測定する (Habuchi, H. 他 (1995) J. Biol. Chem. 270:4172-4179)。

10

別法では、DMEのヘパラン硫酸6-スルホトランスフェラーゼ活性を、ドデシル硫酸ナトリウムポリアクリアミドゲル電気泳動法(SDS-PAGE)によって分離した後、ゲルから酵素を抽出して再生し測定する。分離した後、ゲルをバッファ(0.05 M Tris-HCl, pH 8.0)で洗浄し、3~5 mmのセグメントに切断し、0.15 M NaClを含む同じバッファを100 μlで4 °Cで48時間攪拌する。溶出した酵素を遠心分離して収集し、上記したように、スルホトランスフェラーゼ活性についてアッセイする (Habuchi, H. 他 (1995) J. Biol. Chem. 270:4172-4179)。

20

【0396】

別法では、DMEのスルホトランスフェラーゼ活性を、 $[^3\ ^5\text{S}]$ PAPSから固定されたペプチドへの $[^3\ ^5\text{S}]$ 硫酸の転移を測定して決定する。ある例においては、ペプチド(QATEYELDYLDFLEEC)は、C末端システイン残基が付加された成熟P-セレクチン糖タンパク質リガンド1ポリペプチドのN末端15残基を表す。このペプチドは3つの潜在的なチロシン硫酸化部位に渡っている。このペプチドは、システイン残基によってヨードアセトアミド活性化樹脂に結合される(1 mlの樹脂あたり1.5~3.0 μmolペプチドの密度)。この酵素アッセイは、140 mM Pipes (pH 6.8)、0.3 M NaCl、20 mM MnCl₂、50 mM NaF、1 Triton X-100、および1 mM 5'-AMPを含む最終容量130 μlにおいて10 μlのペプチド誘導体ビーズと2~20 μlのDME含有サンプルを混合させて行った。このアッセイは、0.5 μCiの $[^3\ ^5\text{S}]$ PAPS (1.7 μM; 1 Ci = 37 GBq)を加えて開始させた。37 °Cで30分経過した後に、反応ビーズを65 °Cで6 M グアニジンで洗浄し、ビーズに取り込まれた放射活性を液体シンチレーションカウンタで測定する。ビーズ結合ペプチドに転移された $[^3\ ^5\text{S}]$ 硫酸を測定し、サンプルにおけるDME活性を決定する。1単位の活性は1分間に1 pmolの生成物が生成されると定義する (Ouyang, Y-B. 他 (1998) Biochemistry 95:2896-2901)。

30

40

【0397】

別法では、DMEスルホトランスフェラーゼのアッセイを、50 mM Hepes-NaOH (pH 7.0)、250 mM スクロース、1 mM ジチオスレイトール、14 μM $[^3\ ^5\text{S}]$ PAPS (15 Ci/mmol)、およびドーパミン(25 μM)、-ニトロフェノール(5 μM)またはその他の候補基質を含む最終容量30 μlにおいて硫酸供与体として $[^3\ ^5\text{S}]$ PAPSを用いて行った。アッセイの反応は、精製されたDME酵

50

素試薬或いはDME活性を含むサンプルを加えて開始させ、その反応を37℃で15分間続け、1000rpmで3分間加熱して停止させる。形成した沈降物を遠心分離によって除去する。次に³⁵S硫酸化物を調べるために、上澄みを薄膜クロマトグラフィ或いは二次元薄膜分離方法のいずれかによって分析する。³⁵S硫酸化物の同定ができるように好適な標準を上澄みと平行に流し、反応産物の移動の相対速度に基づいてDME含有サンプルの酵素特異性を決定する(Sakakibara, Y.他(1998) J. Biol. Chem. 273:6242-6247)。

【0398】

DMEのスクアレンエポキシラーゼ活性を、精製したDME(またはDMEを含む未精製の混合液)、20mM Tris-HCl(pH7.5)、0.01mM FAD、0.2単位のNADPH-チトクロームC(P-450)レダクターゼ、0.01mM [¹⁴C]スクアレン(20μlのTween 80を用いて拡散された)、および0.2 Triton X-100を含む混合液においてアッセイする。1mM NADPHを加えて反応を開始させ、37℃で30分間インキュベートする。非けん化脂質を、酢酸エチル/ベンゼン(0.5:99.5, v/v)で作製したシリカゲルTLCによって分析する。反応生成物を、DMEを含まない反応混合液による反応生成物と比較する。2,3(S)-オキシドスクアレンの存在を、好適な脂質標準を用いて確認する(Sakakibara, J.他(1995) 270:17-20)。

【0399】

DMEのエポキシドヒドロラーゼ活性を、エーテル抽出物のガスクロマトグラフィ(GC)分析を用いて基質を除去した後、或いはアセトンで失活させた反応混合液をGC分析により基質を除去してジオールを生成した後に測定する。DMEを含むサンプルまたはエポキシドヒドロラーゼコントロールサンプルを、10mM Tris-HCl(pH8.0)、1mM エチレンジアミンテトラ酢酸(EDTA)、および5mM エポキシド基質(例えば、エチレンオキシド、スチレンオキシド、プロピレンオキシド、一酸化イソプレン(isoprene monoxide)、エピクロロヒドリン、エピプロモヒドリン、エピフルオロヒドリン(epifluorohydrin)、グリシドール、1,2-エポキシブタン、1,2-エポキシヘキサン、または1,2-エポキシオクタン)でインキュベートする。様々な時点でサンプルの一部を反応混合液から取り出して、内部標準を含む1mlの氷冷アセトンに加えてGC分析を行う(例えば、1-ノナノール)。タンパク質および塩を遠心分離(15分、4000×g)によって除去し、抽出物を0.2mm×25m CP-Wax 57-CBカラム(CHROMPACK, Middelburg, The Netherlands)および水素炎イオン化検出器を用いてGCにより分析する。GC産物の同定は、当分野で周知の好適な標準およびコントロールを用いて行う。1単位のDME活性は、1分間に1μmolのジオール生成を触媒する酵素の量と定義する(Rink, R.他(1997) J. Biol. Chem. 272:14650-14657)。

【0400】

DMEのアミノトランスフェラーゼ活性は、DMEを含むサンプルを、1mM L-キヌレニンおよび1mM 2-オキソグルタル酸の存在下で、最終容量200μlの70μM PLPを含む150mM Tris 酢酸バッファ(pH8.0)で37℃で1時間インキュベートしてアッセイする。キヌレニン酸の生成は、当分野で周知の好適な標準およびコントロールを用いて330nmで分光光度検出によりHPLCで定量する。別法では、L-3-ヒドロキシキヌレニンを基質として使い、キサンツレン酸の生成を340nmでのUV検出で生成物のHPLC分析により測定する。キヌレニン酸およびキサンツレン酸の生成はそれぞれ、アミノトランスフェラーゼ活性を示す(Buchli, R.他(1995) J. Biol. Chem. 270:29330-29335)。

【0401】

別法では、DMEのアミノトランスフェラーゼ活性の測定は、酵素結合補助因子であるPLPのUV/VIS吸収スペクトルにおける変化をモニタリングして、一回のターンオー

バーの条件下で精製したDMEのサンプル或いはDMEを含む未精製のサンプルの様々なアミノ酸およびオキソ酸基質に対する活性をして決定して行う。この反応は、9 μ M 精製DME またはサンプル含有DMEおよび検査する基質(アミノ酸およびオキソ酸基質)を含む50 mM 4-メチルモルフォリン(pH 7.5)において25 で行う。アミノ酸からオキソ酸への半反応の後に、酵素結合PLPのPMPへの変換により生じる360 nmにおける吸収率の低下および330 nmにおける吸収率の増加を測定する。DMEの特異性および相対的な活性を、特定の基質に対する酵素活性によって測定する(Vacca, R. A. 他(1997) J. Biol. Chem. 272: 21932-21937)

DMEのスーパーオキシドジスムターゼ活性を、細胞ペレット、培養した上澄み、または精製したタンパク質試薬からアッセイする。サンプルすなわち溶解物を15%非変性ポリアクリルアミドゲル上での電気泳動法によって分離する。このゲルを、2.5 mM ニトロブルーテトラゾリウムにおいて30分間インキュベートし、その後30 mM リン酸カリウム、30 mM TEMED、および30 μ M リボフラビン(pH 7.8)において20分間インキュベートする。スーパーオキシドジスムターゼ活性は、背景のブルーに対して白いバンドとして見ることができ、ライトボックスでゲルに照明を当てる。スーパーオキシドジスムターゼ活性の定量は、好適なスーパーオキシドジスムターゼのポジティブおよびネガティブコントロール(例えば、様々な量の市販されている大腸菌スーパーオキシドジスムターゼなど)を用いて活性ゲルのデンシトメータースキャンにより行う(Harth, G. および Horwitz, M. A. (1999) J. Biol. Chem. 274: 4281-4292)。

【0402】

DMEのカテコール-O-メチルトランスフェラーゼ活性を、50 mM Tris-HCl (pH 7.4)、1.2 mM MgCl₂、200 μ M SAM (S-アデノシル-L-メチオニン) ヨウ化物(0.5 μ Ciのメチル-[³H]SAMを含む)、1 mM ジチオスレイトール、および様々な濃度のカテコール基質(例えば、レボドパ、ドーパミン、またはDBA)から成る最終容量が1.0 mlの反応液において測定する。この反応は、精製したDMEまたは未精製のDMEを含むサンプル250~500 μ gを加えて開始し、37 で30分間行った。この反応は、氷上で素早く冷却して停止させ、直後に氷冷n-ヘプタンで抽出する。次に10分間1000 x gで遠心分離した後、液体シンチレーションカウンタで有機抽出物の3 mlのアリコットをその中に含まれている放射活性について分析した。有機相におけるカテコール関連放射活性のレベルがDMEのカテコール-O-メチルトランスフェラーゼ活性に比例する(Zhu, B. T. Liehr, J. G. (1996) 271: 1357-1363)。

【0403】

DMEのDHFR活性を、340 nm ($\epsilon_{340} = 11,800 \text{ M}^{-1} \cdot \text{cm}^{-1}$)においてNADPHの消滅の後に15 で分光光度法で測定する。標準的なアッセイ混合液は、100 μ M NADPH、14 mM 2-メルカプトエタノール、MTEN バッファ(50 mM 2-morpholinoethanesulfonic acid、25 mM tris(hydroxymethyl)aminomethane、25 mM エタノールアミン、および100 mM NaCl、pH 7.0)を含み最終容量を2.0 mlとする。反応は50 μ Mのジヒドロ葉酸(基質として)を追加して開始させる。反応液におけるNADPHのNADP⁺への酸化がジヒドロ葉酸の還元に一致し、サンプルにおけるDHFR活性の量に比例する(Nakamura, T. および Iwakura, M. (1999) J. Biol. Chem. 274: 19041-19047)。

【0404】

DMEのスルホトランスフェラーゼ活性を、[³⁵S]PAPSからフェノールなどのモデル基質内への³⁵Sの取り込みを用いて測定する(Folds, A. および Meek, J. L. (1973) Biochim. Biophys. Acta 327: 365-374)。酵素のアリコットを1 mLの10 mM リン酸バッファ、pH 6.4、

50 μ M フェノール、および 0.4 ~ 4.0 μ M [35 S] PAPS で 37 でインキュベートする。放射標識の 5 ~ 20 が基質に転移するのに十分な時間が経過した後、0.2 mL の 0.1 M 酢酸バリウムを加えタンパク質およびリン酸バッファを沈降させる。次に 0.2 mL の 0.1 M Ba(OH)₂ を加え、その後 0.2 mL の ZnSO₄ を加える。遠心分離して上澄みを除き、タンパク質および未反応の [35 S] PAPS を除去する。上澄みの放射活性をシンチレーションカウンタで測定する。酵素活性は、反応生成物における放射活性のモル数から決定する。

【0405】

1.8 DME インヒビターの同定

実施例 17 のアッセイで説明したように、検査されるべき化合物を、好適なバッファおよび基質と共に様々な濃度でマルチウェルプレートのウェルに注入する。DME 活性はそれぞれのウェルについて測定し、DME 活性を阻害するそれぞれの化合物の能力および用量反応動態 (dose-response kinetics) を決定することができる。DME 活性を促進する分子の同定にも、このアッセイを用いることができる。

【0406】

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

【0407】

(表の簡単な説明)

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

【0408】

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

【0409】

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

【0410】

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

【0411】

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

【0412】

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

【0413】

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

【表 1】

10

20

30

40

表 1

Incyte プロセス カト ID	ポリペプチド SEQ ID NO.	Incyte ポリペプチド ID	ポリペプチド SEQ ID NO.	Incyte ポリペプチド ID
7248285	1	7248285CD1	20	7248285CB1
7472835	2	7472835CD1	21	7472835CB1
7476203	3	7476203CD1	22	7476203CB1
7478583	4	7478583CD1	23	7478583CB1
7478585	5	7478585CD1	24	7478585CB1
7479904	6	7479904CD1	25	7479904CB1
7480367	7	7480367CD1	26	7480367CB1
8069390	8	8069390CD1	27	8069390CB1
7473869	9	7473869CD1	28	7473869CB1
7478588	10	7478588CD1	29	7478588CB1
55046125	11	55046125CD1	30	55046125CB1
3538709	12	3538709CD1	31	3538709CB1
71563101	13	71563101CD1	32	71563101CB1
7472027	14	7472027CD1	33	7472027CB1
7480358	15	7480358CD1	34	7480358CB1
1618256	16	1618256CD1	35	1618256CB1
3387823	17	3387823CD1	36	3387823CB1
55142051	18	55142051CD1	37	55142051CB1
7395274	19	7395274CD1	38	7395274CB1

10

20

30

40

【表 2】

表2-1

ポリバブチド SEQ ID NO:	Incyte ポリバブチド ID	GenBank ID NO:	確率スコア	GenBank 相同体
1	7248285CD1	94929199	1.90E-248	[アミノ酸産黄色ススキ] リシオルオキシダナーゼ関連タンパク質相同体 Langenau, D.M. 他 (1999) J. Mol. Endocrinol. 23:137-152.
2	7472835CD1	9559046	4.50E-181	[ヒト] フラビン含有モノオキシゲナーゼ 5 Overby, L.H. 他 (1995) Arch. Biochem. Biophys. 317:275-284.
3	7476203CD1	92653663	5.30E-91	[ハツカネズミ] チトクローム F450 モノオキシゲナーゼ
4	7478583CD1	914475602	3.00E-72	[E1] [セラウス菌 (Bacillus cereus)] アセチルアセトインレダクターゼ
4	7478583CD1	9529564	2.00E-61	■ [緑膿菌 (Pseudomonas putida)] 2,3-ブタンジオールデヒドロゲナーゼ Huang, M. 他 (1994) FEMS Microbiol. Lett. 124:141-150.
5	7478585CD1	914582559	1.00E-125	[E1] [黄色ブドウ球菌 (Staphylococcus aureus)] (AF282920) キノンオキシンドレダクターゼ
6	7479904CD1	96683558	2.30E-255	[ワカネズミ] ヘパラン硫酸 6-スルホトランスフェラーゼ 2

10

20

30

40

表2-2

ポルパブチドSEQ ID NO:	Incyte ポルパブチド ID	GenBank ID NO:	確率スコア	GenBank 相同体
7	7480367CD1	9242749	1.30E-101	[ヒト] グルタチオンSトランスフェラーゼ Ha1サブユニット (EC 2.5.1.18) Rozen, F. 他 (1992) Arch. Biochem. Biophys. 292:589-593.
8	8069390CD1	95531815	1.90E-56	[ヒト] ステロイドデヒドロゲナーゼ相同体
8	8069390CD1	913183088	1.00E-179	[f1] [ヒト] ステロイドデヒドロゲナーゼ様タンパク質
9	747869CD1	914331125	2.00E-78	[f1] [イヌ] カルボキシシルエステラーゼ D1
10	747858CD1	914331125	1.00E-140	[f1] [イヌ] カルボキシシルエステラーゼ D1
11	55046125CD1	96683558	1.80E-236	[ハツカネズミ] ヘパリン硫酸 6-スルホトランスフェラーゼ 2
12	3538709CD1	92633182	1.30E-11	[枯草菌 (Bacillus subtilis)] エボキシドヒドラーゼに類似
12	3538709CD1	92649734	2.00E-14	Kunst, F. 他 Nature 390:249-256. [高熱性硫黄古細菌 (Archaeoglobus fulgidus)] カルボキシシルエステラーゼ (est-1)
13	71563101CD1	91944136	1.10E-234	[ドブネズミ] ベータアラニンゼルビン酸アミノトランスフェラーゼ
14	7472027CD1	91151081	5.40E-34	[モルモット] ヒドロキシステロイドスルホトランスフェラーゼ 2 Luu, N.X. (1995) Biochem. Biophys. Res. Commun. 217:1078-1086.

10

20

30

40

表2-3

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	確率スコア	GenBank 相同体
15	7480358CD1	91066120	2.10E-25	[モルモット] 銅亜鉛スーパーオキシドジスムターゼ Yuan, H.T. 他 (1996) Biochim. Biophys. Acta 1305:163-171.
16	1618256CD1	910880453	2.00E-68	[ヒト] 糖タンパク質ペーダー Gal 3 リズルホトランスフェラーゼ
17	3387823CD1	96782275	3.00E-99	[線虫 (Caenorhabditis elegans)] モノアミンオキシダーゼに類似-cDNA EST The C. elegans Sequencing Consortium. Genome sequence of the nematode C. elegans: a platform for investigating biology. Science 282:2012-2018.
18	55142051CD1	9179955	9.50E-41	[ヒト] カテコール Oメチルトランスフェラーゼ Bertocci, B. 他 (1991) Human catechol-O-methyltransferase: cloning and expression of the membrane-associated form. Proc. Natl. Acad. Sci. U.S.A. 88:1416-1420.

10

20

30

40

表2-4

ポリペプチド SEQ ID NO:	Incyte ポリペプチド ID	GenBank ID NO:	標準入ロア	GenBank 相同体
19	7395274CD1	95852342	4.60E-99	【マミチヨク】(Fundulus heteroclitus) マトク ローム P450 2M1 Oleksiak, M.F. 他(2000) Identification, functional characterization, and regulation of a new cytochrome P450 subfamily, the CYP2Ns. J. Biol. Chem. 275:2312- 2321.

10

20

30

40

表3-1

SEQ ID NO.	Incyte ポリブチド ID	アミノ酸残基数	潜在的に酸化部位	潜在的グリコシル化部位	シグネチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
1	7248285CD1	756	T29, S39, T56, S85, S107, S123, S167, T239, S304, T334, S356, T386, T484, S498, S527, S531, T600, T621, T673	N198, N629	オキシダーゼ; リジン; リシル; 銅ドメイン; DM04978 Q05063 I-419; Q505-C731 リシルオキシダーゼ タンパク質リジン 前駆体 シグナル 6 オキシダーゼ オキシドレダクターゼ 糖タンパク質 相同体; PD012364; P533-C731 リシルオキシダーゼシグネチャ: PR00074; R593-T621, T621-C648, D679-E707, S708-N735 シグナルペプチド (signal_peptide): M1-P24 リシルオキシダーゼ Lysyl oxidase: P533-S736 シグナル切断: M1-P24	BLAST-DOMO BLAST-PRODOM BLIMPS-PRINTS HMMER HMMER-PFAM SFSCAN
2	7472835CD1	544	T30, S49, T63, S113, S179, S184, S196, S215, T216, S220, Y231, T285, S296, S304, T309, S318, T319, S338, T347, Y421, T424, T491, T505	N47, N62, N489	ジメチルアリンモノオキシゲナーゼ(N-OXIDE-FORMING) DM02407 P49109 I-335; K3-F337 フラビン含有モノオキシゲナーゼ PR00370; A5-K20, E28-K46, F141-P157, K186-L204, N23-D341, G377-G396, Y474-R495 ピリジンスクワレオチドジスルフィドレダクターゼ クラス II シグネチャ: PR00469; R187-K209, D325-Y333, F183-V207, V326-T347 フラビン結合モノオキシゲナーゼ様 FMO 様; K3-F535 ピリジンスクワレオチドジスルフィドレダクターゼ クラス II 活性部位 (pyridine_redox_2.prf): P157-I214 シグナル切断:M1-A18	BLAST-DOMO BLIMPS-PRINTS BLIMPS-PRINTS HMMER-PFAM PROFILESKAN SFSCAN
3	7476203CD1	501	S54, T118, T174, S372, T402, S412, T454		チトクロームP450 DM00022 P52786 83-492; S71-P478 チトクロームP450 モノオキシゲナーゼ オキシドレダクターゼ ヘム 電子 輸送 膜 ミクロソーム ENDOPLASMIC; PD000021; Q237-L351, D396-P464, P31-K60, E339-F411, N108-E158 P450 スーパーオキシドレダクターゼ: PR00385; A349-V360, I425-Y434, G296-I313, Q314-D327 ミトコンドリアP450 シグネチャ: PR00408; S114-R124, G296-I313, K342-V360	BLAST-DOMO BLAST-PRODOM BLIMPS-PRINTS BLIMPS-PRINTS

10

20

30

40

表3-2

SEQ ID NO.	Incyte ポリヌクレオチド ID	アミノ酸残基数	潜在的にリン酸化部位	潜在的グリコシル化部位	シグナチャ配列、ドメインおよびモチーフ	分析方法及びデータベース
4	7478583CD1	345	S143 S219 S251 S327 S99 T248 T53	N141	E-クラス P450 グループ I : PR00463; S173-D191, N285-A302, L305-G331, M348-V366, N389-D413, F424-Y434, Y434-F457 シグナルペプチド (signal_peptide): M1-C24 膜貫通ドメイン (transmem_domain): L4-F23, L197-Q214 チトクロム P450 (p450): F31-M67, R107-A485 チトクロム P450 システムインヘム鉄リガンドシグネチャ: F406-R456 シグナル切断: M1-Q49 亜結合有アルコールデヒドロゲナーゼ DM00064 P39714 18-373; E105-I343, E21-S99 オキシドレダクターゼ 亜結合有アルコールデヒドロゲナーゼ アルコール タンパク質 フラミリ 多重伝子 NADP ホルムアルデヒド: PD000104; A3-V265, G168-A246 亜結合有アルコールデヒドロゲナーゼタンパク質: BL00059; E26-H42, P65-V92, G117-Q158, G168-T214 亜結合有デヒドロゲナーゼ (adh_zinc): K2-K344 Adh_zinc G69-V83 亜結合有アルコールデヒドロゲナーゼシグネチャ (adh_zinc.prf): G38-E102 DM00064 S57611 3-340; Q32-V356 亜結合有デヒドロゲナーゼ (adh_zinc): D46-K357 (E-value = 4.4e-06, Score = -68.6) キノンオキシドレダクターゼ: BL01162; S150-K193 (p < 0.057) ヘパタン硫黄 6 スルホトランスフェラーゼ トランスフェラーゼ: PD139786; K207-R412, M1-K182 シグナルペプチド (signal_peptide): M1-Q24 膜貫通ドメイン (transmem_domain): I8-Y25	BLIMPS-PRINTS HMMER HMMER HMMER-PFAM PROFILESSCAN SPSCAN BLAST-DOMO BLAST-PRODOM BLIMPS-BLOCKS HMMER-PFAM MOTIFS PROFILESSCAN BLAST-DOMO HMMER-PFAM BLIMPS-BLOCKS BLAST-PRODOM HMMER HMMER
5	7478585CD1	361	T14, Y86, S107, S113, T330	N81		
6	7479904CD1	499	S5, S50, S55, T65, T120, T129, S156, S187, S199, T205, S208,	N63, N204, N296, N354, N438, N450, N458, N483, N486		

10

20

30

40

【表 8】

表3-3

SEQ ID NO:	Incyte ポリペプチドID	アミノ酸残基数	潜在的にリン酸化部位	潜在的グリコシル化部位	シグナルペプチド配列、ドメインおよびモチーフ	分析方法及びデータベース
7	7480367CD1	222	T237, T251, T260, Y277, S314, T355, S359, S464, S471, S476		シグナル切断: M1-Q24 グルタチオンSトランスフェラーゼ DM00127 P08263 I1-190; L72-P192 グルタチオンSトランスフェラーゼ Sトランスフェラーゼ 多重遺伝子 ファミリー タンパク質 クラスアルファ S クリスタリン GST LYASE: PD000312; L81-R204, K6-L102 (p = 8.0e-10) グルタチオンSトランスフェラーゼ: PF00043; Q54-G83 グルタチオンSトランスフェラーゼ: L7-P192 短縮アルコールヒドロゲナーゼアミノDM00034 P37058 43-275; I63-I297 短縮ヒドロゲナーゼ: BL00061; G198-E235 アルコールヒドロゲナーゼ: PRO0080; F218-A237, K143-V154, G198-G206 グルコース/ピロリドヒドロゲナーゼ: PR00081; W69-E86, K143-V154, M192-C208, F218-A237, K239-T256 短縮ヒドロゲナーゼ(adh_short): R68-S254 短縮アルコールヒドロゲナーゼアミノシグネチヤ(adh_short.prf): G198-P248 シグナル切断: M1-K36 カルボキシシルエステラーゼ タイプB: DM00175 P16303 20-361; M1-M264 エステラーゼ ヒドロラーゼ 前駆体 シグナル 糖タンパク質 セリン タンパク質 カルボキシシルエステラーゼ ファミリー 多重遺伝子 PD000169; S4-M281 カルボキシシルエステラーゼ タイプB: BL00122; E16-P26, P36-F46, V65-L80, G89-I129, L141-G150 カルボキシシルエステラーゼ (COesterase): L12-F266	SPSCAN BLAST-DOMO BLAST-PRODOM BLIMPS-PFAM HMMER-PFAM BLAST-DOMO BLIMPS-BLOCKS BLIMPS-PRINTS BLIMPS-PRINTS HMMER-PFAM PROFILESCAN SPSCAN BLAST-DOMO BLAST-PRODOM BLIMPS-BLOCKS HMMER-PFAM
8	8069390CD1	330	T33, S58, S98, T114, S128, S271, S286, T290, S320	N318		
9	7473869CD1	303	S4, Y20, S52, S81, S199, T277, S284	N116, N178		

10

20

30

40

【表 9】

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
10	7478588CD1	584	S23, T47, Y107, N237, N299, S125, Y141, S173, S202, S320, T398, S405, Y548, T554, S575	N112, N307, N363, N447, N459, N467, N492, N495	カルボキシシルエステラーゼモチーフ: (carboxylesterase_B_1): F110-G125 (carboxylesterase_B_2): E16-P26 カルボキシシルエステラーゼ タイプ B シグネチャ (carboxylesterase_b.prf): N90-A145 カルボキシシルエステラーゼ タイプ B DM00175 Q04791 26-353; K48-N385 エステラーゼ ヒドラーゼ 前駆体 シグナル タンパク質 セリン タンパク質 カルボキシシルエステラーゼ ファミリー 多重伝子: PD000169; M37-E430, W26-T203, M394-W569, A488-P578 カルボキシシルエステラーゼ タイプ B: BL00122; F74-P94, E137-P147, P157-F167, V186-L201, G210-I250, L262-G271, W522-N532 カルボキシシルエステラーゼ (COesterase): R25-W569 Carboxylesterase_B_1: F231-G246; Carboxylesterase_B_2: E137-P147 カルボキシシルエステラーゼ タイプ B シグネチャ (carboxylesterase_b.prf): N211-A266 シグナル切断: M1-G43	MOTIFS PROFILESCAN BLAST-DOMO BLAST-PRODOM BLIMPS-BLOCKS HMMER-PFAM MOTIFS PROFILESCAN SPSCAN BLAST-PRODOM HMMER
11	55046125CD1	508	S26, S54, S99, S104, T114, T169, T178, S205, T246, T260, T269, Y286, S323, T364, S368, S473, S480, S485	N112, N307, N363, N447, N459, N467, N492, N495	ヘパラン硫酸 6 スルホトランスフェラーゼ トランスフェラーゼ: PD139786;M50-R421 膜貫通ドメイン (transmem_domain): L57-Y74	HMMER
12	3538709CD1	439	T34, T83, T167, S168, T281, S423	N108, N332	アルファ/ベータ ヒドラーゼファミリー: PR00111; G202-S217, A262-I275, V362-E376 エポキシヒドラーゼシグネチャ: PR00412; H182-R200, G202-S217, L248-L261, I393-F415 アルファ/ベータヒドラーゼファミリー(abbhydrolase): Y203-L416	BLIMPS-PRINTS BLIMPS-PRINTS HMMER-PFAM

【表 10】

10

20

30

40

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
13	71563101CD1	514	T32, S169, S192, T246, S253, T363, S448, T480, T490		アミノトランスフェラーゼIII ピリドキサル-リン酸 結合 DM00188 P16932 2-426; D271-A504, K84-R233 アラニングリオキシレート アミノトランスフェラーゼ 2 前駆体 EC 2.6.1.44 AGT ベータアラニンピルビン酸 BETALALAT II トランスフェラーゼ ピリドキサル リン酸 ミトコンドリア トランジット ペプチド; PD032383; M1-E94 アミノトランスフェラーゼ 3: BL00600; F343-G355, H377-L395, D98-P121, P160-S185, I190-L205, I281-G294, F302-G330 膜貫通ドメイン: Y104-N124, aminotran_3: W95-T208 I281-M487 アミノトランスフェラーゼIII ピリドキサルリン酸結合部位 (aa_transfer_class_3.psf): E308-C374 ホスホアラニン-ホスホサルフェート シンターゼ (PAPS) 結合部位 結合 DM00981 JC4531 1-287; M50-E193	BLAST-DOMO BLAST-PRODOM
14	7472027CD1	226	S17, T23, T28, S106, S136, T195	N21, N66	トランスフェラーゼ スルホトランスフェラーゼ ステロイド 代謝 ヒドロキシステロイド アルコール フェノール エストロゲン タンパク質 ステロイド結合 PD001218; L73-Y165	BLIMPS-BLOCKS HMMER HMMER-PFAM PROFILESSCAN BLAST-DOMO
15	7480358CD1	121	T38, S96, T119	N69	銅/亜鉛 スーパーオキシドジムスターゼ DM00227 P00442 3-147; K3-V116 スーパーオキシド ジムスターゼ 銅 亜鉛 オキシドレダクターゼ 銅 亜鉛 前駆体 シグナル ペリプラズム 細胞外; PD000469; M2-V116 銅/亜鉛 スーパーオキシドジムスターゼタンパク質; BL00087; K3-Q22, R40-W83 銅 亜鉛 スーパーオキシドジムスターゼ; PR00068; R43-P65, V97-T119 銅/亜鉛 スーパーオキシドジムスターゼ(SODC); (sodcu); M2-P73, S96-V116	BLAST-DOMO BLAST-PRODOM BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM

10

20

30

40

【表 1 1】

表3-6

SEQ ID NO.	Incyte ポリヌクレオチド ID	アミノ酸残基数	潜在的に酸化部位	潜在的グリコシル化部位	シグナチン配列、ドメインおよびモチーフ	分析方法及びデータベース
16	1618256CD1	486	T9, S55, T74, T265, S296, T349, T356, T420, S431, S444, T480	N374	シグナルペプチド: M1-G29 シグナル切断: M1-G36 セプトロシドスルボトランスフェラーゼ; PD140372; M13-G215, F279-Y461 タンパク質 類似性 キナーゼ 結菌 R13G10.2 染色体 III オキシレダクターゼ フラビンタンパク質; PD041935; L358-D433, D250-G337	SIGPEPT SFSCAN BLAST-PRODOM
17	3387823CD1	649	T64, T184, T236, S253, S368, S382, S431, T436, S471, S534, S356, S516, S547, T550, T563, T624	N209	シグナルペプチド: M1-E22 フラビン含有アミノキシダーゼ; Amino oxidase; P219-A647 アドレノキシシン レダクターゼ ファミリ (ミトコンドリアモノオキシシグナーゼ系のフラビン含有成分): PR00419A; S211-I233, PR00419B; K234-R247, PR00419D; V212-R226 芳香環 ヒドロキシラーゼ (フラビンタンパク質モノオキシシグナーゼ); PR00420A; S211-I233, PR00420C; Q614-P629 膜貫通ドメイン: M1-V18	BLAST-PRODOM SIGPEPT HMMER-PFAM BLIMPS-PRINTS
18	55142051CD1	258	S35, T67, S73, S74, S159, S174		芳香環 ヒドロキシラーゼ (フラビンタンパク質モノオキシシグナーゼ); PR00420A; S211-I233, PR00420C; Q614-P629 膜貫通ドメイン: M1-V18 O-メチルトランスフェラーゼ トランスフェラーゼ メチルトランスフェラーゼ; PD173699; I65-D162 (P=9.5e-10) カテコール-O-メチルトランスフェラーゼ; DM02481 P22734 41-139; E46-A137	HMMER BLAST-PRODOM BLIMPS-PRINTS

【表 1 2】

10

20

30

40

表3-7

SEQ ID NO:	Incyte ポリハ ブド 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
19	7395274CD1	544	S85, S88, S135, S148, S181, T184, Y205, T244, S335, S336, S393, T395, T406, T479, S519, T525, S542	N331, N334, N538	チトクローム P450 DM00022 P10635 52-480: L110-F528 チトクローム P450 モノオキシゲナーゼ オキシドレダクターゼ ヘム 電子 輸送 膜 ミクロソーム ENDOPLASMICPD000921: K285-G433, I386-O471, H450-I539, I436-I539, A227-D338, G58-I231 チトクローム P450 システイン BL00086: F480-F511 ミトコンドリア P450 シグネチャ PR00408:L370-E383, A398-V416, F462-D470, I481-C490, C490-F501, W50-P65, H169-K179, A352-S369 E-クラス P450 グループ I PR00463: A111-L130, S135-I156, A227-N245, Y341-T358, L361-G387, E404-P422, N445-D469, F480-C490, C490-F513 E-クラス P450 グループ IV PR00465: V105-V128, H450-I468, L474-C490, C490-M508 チトクローム P450: L107-R543 チトクローム P450: F483-G492 チトクローム P450 システイン ヘム鉄 リガンドシグネチャ (cytochrome_p450.prf): P464-A512 Signal_cleavage: M1-G47	BLAST-DOMO BLAST-PRODOM BLIMPS-BLOCKS BLIMPS_PRINTS BLIMPS-PRINTS BLIMPS-PRINTS HMMER-PFAM MOTIFS PROFILESSCAN SPSCAN

10

20

30

40

表4-1

ポリクレオチド SEQ ID NO.	Incyte ポリクレオチド ID	配列長	選択された断片	配列断片	5'位置	3'位置
20	7248285CB1	2603	2496-2603, 526-572	7690384JI (PROSTME06)	652	1330
				6460535F7 (OSTEONC01)	1144	1999
				8035282H1 (SMCRUNE01)	276	892
				2796172F6 (NFOLNCT01)	2072	2603
				7690384HI (PROSTME06)	1650	2185
				FL7248285_g7939436_0000006_g1890108_1_7-8	1106	1428
				GNN:g7939436_0000006_002.e dit	1	456
21	7472835CB1	1745	1-57, 412-1745	GBI.g7107731_0001.edit_2.comp	1445	1745
				GBI.g7107731_0001.edit_1	1	1674
22	7476203CB1	1587	1-209, 280-1587	GBI.g8567322_0000003_000004_edit	1	1139
				GBI.g8567322_0000003_000004_edit_2	960	1587
23	7478583CB1	1038	1-1038	GNN.g8576127_000058_002	1	1038
				GNN.g8576127_CDS_1.edit	1	1467
24	7478585CB1	1584	1-1584	GNN.7478585CT1	596	1584
				g1655588_g3191972_g856919_3_edit	1	1500
26	7480367CB1	669		GBI.g8348076_edit_1	1	669
				2090065H1 (PANCR0T04)	508	669
27	8069390CB1	3551	1609-1783, 701-898, 1-21, 3101-3551	7360250H1 (BRAIFEE05)	607	1056
				7660957H1 (OVARNOE02)	2157	2645
				7660957J1 (OVARNOE02)	1	537
				6450195H1 (BRAINOC01)	330	989
				1759470R6 (PITUNOT03)	2966	3512
				1619329T6 (BRAITUT12)	2859	3484
				4249653R6 (BRADDIRO1)	2357	2781
				g1203094	3095	3551
				1576067T6 (LNODNCT03)	2736	3390
				7365248H1 (OVARDCO01)	920	1560
4249653F6 (BRADDIRO1)	1748	2196				
7100733H1 (BRATDRO2)	1597	2148				
3333887F6 (BRAIFET01)	1068	1673				

10

20

30

40

【表 1 4】

表4-2

ポリスクレオチド SEQ ID NO:	Incyte ポリスクレオチド ID	配列長	選択された断片	配列断片	5'位置	3'位置
28	7473869CB1	2178	1-231, 1549-1579, 1949-2178, 654-732	55023671J1	1360	2178
				55031441J1	1	1032
				55071187J1	1045	1662
29	7478588CB1	2081	1-97, 1256-1450	55071183H1	987	1606
				5501748J1	1496	1970
				7580667H1 (BRAIFEC01)	543	1108
				60251275D1	1333	1837
				55058627J1	157	714
				6205079H1 (PIVUNON01)	1562	2081
				4365390H1 (SKIRNOT01)	1129	1391
				55111596H1	1	540
				55071187J1	752	1365
				71975318V1	1694	2315
30	55046125CB1	2642	2250-2642, 403-468	2783360F6 (BRSTNOT13)	2121	2642
				55071237J1	1	548
				55046165J2	720	1572
				55046125H1	314	1143
				5626585H1 (PLACFER01)	1512	2108
				7067524H1 (BRATNOR01)	624	1178
				2225614F6 (SEMVNOT01)	1797	2080
				1668310H1 (BMARNOT03)	1141	1376
				3763474H1 (BRAINOT21)	1166	1461
				7764424H1 (UREFTTUE01)	1	463
31	3538709CB1	2080	863-1020, 2027-2080, 580-639	6883439J1 (BRAHTR03)	1315	1932
				7191721H2 (BRATD1C01)	253	855
				2560864H1 (ADRETTU01)	1656	1947
				92985189	1734	2212
				71563534V1	684	1366
				71562178V1	1184	1840
				6926859H1 (SINLIUMR01)	1	662
				293718H1 (LIVNNOF04)	1828	2066
				71563101V1	529	1354
				FL7472027CT1_00001	1	681
32	71563101CB1	2219	1-44, 541-1063	GNV_92211248_000026_002	1	399
				2241-2302, 1-	1	604
					1	
33	7472027CB1	681	148-681			
34	7480358CB1	399	1-35			
35	1618256CB1	2302	2241-2302, 1-			

10

20

30

40

【表 15】

表4-3

ポリヌクレオチド SEQ ID NO.	Incyte ポリヌクレオチド ID	配列長	選択された断片	配列断片	5'位置	3'位置
36	3387823CBI	3341	260, 1172- 1194	1618256F6 (BRAITUT12)	1329	1876
				6330412F8 (BRANDIN01)	359	1086
				7403247H1 (SINDIME01)	1475	2112
				4922615F9 (TESINOT11)	646	1374
				2605131F6 (LUNGPUT07)	1835	2302
				7741282H1 (THYMOE01)	1	662
				72335796V1	2819	3341
				70858392V1	1919	2549
				70857153V1	2459	2872
				70856101V1	1381	1902
37	55142051CBI	777	1-447, 3270- 3341, 876- 1415	657765H1 (EOSINOT03)	2322	2590
				70002313D1	1551	2004
				9959568H1 (KIDNUT01)	2584	2902
				6771554J1 (BRAUNOR01)	788	1218
				70857616V1	1116	1524
				7087862H1 (BRAUTDR03)	438	906
				55142051J1	72	337
				GBI.98567879_edit	1	777
				6779264H1 (OVARDIR01)	2278	2967
				92064300	3001	3596
38	7395274CBI	3600	1-1390, 2414- 3076, 2342- 2372, 3576- 3600	55052831H1	787	1597
				7394114R8 (BRABDIE02)	454	1371
				8128245H1 (SCOMDICO1)	1	693
				2300407R6 (BRSTNOT05)	2076	2545
				2300407T6 (BRSTNOT05)	3044	3568
				7996595H1 (ADRETUCO1)	1493	2221
				92079829	3243	3600
				7458912H1 (LIVRTUE01)	1677	2255
				1259839T6 (MENLUT03)	2869	3567
				6544865H1 (LNOBNO02)	3124	3590

10

20

30

40

表 5

ポリアクレオザド SEQ ID NO:	Incyte プロジェクト ID	代表的ライブラリ
20	7248285CB1	PROSTME06
26	7480367CB1	SININCT02
27	8069390CB1	HNT2AGT01
28	7473869CB1	BRAOTDR02
29	7476588CB1	OVARDIRO1
30	55046125CB1	BRSTNOT13
31	3538709CB1	BRAHDR03
32	71563101CB1	SINITMR01
35	1618256CB1	BRAITUT12
36	3387823CB1	BRAINOT19
38	7395274CB1	BRAEDIE02

10

20

30

40

表6-1

ライブラリ BRADDIE02	ベクター pINCY	ライブラリの説明 この5'UTRのランダムプライムされたライブラリは脳血管事故死亡の57才の白人男性の脳から除去した罹病小脳組織から単離したRNAを用いて作製した。血清検査は陰性だった。患者の病歴は、ハンチントン病、気腫およびタバコの使用(40年間1日3~4箱)などであった。
BRAHTER03	PCDNA2.1	このランダムプライムされたライブラリは胆管癌で死亡した55才女性から抽出した原始皮質、前海馬組織から単離したRNAから作製した。病理学検査では、主に凸面上に軽い顆粒線維症、帯状皮質と褶床の白質内に散在する輪索スフィロイド、内側嗅皮質および中脳水道周囲灰色領域にいくつかの散在する神経原線維変化が見られた。関連する腫瘍の病理学検査によると、残存または再発腫瘍を伴うよく分化した胆管癌が見られた。患者の病歴としては、胆管癌、術後 Budd-Chiari 症候群、胆道腫水症、胸水症、脱水症、栄養不良、減尿症、および急性腎不全などがあった。以前の手術としては、胆嚢摘出術および肝臓の85%の切除などがあつた。
BRAINOT19	pINCY	ライブラリは27才の白人男性の脳葉切除中に左前頭葉から採取した罹病脳組織から単離したRNAを用いて作製した。病理学検査では、顕著なグリーニッシュ、石灰化、およびヘモジデリン含有マクロファージを特徴とする限局性の深い白質病変が見られ、これは遠隔局所性興奮と合致する。この組織はまた、主に軟膜下および皮質下に軽度から中度の全脳性グリーニッシュを示している。これは慢性的な発作傷害と合致する。近心側頭葉を含む左側頭葉は、海馬セクサーCA1において限局性の顕著な錐体細胞の喪失とグリーニッシュを示していたが、これは近心側頭硬化と合致する。グリア酸性線維素タンパク質は星状細胞において陽性であった。患者は難治性癲癇、限局性癲癇、片麻痺、および不特定の脳傷害を示していた。患者の病歴は、脳性麻痺、歩行異常、および抑鬱性障害を含んでいた。家族の病歴には脳の癌が含まれていた。

10

20

30

40

表6-2

ライブラリ	ベクター	ライブラリの説明
BRAITUT12	P1NCY	ライブラリは、40才の白人女性の脳腫瘍の切除時に、左前頭葉から採取した脳腫瘍組織から単離した RNA を用いて作製した。病理は大円形細胞性星細胞腫のグレード4を示した。
BRAQTR02	PCDNA2.1	このランダムライブラリは、腫瘍で死亡した55才女性から抽出した混合中枢神経系(大脳皮質、水遁周囲灰質および黒質)から単離した RNA から作製した。病理学検査では、主に凸面上に軽い腫瘍性変化、帯状皮質と視床の白質内に散在する軸索スワロイド、内側嗅皮質および中脳水遁周囲灰質領域にいくつかの散在する神経線維変化が認められた。関連する腫瘍の病理学検査によくと、残存または再発腫瘍を伴うよく分化した胆管癌が見られた。患者の病歴としては、胆管癌、術後 Budd-Chiari 症候群、胆道腫瘍、胸水症、脱水症、栄養不良、減尿症、および急性腎不全などがあった。以前の手術としては、胆管癌切除および肝臓の85%の切除などがあった。
BRSENOT13	P1NCY	ライブラリは36才の白人女性の高側単純乳房切除時に採取した乳房組織から単離した RNA を用いて作製した。患者の病歴としては、乳房腫瘍、抑鬱性障害、高脂血症および慢性胃潰瘍などがあった。家族の病歴としては、心臓血管および脳血管病、高脂血症、皮膚癌、乳癌、食道癌、膀胱癌、骨癌およびホジキン病などがあった。
HNT2AGT01	PBLUESCRIPT	ライブラリは、hNT2細胞株(前駆神経先駆物質の特性を示した人骨形成から由来する)から単離した RNA を使用して Stratagene (STR937233)で作製した。細胞はシナイン様で5週間、有糸分裂阻害剤で2週間処理され、さらに条件付けられた培養液中で4週間培養された。
OVARDIR01	PCDNA2.1	この無作為ライブラリは45才の白人女性の腹式子宮全摘出術、両側性卵巣摘出術、腫瘍重と固定化、および付随的虫垂切除術時に除去された右卵巣組織から単離された RNA を用いて作製された。病理は左右卵巣の間質性卵巣萎縮増殖を示し、また一致する腫瘍組織の病理は左卵巣に類皮嚢腫(良性嚢胞性奇形腫)があった。複数(3個)の壁内平滑筋腫瘍を認め、頰部には扁平上皮異形成があった。患者の病歴には、不正子宮出血、女性ストレス性失調症、脱毛症、抑鬱病、肺炎、正常分娩および欠乏性貧血が含まれる。家族歴には良性高血圧、アテローム硬化型冠動脈疾患、高脂血症および皮膚初感染徴候がある。

【表 19】

10

20

30

40

表6-3

ライブラリ	バクター	ライブラリの説明
PROSTM06	PCDNA2.1	この5'パイアスでランダムプライムされたライブラリは、57才の白人男性から閉鎖型前立腺生検、前立腺全摘出術および局所リンパ節切除時に摘出した病変前立腺組織から単離した RNA を用いて作製した。病理学検査では腺癌増殖症が見られた。一致する腫瘍組織の病理検査では、腺癌 (Gleason グレード 3+3) が右側を中心に主要な塊を形成していた。この患者は前立腺特異抗原の上昇および前立腺癌を示した。患者の病歴は寛解期にあるタリド中毒であった。以前の手術としては、胆嚢摘出、精嚢腺入ルニアの修復、および脊椎骨折の修復が含まれていた。患者の使用薬剤としては、Pepsid、Omnipen および Eulexin などがあつた。家族の病歴としては、良性高血圧、脳血管事故、アテローム硬化性冠動脈疾患、子宮癌およびタイプ II 糖尿病が母親にあり、前立腺癌が父親にあり、兄弟姉妹に麻薬濫用、前立腺癌および乳癌があつた。
SINTM01	PCDNA2.1	このランダムプライムされたライブラリは70才白人女性の右半結腸切除、閉鎖型肝生検、S 状結腸鏡検査、結腸鏡検査および永久的結腸造設術時に摘出された回腸組織から単離された mRNA を用いて作製した。一致する腫瘍の病理は侵襲度 2 の腺癌で、回盲弁の 2cm 遠位部位に潰瘍性病変の形成を示した。患者の病歴には、悪性乳房腫瘍、II 型糖尿病、高脂血症、ウイルス性肝炎、非特定甲狀腺疾患、骨関節炎、悪性皮膚腫瘍、欠乏性貧血、および正常分娩が含まれている。家族の病歴には、乳癌、アテローム硬化性冠動脈疾患、良性の高血圧、脳血管疾患、卵巣癌、および高脂血症が見られた。
SINTNOT02	PBLUESCRIPT	ライブラリは、くも膜下出血で死した 55 才白人女性の小脳から単離された RNA を用いて作製された。血清学はサイトメガロウイルス (CMV) にて陽性であった。以前の手術としては、子宮摘出などがあつた。

10

20

30

40

表7-1

プログラム	説明	引用文献	パラメータ閾値
ABI FACTURA	核酸配列においてベクター配列を除去して、不定の塩基をマスクするプログラム	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finderは、アミノ酸または核酸配列の比較、注釈付に有用である。	Applied Biosystems, Foster City, CA. Paracel Inc., Pasadena, CA.	不一致<50%
ABI AutoAssembler	核酸配列を構築するプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	Basic Local Alignment Search Tool はアミノ酸および核酸配列の配列類似性検索に有用であり、BLASTには5つの機能がある: blastp, blastn, blastx, tblastnおよびtblastx。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) (1997) Nucleic Acids Res. 25:3389-3402.	ESTs確率値=1.0E-8 以下 完全原配列確率値= 1.0E-10以下
FASTA	問合せ配列と同種の配列群との類似性を検索する Pearson およびLipman アルゴリズム。FASTAは最低5つの機能からなる: fasta, tfasta, fastx, tfastxおよびssearch。	Pearson, W.R.及びD.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; およびSmith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTsfasta F 値=1.0E-6 構築された ESTs:fasta 同一性=95%以上 一致した長さ=200塩基以上 fastx E 値=1.0E-8 以下 完全原配列: 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、配列相同性および構造的フィンガープリント領域を検索するBLOCKS IMPROVED サーチャャー。	Henikoff, S.及びJ.G. Henikoff (1991) Nucleic Acids Res. 19:8665-8672; Henikoff, J.G.及びS. Henikoff (1996) Methods Enzymol. 266:88-105; Attwood, T.K. 他. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	確率値= 1.0E-3以下
HMMER	PFAMのようなタンパク質ファミリーコンセンサス配列の隠れマルコフモデル(HMM)に基づいたデータベースに対して問合せ配列を検索するアルゴリズム。	Krogh, A. 他. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. 他 (1988) Nucleic Acids Res. 26:320-322; 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

40

【表 2 1】

表7-2

プログラム	説明	引用文献	パラメータ閾値
ProfileScan	Prositeで定義された配列パターンと一致するタンパク質配列の構造的モチーフおよび配列モチーフを探索するアルゴリズム。	Gribskov, M. 他 (1988) CABIOS 4:61-66; Gribskov, M. 他 (1989) Methods Enzymol. 183:146-159; Bairoch, A. 他. (1997) Nucleic Acids Res. 25:217-221.	標準化された質のスコア>その特定のPrositeモチーフに対するGCC指定” HIGLI” 値 通常, スコア=1.4-2.1.
Phred	高い感度と確率で自動配列決定機のアレーを調べる塩基誤出しアルゴリズム。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. 及び P. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的なインプリメンテーションに基づくSWATやCrossMatchを含むPhils Revised Assembly プログラムで、配列相同性の検索やDNA配列の構築に有用である。	Smith, T.F. および M.S. Waterman (1981) Adv. Appl. Math. 2:482-488; Smith, T.F. および M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	スコア=120以上 一致した長さは56以上
Consed	Phrap で構築したものの表示、編集用グラフィカルツール。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキャンして、分泌シグナルペプチドの存在を調べる重み付けマトリックス解析プログラム。	Nielson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J.M. 及び S. Audic (1997) CABIOS 12:431-439.	スコア=3.5以上
TMAP	加重マトリックスを用いて蛋白質配列での膜貫通セグメントの描写および配向を決定するプログラム。	Persson, B. および P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. および P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	隠れMarkov モデル(HMM)を使って蛋白質配列の膜貫通セグメントを描写し、配向を決定するプログラム。	Sonnhammer, E.L. 他 (1998) Proc.Sixth Intl. Conf. on Intelligent Systems for Mol.Biol., Glasgow 他, eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA. 175-182.	
Motifs	Prositeで定義された配列と一致したタンパク質のアミノ酸配列を検索するプログラム。	Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, 第9版, M51-59頁, Genetics Computer Group, Madison, WI	

10

20

30

40

【国際公開パンフレット】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
14 February 2002 (14.02.2002)

PCT

(10) International Publication Number
WO 02/12467 A2

(51) International Patent Classification: C12N 9/00

(21) International Application Number: PCT/US01/24382

(22) International Filing Date: 3 August 2001 (03.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/223,055	4 August 2000 (04.08.2000)	US
60/224,728	11 August 2000 (11.08.2000)	US
60/226,440	18 August 2000 (18.08.2000)	US
60/228,067	24 August 2000 (24.08.2000)	US
60/230,063	31 August 2000 (31.08.2000)	US
60/232,244	13 September 2000 (13.09.2000)	US
60/234,269	20 September 2000 (20.09.2000)	US

CA 94025 (US). RAMKUMAR, Jayalaxmi [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). RING, Huijun, Z. [US/US]; 625 Orange Avenue, Los Altos, CA 94022 (US). SANJANWALA, Madhu, S. [US/US]; 210 Sylvia Court, Los Altos, CA 94024 (US). YANG, Y. Tom [US/US]; 4230 Ramwick Court, San Jose, CA 95118 (US). THANGAVELU, Kavitha [IN/US]; 1950 Montecito Avenue 23, Mountain View, CA 94043 (US). THORNTON, Michael [US/US]; 9 Medway Road, Woodside, CA 94062 (US). TRIBOLEY, Catherine, M. [FR/US]; 1121 Tennessee Street, #5, San Francisco, CA 94107 (US). WALLA, Narinder, K. [US/US]; 890 Davis Street #205, San Leandro, CA 94577 (US). WARREN, Bridget, A. [US/US]; 10130 Parwood Drive #2, Cupertino, CA 95014 (US). YANG, Junming [CN/US]; 1739 Bank Lane, San Jose, CA 95129 (US). YAO, Monique, G. [US/US]; 111 Frederick Court, Mountain View, CA 94043 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US).

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

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics, Inc., 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). BRUNS, Christopher, M. [US/US]; 575 S. Rengstorff Avenue #126, Mountain View, CA 94040 (US). DAS, Debopriya [IN/US]; 1179 Bonita Avenue, Apt. 3, Mountain View, CA 94040 (US). DELEGEANE, Angelo, M. [US/US]; 594 Angus Drive, Milpitas, CA 95035 (US). DING, Li [CN/US]; 3353 Alma Street, #146, Palo Alto, CA 94306 (US). ELLIOT, Vicki, S. [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). GANDHI, Ameena, R. [US/US]; 837 Roble Avenue, #1, Menlo Park, CA 94025 (US). GRIFFIN, Jennifer, A. [US/US]; 33691 Mello Way #11, Fremont, CA 94555 (US). HAPALIA, April, J., A. [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). KHAN, Farrah, A. [IN/US]; 3617 Central Road #102, Glenview, Illinois 60025 (US). LAL, Preeti [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). LEE, Sathy [US/US]; 825 East Evelyn, #425, Sunnyvale, CA 94086 (US). LI, Dying, Aina, M. [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). LI, Yan [CN/US]; 3885 Corina Way, Palo Alto, CA 94303 (US). PATTERSON, Chandra [US/US]; 490 Sherwood Way #1, Menlo Park,

(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, EE, ES, FI, GB, GD, GH, GI, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

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

Published:

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

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



WO 02/12467 A2

(54) Title: DRUG METABOLIZING ENZYMES

(57) Abstract: The invention provides human drug metabolizing enzymes (DME) and polymucleotides which identify and encode DME. 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 DME.

WO 02/12467

PCT/US01/24382

DRUG METABOLIZING ENZYMES**TECHNICAL FIELD**

This invention relates to nucleic acid and amino acid sequences of drug metabolizing enzymes and to the use of these sequences in the diagnosis, treatment, and prevention of autoimmune/inflammatory, cell proliferative, developmental, endocrine, eye, metabolic, and gastrointestinal disorders, including liver disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of drug metabolizing enzymes.

BACKGROUND OF THE INVENTION

The metabolism of a drug and its movement through the body (pharmacokinetics) are important in determining its effects, toxicity, and interactions with other drugs. The three processes governing pharmacokinetics are the absorption of the drug, distribution to various tissues, and elimination of drug metabolites. These processes are intimately coupled to drug metabolism, since a variety of metabolic modifications alter most of the physicochemical and pharmacological properties of drugs, including solubility, binding to receptors, and excretion rates. The metabolic pathways which modify drugs also accept a variety of naturally occurring substrates such as steroids, fatty acids, prostaglandins, leukotrienes, and vitamins. The enzymes in these pathways are therefore important sites of biochemical and pharmacological interaction between natural compounds, drugs, carcinogens, mutagens, and xenobiotics.

It has long been appreciated that inherited differences in drug metabolism lead to drastically different levels of drug efficacy and toxicity among individuals. For drugs with narrow therapeutic indices, or drugs which require bioactivation (such as codeine), these polymorphisms can be critical. Moreover, promising new drugs are frequently eliminated in clinical trials based on toxicities which may only affect a segment of the patient group. Advances in pharmacogenomics research, of which drug metabolizing enzymes constitute an important part, are promising to expand the tools and information that can be brought to bear on questions of drug efficacy and toxicity (See Evans, W. E. and R. V. Relling (1999) *Science* 286:487-491).

Drug metabolic reactions are categorized as Phase I, which functionalize the drug molecule and prepare it for further metabolism, and Phase II, which are conjugative. In general, Phase I reaction products are partially or fully inactive, and Phase II reaction products are the chief excreted species. However, Phase I reaction products are sometimes more active than the original administered drugs; this metabolic activation principle is exploited by pro-drugs (e.g. L-dopa).

WO 02/12467

PCT/US01/24382

Additionally, some nontoxic compounds (e.g. aflatoxin, benzo[*a*]pyrene) are metabolized to toxic intermediates through these pathways. Phase I reactions are usually rate-limiting in drug metabolism. Prior exposure to the compound, or other compounds, can induce the expression of Phase I enzymes however, and thereby increase substrate flux through the metabolic pathways. (See Klaassen, C. D., 5 Amdur, M. O. and J. Doull (1996) Casarett and Doull's Toxicology: The Basic Science of Poisons, McGraw-Hill, New York, NY, pp. 113-186; B. G. Katzung (1995) Basic and Clinical Pharmacology, Appleton and Lange, Norwalk, CT, pp. 48-59; G. G. Gibson and P. Skett (1994) Introduction to Drug Metabolism, Blackie Academic and Professional, London.)

Drug metabolizing enzymes (DMEs) have broad substrate specificities. This can be 10 contrasted to the immune system, where a large and diverse population of antibodies are highly specific for their antigens. The ability of DMEs to metabolize a wide variety of molecules creates the potential for drug interactions at the level of metabolism. For example, the induction of a DME by one compound may affect the metabolism of another compound by the enzyme.

DMEs have been classified according to the type of reaction they catalyze and the cofactors 15 involved. The major classes of Phase I enzymes include, but are not limited to, cytochrome P450 and flavin-containing monooxygenase. Other enzyme classes involved in Phase I-type catalytic cycles and reactions include, but are not limited to, NADPH cytochrome P450 reductase (CPR), the microsomal cytochrome b5/NADH cytochrome b5 reductase system, the ferredoxin/ferredoxin reductase redox pair, aldo/keto reductases, and alcohol dehydrogenases. The major classes of Phase II enzymes 20 include, but are not limited to, UDP glucuronyltransferase, sulfotransferase, glutathione S-transferase, N-acyltransferase, and N-acetyl transferase.

Cytochrome P450 and P450 catalytic cycle-associated enzymes

Members of the cytochrome P450 superfamily of enzymes catalyze the oxidative metabolism 25 of a variety of substrates, including natural compounds such as steroids, fatty acids, prostaglandins, leukotrienes, and vitamins, as well as drugs, carcinogens, mutagens, and xenobiotics. Cytochromes P450, also known as P450 heme-thiolate proteins, usually act as terminal oxidases in multi-component electron transfer chains, called P450-containing monooxygenase systems. Specific reactions catalyzed include hydroxylation, epoxidation, N-oxidation, sulfoxidation, N-, S-, and O-dealkylations, 30 desulfation, deamination, and reduction of azo, nitro, and N-oxide groups. These reactions are involved in steroidogenesis of glucocorticoids, cortisols, estrogens, and androgens in animals; insecticide resistance in insects; herbicide resistance and flower coloring in plants; and environmental bioremediation by microorganisms. Cytochrome P450 actions on drugs, carcinogens, mutagens, and xenobiotics can result in detoxification or in conversion of the substance to a more toxic product.

WO 02/12467

PCT/US01/24382

Cytochromes P450 are abundant in the liver, but also occur in other tissues; the enzymes are located in microsomes. (See ExPASy ENZYME EC 1.14.14.1; Prosite PDOC00081 Cytochrome P450 cysteine heme-iron ligand signature; PRINTS EP450I E-Class P450 Group I signature; Graham-Lorence, S. and Peterson, J.A. (1996) FASEB J. 10:206-214.)

5 Four hundred cytochromes P450 have been identified in diverse organisms including bacteria, fungi, plants, and animals (Graham-Lorence, *supra*). The B-class is found in prokaryotes and fungi, while the E-class is found in bacteria, plants, insects, vertebrates, and mammals. Five subclasses or groups are found within the larger family of E-class cytochromes P450 (PRINTS EP450I E-Class P450 Group I signature).

10 All cytochromes P450 use a heme cofactor and share structural attributes. Most cytochromes P450 are 400 to 530 amino acids in length. The secondary structure of the enzyme is about 70% alpha-helical and about 22% beta-sheet. The region around the heme-binding site in the C-terminal part of the protein is conserved among cytochromes P450. A ten amino acid signature sequence in this heme-iron ligand region has been identified which includes a conserved cysteine
15 involved in binding the heme iron in the fifth coordination site. In eukaryotic cytochromes P450, a membrane-spanning region is usually found in the first 15-20 amino acids of the protein, generally consisting of approximately 15 hydrophobic residues followed by a positively charged residue. (See Prosite PDOC00081, *supra*; Graham-Lorence, *supra*.)

Cytochrome P450 enzymes are involved in cell proliferation and development. The enzymes
20 have roles in chemical mutagenesis and carcinogenesis by metabolizing chemicals to reactive intermediates that form adducts with DNA (Nebert, D.W. and Gonzalez, F.J. (1987) Ann. Rev. Biochem. 56:945-993). These adducts can cause nucleotide changes and DNA rearrangements that lead to oncogenesis. Cytochrome P450 expression in liver and other tissues is induced by xenobiotics such as polycyclic aromatic hydrocarbons, peroxisomal proliferators, phenobarbital, and the
25 glucocorticoid dexamethasone (Dogra, S.C. et al. (1998) Clin. Exp. Pharmacol. Physiol. 25:1-9). A cytochrome P450 protein may participate in eye development as mutations in the P450 gene CYP1B1 cause primary congenital glaucoma (Online Mendelian Inheritance in Man (OMIM) *601771 Cytochrome P450, subfamily I (dioxin-inducible), polypeptide 1; CYP1B1).

Cytochromes P450 are associated with inflammation and infection. Hepatic cytochrome P450
30 activities are profoundly affected by various infections and inflammatory stimuli, some of which are suppressed and some induced (Morgan, E.T. (1997) Drug Metab. Rev. 29:1129-1188). Effects observed *in vivo* can be mimicked by proinflammatory cytokines and interferons. Autoantibodies to two cytochrome P450 proteins were found in patients with autoimmune polyendocrinopathy-

WO 02/12467

PCT/US01/24382

candidiasis-ectodermal dystrophy (APECED), a polyglandular autoimmune syndrome (OMIM *240300 Autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy).

Mutations in cytochromes P450 have been linked to metabolic disorders, including congenital adrenal hyperplasia, the most common adrenal disorder of infancy and childhood; pseudovitamin D-deficiency rickets; cerebrotendinous xanthomatosis, a lipid storage disease characterized by progressive neurologic dysfunction, premature atherosclerosis, and cataracts; and an inherited resistance to the anticoagulant drugs coumarin and warfarin (Isselbacher, K.J. et al. (1994) Harrison's Principles of Internal Medicine, McGraw-Hill, Inc. New York, NY, pp. 1968-1970; Takeyama, K. et al. (1997) *Science* 277:1827-1830; Kitataka, S. et al. (1998) *N. Engl. J. Med.* 338:653-661; OMIM *213700 Cerebrotendinous xanthomatosis; and OMIM #122700 Coumarin resistance). Extremely high levels of expression of the cytochrome P450 protein aromatase were found in a fibrolamellar hepatocellular carcinoma from a boy with severe gynecomastia (feminization) (Agarwal, V.R. (1998) *J. Clin. Endocrinol. Metab.* 83:1797-1800).

The cytochrome P450 catalytic cycle is completed through reduction of cytochrome P450 by NADPH cytochrome P450 reductase (CPR). Another microsomal electron transport system consisting of cytochrome b5 and NADPH cytochrome b5 reductase has been widely viewed as a minor contributor of electrons to the cytochrome P450 catalytic cycle. However, a recent report by Lamb, D. C. et al. (1999; *FEBS Lett.* 462:283-8) identifies a Candida albicans cytochrome P450 (CYP51) which can be efficiently reduced and supported by the microsomal cytochrome b5/NADPH cytochrome b5 reductase system. Therefore, there are likely many cytochromes P450 which are supported by this alternative electron donor system.

Cytochrome b5 reductase is also responsible for the reduction of oxidized hemoglobin (methemoglobin, or ferrihemoglobin, which is unable to carry oxygen) to the active hemoglobin (ferrohämoglobin) in red blood cells. Methemoglobinemia results when there is a high level of oxidant drugs or an abnormal hemoglobin (hemoglobin M) which is not efficiently reduced. Methemoglobinemia can also result from a hereditary deficiency in red cell cytochrome b5 reductase (Reviewed in Mansour, A. and Lurie, A. A. (1993) *Am. J. Hematol.* 42:7-12).

Members of the cytochrome P450 family are also closely associated with vitamin D synthesis and catabolism. Vitamin D exists as two biologically equivalent prohormones, ergocalciferol (vitamin D₂), produced in plant tissues, and cholecalciferol (vitamin D₃), produced in animal tissues. The latter form, cholecalciferol, is formed upon the exposure of 7-dehydrocholesterol to near ultraviolet light (i.e., 290-310 nm), normally resulting from even minimal periods of skin exposure to sunlight (reviewed in Miller, W.L. and Portale, A.A. (2000) *Trends Endocrinol. Metab.* 11:315-319).

WO 02/12467

PCT/US01/24382

Both prohormone forms are further metabolized in the liver to 25-hydroxyvitamin D (25(OH)D) by the enzyme 25-hydroxylase. 25(OH)D is the most abundant precursor form of vitamin D which must be further metabolized in the kidney to the active form, 1 α ,25-dihydroxyvitamin D (1 α ,25(OH)₂D), by the enzyme 25-hydroxyvitamin D 1 α -hydroxylase (1 α -hydroxylase). Regulation of 5 1 α ,25(OH)₂D production is primarily at this final step in the synthetic pathway. The activity of 1 α -hydroxylase depends upon several physiological factors including the circulating level of the enzyme product (1 α ,25(OH)₂D) and the levels of parathyroid hormone (PTH), calcitonin, insulin, calcium, phosphorus, growth hormone, and prolactin. Furthermore, extrarenal 1 α -hydroxylase activity has been reported, suggesting that tissue-specific, local regulation of 1 α ,25(OH)₂D production may also be biologically important. The catalysis of 1 α ,25(OH)₂D to 24,25-dihydroxyvitamin D 10 (24,25(OH)₂D), involving the enzyme 25-hydroxyvitamin D 24-hydroxylase (24-hydroxylase), also occurs in the kidney. 24-hydroxylase can also use 25(OH)D as a substrate (Shinki, T. *et al.* (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94:12920-12925; Miller, W.L. and Portale, A.A. *supra*; and references within).

15 Vitamin D 25-hydroxylase, 1 α -hydroxylase, and 24-hydroxylase are all NADPH-dependent, type I (mitochondrial) cytochrome P450 enzymes that show a high degree of homology with other members of the family. Vitamin D 25-hydroxylase also shows a broad substrate specificity and may also perform 26-hydroxylation of bile acid intermediates and 25, 26, and 27-hydroxylation of cholesterol (Dilworth, F.J. *et al.* (1995) *J. Biol. Chem.* 270:16766-16774; Miller, W.L. and Portale, 20 A.A. *supra*; and references within).

The active form of vitamin D (1 α ,25(OH)₂D) is involved in calcium and phosphate homeostasis and promotes the differentiation of myeloid and skin cells. Vitamin D deficiency resulting from deficiencies in the enzymes involved in vitamin D metabolism (e.g., 1 α -hydroxylase) causes hypocalcemia, hypophosphatemia, and vitamin D-dependent (sensitive) rickets, a disease characterized 25 by loss of bone density and distinctive clinical features, including bandy or bow leggedness accompanied by a waddling gait. Deficiencies in vitamin D 25-hydroxylase cause cerebrotendinous xanthomatosis, a lipid-storage disease characterized by the deposition of cholesterol and cholestanol in the Achilles' tendons, brain, lungs, and many other tissues. The disease presents with progressive neurologic dysfunction, including postpubescent cerebellar ataxia, atherosclerosis, and cataracts. 30 Vitamin D 25-hydroxylase deficiency does not result in rickets, suggesting the existence of alternative pathways for the synthesis of 25(OH)D (Griffin, J.E. and Zerwekh, J.E. (1983) *J. Clin. Invest.* 72:1190-1199; Gambliin, G.T. *et al.* (1985) *J. Clin. Invest.* 75:954-960; and W.L. and Portale, A.A. *supra*).

WO 02/12467

PCT/US01/24382

Ferredoxin and ferredoxin reductase are electron transport accessory proteins which support at least one human cytochrome P450 species, cytochrome P450c27 encoded by the CYP27 gene (Dilworth, F. J. et al. (1996) *Biochem. J.* 320:267-71). A *Streptomyces griseus* cytochrome P450, CYP104D1, was heterologously expressed in *E. coli* and found to be reduced by the endogenous

5 ferredoxin and ferredoxin reductase enzymes (Taylor, M. et al. (1999) *Biochem. Biophys. Res. Commun.* 263:838-42), suggesting that many cytochrome P450 species may be supported by the ferredoxin/ferredoxin reductase pair. Ferredoxin reductase has also been found in a model drug metabolism system to reduce actinomycin D, an antitumor antibiotic, to a reactive free radical species (Flitter, W. D. and Mason, R. P. (1988) *Arch. Biochem. Biophys.* 267:632-9).

10 Flavin-containing monooxygenase (FMO)

Flavin-containing monooxygenases oxidize the nucleophilic nitrogen, sulfur, and phosphorus heteroatom of an exceptional range of substrates. Like cytochromes P450, FMOs are microsomal and use NADPH and O₂; there is also a great deal of substrate overlap with cytochromes P450. The tissue distribution of FMOs includes liver, kidney, and lung.

15 There are five different known isoforms of FMO in mammals (FMO1, FMO2, FMO3, FMO4, and FMO5), which are expressed in a tissue-specific manner. The isoforms differ in their substrate specificities and other properties such as inhibition by various compounds and stereospecificity of reaction. FMOs have a 13 amino acid signature sequence, the components of which span the N-terminal two-thirds of the sequences and include the FAD binding region and the FATGY motif which

20 has been found in many N-hydroxylating enzymes (Stehr, M. et al. (1998) *Trends Biochem. Sci.* 23:56-57; PRINTS FMOXYGENASE Flavin-containing monooxygenase signature).

Specific reactions include oxidation of nucleophilic tertiary amines to N-oxides, secondary amines to hydroxylamines and nitrones, primary amines to hydroxylamines and oximes, and sulfur-containing compounds and phosphines to S- and P-oxides. Hydrazines, iodides, selenides, and boron-

25 containing compounds are also substrates. Although FMOs appear similar to cytochromes P450 in their chemistry, they can generally be distinguished from cytochromes P450 *in vitro* based on, for example, the higher heat lability of FMOs and the nonionic detergent sensitivity of cytochromes P450; however, use of these properties in identification is complicated by further variation among FMO isoforms with respect to thermal stability and detergent sensitivity.

30 FMOs play important roles in the metabolism of several drugs and xenobiotics. FMO (FMO3 in liver) is predominantly responsible for metabolizing (S)-nicotine to (S)-nicotine N-1'-oxide, which is excreted in urine. FMO is also involved in S-oxygenation of cimetidine, an H₂-antagonist widely used for the treatment of gastric ulcers. Liver-expressed forms of FMO are not under the same regulatory

WO 02/12467

PCT/US01/24382

control as cytochrome P450. In rats, for example, phenobarbital treatment leads to the induction of cytochrome P450, but the repression of FMO1.

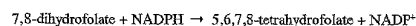
Endogenous substrates of FMO include cysteamine, which is oxidized to the disulfide, cystamine, and trimethylamine (TMA), which is metabolized to trimethylamine N-oxide. TMA smells like rotting fish, and mutations in the FMO3 isoform lead to large amounts of the malodorous free amine being excreted in sweat, urine, and breath. These symptoms have led to the designation fish-odor syndrome (OMIM 602079 Trimethylaminuria).

Lysyl oxidase:

Lysyl oxidase (lysine 6-oxidase, LO) is a copper-dependent amine oxidase involved in the formation of connective tissue matrices by crosslinking collagen and elastin. LO is secreted as a N-glycosylated precursor protein of approximately 50 kDa Levels and cleaved to the mature form of the enzyme by a metalloprotease, although the precursor form is also active. The copper atom in LO is involved in the transport of electron to and from oxygen to facilitate the oxidative deamination of lysine residues in these extracellular matrix proteins. While the coordination of copper is essential to LO activity, insufficient dietary intake of copper does not influence the expression of the apoenzyme. However, the absence of the functional LO is linked to the skeletal and vascular tissue disorders that are associated with dietary copper deficiency. LO is also inhibited by a variety of semicarbazides, hydrazines, and amino nitrites, as well as heparin. Beta-aminopropionitrile is a commonly used inhibitor. LO activity is increased in response to ozone, cadmium, and elevated levels of hormones released in response to local tissue trauma, such as transforming growth factor-beta, platelet-derived growth factor, angiotensin II, and fibroblast growth factor. Abnormalities in LO activity has been linked to Menkes syndrome and occipital horn syndrome. Cytosolic forms of the enzyme have been implicated in abnormal cell proliferation (reviewed in Rucker, R.B. et al. (1998) Am. J. Clin. Nutr. 67:996S-1002S and Smith-Mungo, L.I. and Kagan, H.M. (1998) Matrix Biol. 16:387-398).

Dihydrofolate reductases

Dihydrofolate reductases (DHFR) are ubiquitous enzymes that catalyze the NADPH-dependent reduction of dihydrofolate to tetrahydrofolate, an essential step in the *de novo* synthesis of glycine and purines as well as the conversion of deoxyuridine monophosphate (dUMP) to deoxythymidine monophosphate (dTMP). The basic reaction is as follows:



The enzymes can be inhibited by a number of dihydrofolate analogs, including trimethoprim and

WO 02/12467

PCT/US01/24382

methotrexate. Since an abundance of TMP is required for DNA synthesis, rapidly dividing cells require the activity of DHFR. The replication of DNA viruses (i.e., herpesvirus) also requires high levels of DHFR activity. As a result, drugs that target DHFR have been used for cancer chemotherapy and to inhibit DNA virus replication. (For similar reasons, thymidylate synthetases are also target enzymes.) Drugs that inhibit DHFR are preferentially cytotoxic for rapidly dividing cells (or DNA virus-infected cells) but have no specificity, resulting in the indiscriminate destruction of dividing cells. Furthermore, cancer cells may become resistant to drugs such as methotrexate as a result of acquired transport defects or the duplication of one or more DHFR genes (Stryer, L. (1988) Biochemistry: W.H Freeman and Co., Inc. New York. pp. 511-5619).

10 Aldo/keto reductases

Aldo/keto reductases are monomeric NADPH-dependent oxidoreductases with broad substrate specificities (Bohren, K. M. et al. (1989) *J. Biol. Chem.* 264:9547-51). These enzymes catalyze the reduction of carbonyl-containing compounds, including carbonyl-containing sugars and aromatic compounds, to the corresponding alcohols. Therefore, a variety of carbonyl-containing drugs and xenobiotics are likely metabolized by enzymes of this class.

One known reaction catalyzed by a family member, aldose reductase, is the reduction of glucose to sorbitol, which is then further metabolized to fructose by sorbitol dehydrogenase. Under normal conditions, the reduction of glucose to sorbitol is a minor pathway. In hyperglycemic states, however, the accumulation of sorbitol is implicated in the development of diabetic complications (OMIM *103880 Aldo-keto reductase family 1, member B1). Members of this enzyme family are also highly expressed in some liver cancers (Cao, D. et al. (1998) *J. Biol. Chem.* 273:11429-35).

20 Alcohol dehydrogenases

Alcohol dehydrogenases (ADHs) oxidize simple alcohols to the corresponding aldehydes. ADH is a cytosolic enzyme, prefers the cofactor NAD⁺, and also binds zinc ion. Liver contains the highest levels of ADH, with lower levels in kidney, lung, and the gastric mucosa.

Known ADH isoforms are dimeric proteins composed of 40 kDa subunits. There are five known gene loci which encode these subunits (a, b, g, p, c), and some of the loci have characterized allelic variants (b₁, b₂, b₃, g₁, g₂). The subunits can form homodimers and heterodimers; the subunit composition determines the specific properties of the active enzyme. The holoenzymes have therefore been categorized as Class I (subunit compositions aa, ab, ag, bg, gg), Class II (pp), and Class III (cc). Class I ADH isozymes oxidize ethanol and other small aliphatic alcohols, and are inhibited by pyrazole. Class II isozymes prefer longer chain aliphatic and aromatic alcohols, and are unable to oxidize methanol, and are not inhibited by pyrazole. Class III isozymes prefer even longer chain aliphatic alcohols (five

WO 02/12467

PCT/US01/24382

carbons and longer) and aromatic alcohols, and are not inhibited by pyrazole.

The short-chain alcohol dehydrogenases include a number of related enzymes with a variety of substrate specificities. Included in this group are the mammalian enzymes D-beta-hydroxybutyrate dehydrogenase, (R)-3-hydroxybutyrate dehydrogenase, 15-hydroxyprostaglandin dehydrogenase, NADPH-dependent carbonyl reductase, corticosteroid 11-beta-dehydrogenase, and estradiol 17-beta-dehydrogenase, as well as the bacterial enzymes acetoacetyl-CoA reductase, glucose 1-dehydrogenase, 3-beta-hydroxysteroid dehydrogenase, 20-beta-hydroxysteroid dehydrogenase, ribitol dehydrogenase, 3-oxoacyl reductase, 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase, sorbitol-6-phosphate 2-dehydrogenase, 7-alpha-hydroxysteroid dehydrogenase, cis-1,2-dihydroxy-3,4-cyclohexadiene-1-carboxylate dehydrogenase, cis-toluene dihydrodiol dehydrogenase, cis-benzene glycol dehydrogenase, biphenyl-2,3-dihydro-2,3-diol dehydrogenase, N-acylmannosamine 1-dehydrogenase, and 2-deoxy-D-gluconate 3-dehydrogenase (Krozowski, Z. (1994) J. Steroid Biochem. Mol. Biol. 51:125-130; Krozowski, Z. (1992) Mol. Cell Endocrinol. 84:C25-31; and Marks, A.R. et al. (1992) J. Biol. Chem. 267:15459-15463).

15 UDP glucuronyltransferase

Members of the UDP glucuronyltransferase family (UGTs) catalyze the transfer of a glucuronic acid group from the cofactor uridine diphosphate-glucuronic acid (UDP-glucuronic acid) to a substrate. The transfer is generally to a nucleophilic heteroatom (O, N, or S). Substrates include xenobiotics which have been functionalized by Phase I reactions, as well as endogenous compounds such as bilirubin, steroid hormones, and thyroid hormones. Products of glucuronidation are excreted in urine if the molecular weight of the substrate is less than about 250 g/mol, whereas larger glucuronidated substrates are excreted in bile.

UGTs are located in the microsomes of liver, kidney, intestine, skin, brain, spleen, and nasal mucosa, where they are on the same side of the endoplasmic reticulum membrane as cytochrome P450 enzymes and flavin-containing monooxygenases, and therefore are ideally located to access products of Phase I drug metabolism. UGTs have a C-terminal membrane-spanning domain which anchors them in the endoplasmic reticulum membrane, and a conserved signature domain of about 50 amino acid residues in their C terminal section (Prosite PDOC00359 UDP-glycosyltransferase signature).

30 UGTs involved in drug metabolism are encoded by two gene families, UGT1 and UGT2. Members of the UGT1 family result from alternative splicing of a single gene locus, which has a variable substrate binding domain and constant region involved in cofactor binding and membrane insertion. Members of the UGT2 family are encoded by separate gene loci, and are divided into two

WO 02/12467

PCT/US01/24382

families, UGT2A and UGT2B. The 2A subfamily is expressed in olfactory epithelium, and the 2B subfamily is expressed in liver microsomes. Mutations in UGT genes are associated with hyperbilirubinemia (OMIM #143500 Hyperbilirubinemia I); Crigler-Najjar syndrome, characterized by intense hyperbilirubinemia from birth (OMIM #218800 Crigler-Najjar syndrome); and a milder form of

5 hyperbilirubinemia termed Gilbert's disease (OMIM #191740 UGT1).

Sulfotransferase

Sulfate conjugation occurs on many of the same substrates which undergo O-glucuronidation to produce a highly water-soluble sulfuric acid ester. Sulfotransferases (ST) catalyze this reaction by transferring SO_3^- from the cofactor 3'-phosphoadenosine-5'-phosphosulfate (PAPS) to the substrate.

10 ST substrates are predominantly phenols and aliphatic alcohols, but also include aromatic amines and aliphatic amines, which are conjugated to produce the corresponding sulfamates. The products of these reactions are excreted mainly in urine.

STs are found in a wide range of tissues, including liver, kidney, intestinal tract, lung, platelets, and brain. The enzymes are generally cytosolic, and multiple forms are often co-expressed. For example, there are more than a dozen forms of ST in rat liver cytosol. These biochemically

15 characterized STs fall into five classes based on their substrate preference: arylsulfotransferase, alcohol sulfotransferase, estrogen sulfotransferase, tyrosine ester sulfotransferase, and bile salt sulfotransferase.

ST enzyme activity varies greatly with sex and age in rats. The combined effects of developmental cues and sex-related hormones are thought to lead to these differences in ST

20 expression profiles, as well as the profiles of other DMEs such as cytochromes P450. Notably, the high expression of STs in cats partially compensates for their low level of UDP glucuronyltransferase activity.

Several forms of ST have been purified from human liver cytosol and cloned. There are two

25 phenol sulfotransferases with different thermal stabilities and substrate preferences. The thermostable enzyme catalyzes the sulfation of phenols such as para-nitrophenol, minoxidil, and acetaminophen; the thermolabile enzyme prefers monoamine substrates such as dopamine, epinephrine, and levodopa. Other cloned STs include an estrogen sulfotransferase and an N-acetylglucosamine-6-O-

30 sulfotransferase. This last enzyme is illustrative of the other major role of STs in cellular biochemistry, the modification of carbohydrate structures that may be important in cellular differentiation and maturation of proteoglycans. Indeed, an inherited defect in a sulfotransferase has been implicated in macular corneal dystrophy, a disorder characterized by a failure to synthesize mature keratan sulfate proteoglycans (Nakazawa, K. et al. (1984) *J. Biol. Chem.* 259:13751-7; OMIM #217800 Macular

WO 02/12467

PCT/US01/24382

dystrophy, corneal).

Galactosyltransferases

Galactosyltransferases are a subset of glycosyltransferases that transfer galactose (Gal) to the terminal N-acetylglucosamine (GlcNAc) oligosaccharide chains that are part of glycoproteins or glycolipids that are free in solution (Kolbinger, F. et al. (1998) *J. Biol. Chem.* 273:433-440; Amado, M. et al. (1999) *Biochim. Biophys. Acta* 1473:35-53). Galactosyltransferases have been detected on the cell surface and as soluble extracellular proteins, in addition to being present in the Golgi. β 1,3-galactosyltransferases form Type I carbohydrate chains with Gal (β 1-3)GlcNAc linkages. Known human and mouse β 1,3-galactosyltransferases appear to have a short cytosolic domain, a single transmembrane domain, and a catalytic domain with eight conserved regions. (Kolbinger, F. *supra* and Hennes, T. et al. (1998) *J. Biol. Chem.* 273:58-65). In mouse UDP-galactose: β -N-acetylglucosamine β 1,3-galactosyltransferase-I region 1 is located at amino acid residues 78-83, region 2 is located at amino acid residues 93-102, region 3 is located at amino acid residues 116-119, region 4 is located at amino acid residues 147-158, region 5 is located at amino acid residues 172-183, region 6 is located at amino acid residues 203-206, region 7 is located at amino acid residues 236-246, and region 8 is located at amino acid residues 264-275. A variant of a sequence found within mouse UDP-galactose: β -N-acetylglucosamine β 1,3-galactosyltransferase-I region 8 is also found in bacterial galactosyltransferases, suggesting that this sequence defines a galactosyltransferase sequence motif (Hennes, T. *supra*). Recent work suggests that brainiac protein is a β 1,3-galactosyltransferase. (Yuan, Y. et al. (1997) *Cell* 88:9-11; and Hennes, T. *supra*).

UDP-Gal:GlcNAc-1,4-galactosyltransferase (-1,4-GalT) (Sato, T. et al., (1997) *EMBO J.* 16:1850-1857) catalyzes the formation of Type II carbohydrate chains with Gal (β 1-4)GlcNAc linkages. As is the case with the β 1,3-galactosyltransferase, a soluble form of the enzyme is formed by cleavage of the membrane-bound form. Amino acids conserved among β 1,4-galactosyltransferases include two cysteines linked through a disulfide-bonded and a putative UDP-galactose-binding site in the catalytic domain (Yadav, S. and Brew, K. (1990) *J. Biol. Chem.* 265:14163-14169; Yadav, S.P. and Brew, K. (1991) *J. Biol. Chem.* 266:698-703; and Shaper, N.L. et al. (1997) *J. Biol. Chem.* 272:31389-31399). β 1,4-galactosyltransferases have several specialized roles in addition to synthesizing carbohydrate chains on glycoproteins or glycolipids. In mammals a β 1,4-galactosyltransferase, as part of a heterodimer with α -lactalbumin, functions in lactating mammary gland lactose production. A β 1,4-galactosyltransferase on the surface of sperm functions as a receptor that specifically recognizes the egg. Cell surface β 1,4-galactosyltransferases also function in cell adhesion, cell/basal lamina interaction, and normal and metastatic cell migration. (Shur,

WO 02/12467

PCT/US01/24382

B. (1993) *Curr. Opin. Cell Biol.* 5:854-863; and Shaper, J. (1995) *Adv. Exp. Med. Biol.* 376:95-104).
Glutathione S-transferase

The basic reaction catalyzed by glutathione S-transferases (GST) is the conjugation of an electrophile with reduced glutathione (GSH). GSTs are homodimeric or heterodimeric proteins localized mainly in the cytosol, but some level of activity is present in microsomes as well. The major isozymes share common structural and catalytic properties; in humans they have been classified into four major classes, Alpha, Mu, Pi, and Theta. The two largest classes, Alpha and Mu, are identified by their respective protein isoelectric points; pI = 7.5-9.0 (Alpha), and pI = 6.6 (Mu). Each GST possesses a common binding site for GSH and a variable hydrophobic binding site. The hydrophobic binding site in each isozyme is specific for particular electrophilic substrates. Specific amino acid residues within GSTs have been identified as important for these binding sites and for catalytic activity. Residues Q67, T68, D101, E104, and R131 are important for the binding of GSH (Lee, H-C et al. (1995) *J. Biol. Chem.* 270: 99-109). Residues R13, R20, and R69 are important for the catalytic activity of GST (Stenberg G et al. (1991) *Biochem. J.* 274: 549-55).

In most cases, GSTs perform the beneficial function of deactivation and detoxification of potentially mutagenic and carcinogenic chemicals. However, in some cases their action is detrimental and results in activation of chemicals with consequent mutagenic and carcinogenic effects. Some forms of rat and human GSTs are reliable preneoplastic markers that aid in the detection of carcinogenesis. Expression of human GSTs in bacterial strains, such as Salmonella typhimurium used in the well-known Ames test for mutagenicity, has helped to establish the role of these enzymes in mutagenesis. Dihalomethanes, which produce liver tumors in mice, are believed to be activated by GST. This view is supported by the finding that dihalomethanes are more mutagenic in bacterial cells expressing human GST than in untransfected cells (Thier, R. et al. (1993) *Proc. Natl. Acad. Sci. USA* 90: 8567-80). The mutagenicity of ethylene dibromide and ethylene dichloride is increased in bacterial cells expressing the human Alpha GST, A1-1, while the mutagenicity of aflatoxin B1 is substantially reduced by enhancing the expression of GST (Simula, T.P. et al. (1993) *Carcinogenesis* 14: 1371-6). Thus, control of GST activity may be useful in the control of mutagenesis and carcinogenesis.

GST has been implicated in the acquired resistance of many cancers to drug treatment, the phenomenon known as multi-drug resistance (MDR). MDR occurs when a cancer patient is treated with a cytotoxic drug such as cyclophosphamide and subsequently becomes resistant to this drug and to a variety of other cytotoxic agents as well. Increased GST levels are associated with some of these drug resistant cancers, and it is believed that this increase occurs in response to the drug agent which is then deactivated by the GST catalyzed GSH conjugation reaction. The increased GST levels

WO 02/12467

PCT/US01/24382

then protect the cancer cells from other cytotoxic agents which bind to GST. Increased levels of A1-1 in tumors has been linked to drug resistance induced by cyclophosphamide treatment (Dirven H.A. et al. (1994) *Cancer Res.* 54: 6215-20). Thus control of GST activity in cancerous tissues may be useful in treating MDR in cancer patients.

5 Gamma-glutamyl transpeptidase

Gamma-glutamyl transpeptidases are ubiquitously expressed enzymes that initiate extracellular glutathione (GSH) breakdown by cleaving gamma-glutamyl amide bonds. The breakdown of GSH provides cells with a regional cysteine pool for biosynthetic pathways. Gamma-glutamyl transpeptidases also contribute to cellular antioxidant defenses and expression is induced by oxidative stress. The cell surface-localized glycoproteins are expressed at high levels in cancer cells. Studies have suggested that the high level of gamma-glutamyl transpeptidases activity present on the surface of cancer cells could be exploited to activate precursor drugs, resulting in high local concentrations of anti-cancer therapeutic agents (Hanigan, M.H. (1998) *Chem. Biol. Interact.* 111-112:333-42; Taniguchi, N. and Ikeda, Y. (1998) *Adv. Enzymol. Relat. Areas Mol. Biol.* 72:239-78; Chikhi, N. et al. (1999) *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* 122:367-80).

15 Acyltransferase

N-acyltransferase enzymes catalyze the transfer of an amino acid conjugate to an activated carboxylic group. Endogenous compounds and xenobiotics are activated by acyl-CoA synthetases in the cytosol, microsomes, and mitochondria. The acyl-CoA intermediates are then conjugated with an amino acid (typically glycine, glutamine, or taurine, but also ornithine, arginine, histidine, serine, aspartic acid, and several dipeptides) by N-acyltransferases in the cytosol or mitochondria to form a metabolite with an amide bond. This reaction is complementary to O-glucuronidation, but amino acid conjugation does not produce the reactive and toxic metabolites which often result from glucuronidation.

One well-characterized enzyme of this class is the bile acid-CoA:amino acid N-acyltransferase (BAT) responsible for generating the bile acid conjugates which serve as detergents in the gastrointestinal tract (Falany, C. N. et al. (1994) *J. Biol. Chem.* 269:19375-9; Johnson, M. R. et al. (1991) *J. Biol. Chem.* 266:10227-33). BAT is also useful as a predictive indicator for prognosis of hepatocellular carcinoma patients after partial hepatectomy (Furutani, M. et al. (1996) *Hepatology* 24:1441-5).

30 Acetyltransferases

Acetyltransferases have been extensively studied for their role in histone acetylation. Histone acetylation results in the relaxing of the chromatin structure in eukaryotic cells, allowing transcription factors to gain access to promoter elements of the DNA templates in the affected region of the

WO 02/12467

PCT/US01/24382

genome (or the genome in general). In contrast, histone deacetylation results in a reduction in transcription by closing the chromatin structure and limiting access of transcription factors. To this end, a common means of stimulating cell transcription is the use of chemical agents that inhibit the deacetylation of histones (e.g., sodium butyrate), resulting in a global (albeit artifactual) increase in gene expression. The modulation of gene expression by acetylation also results from the acetylation of other proteins, including but not limited to, p53, GATA-1, MyoD, ACTR, TFIIIE, TFIIIF and the high mobility group proteins (HMG). In the case of p53, acetylation results in increased DNA binding, leading to the stimulation of transcription of genes regulated by p53. The prototypic histone acetylase (HAT) is Gcn5 from *Saccharomyces cerevisiae*. Gcn5 is a member of a family of acetylases that includes Tetrahymena p55, human Gcn5, and human p300/CBP. Histone acetylation is reviewed in (Cheung, W.L. et al. (2000) *Curr. Opin. Cell Biol.* 12:326-333 and Berger, S.L. (1999) *Curr. Opin. Cell Biol.* 11:336-341). Some acetyltransferase enzymes possess the alpha/beta hydrolase fold (Center of Applied Molecular Engineering Inst. of Chemistry and Biochemistry - University of Salzburg, <http://predict.sanger.ac.uk/irbm-course97/Docs/ms/>) common to several other major classes of enzymes, including but not limited to, acetylcholinesterases and carboxylesterases (Structural Classification of Proteins, <http://scop.mrc-lmb.cam.ac.uk/scop/index.html>).

N-acetyltransferase

Aromatic amines and hydrazine-containing compounds are subject to N-acetylation by the N-acetyltransferase enzymes of liver and other tissues. Some xenobiotics can be O-acetylated to some extent by the same enzymes. N-acetyltransferases are cytosolic enzymes which utilize the cofactor acetyl-coenzyme A (acetyl-CoA) to transfer the acetyl group in a two step process. In the first step, the acetyl group is transferred from acetyl-CoA to an active site cysteine residue; in the second step, the acetyl group is transferred to the substrate amino group and the enzyme is regenerated.

In contrast to most other DME classes, there are a limited number of known N-acetyltransferases. In humans, there are two highly similar enzymes, NAT1 and NAT2; mice appear to have a third form of the enzyme, NAT3. The human forms of N-acetyltransferase have independent regulation (NAT1 is widely-expressed, whereas NAT2 is in liver and gut only) and overlapping substrate preferences. Both enzymes appear to accept most substrates to some extent, but NAT1 does prefer some substrates (para-aminobenzoic acid, para-aminosalicylic acid, sulfamethoxazole, and sulfanilamide), while NAT2 prefers others (isoniazid, hydralazine, procainamide, dapson, aminoglutethimide, and sulfamethazine).

Clinical observations of patients taking the antituberculosis drug isoniazid in the 1950s led to the description of fast and slow acetylators of the compound. These phenotypes were shown

WO 02/12467

PCT/US01/24382

subsequently to be due to mutations in the NAT2 gene which affected enzyme activity or stability. The slow isoniazid acetylator phenotype is very prevalent in Middle Eastern populations (approx. 70%), and is less prevalent in Caucasian (approx. 50%) and Asian (<25%) populations. More recently, functional polymorphism in NAT1 has been detected, with approximately 8% of the population tested showing a slow acetylator phenotype (Butcher, N. J. et al. (1998) *Pharmacogenetics* 8:67-72). Since NAT1 can activate some known aromatic amine carcinogens, polymorphism in the widely-expressed NAT1 enzyme may be important in determining cancer risk (OMIM *108345 N-acetyltransferase 1).

Aminotransferases

Aminotransferases comprise a family of pyridoxal 5'-phosphate (PLP) -dependent enzymes that catalyze transformations of amino acids. Aspartate aminotransferase (AspAT) is the most extensively studied PLP-containing enzyme. It catalyzes the reversible transamination of dicarboxylic L-amino acids, aspartate and glutamate, and the corresponding 2-oxo acids, oxalacetate and 2-oxoglutarate. Other members of the family included pyruvate aminotransferase, branched-chain amino acid aminotransferase, tyrosine aminotransferase, aromatic aminotransferase, alanine:glyoxylate aminotransferase (AGT), and kynurenine aminotransferase (Vacca, R.A. et al. (1997) *J. Biol. Chem.* 272:21932-21937).

Primary hyperoxaluria type-1 is an autosomal recessive disorder resulting in a deficiency in the liver-specific peroxisomal enzyme, alanine:glyoxylate aminotransferase-1. The phenotype of the disorder is a deficiency in glyoxylate metabolism. In the absence of AGT, glyoxylate is oxidized to oxalate rather than being transaminated to glycine. The result is the deposition of insoluble calcium oxalate in the kidneys and urinary tract, ultimately causing renal failure (Lumb, M.J. et al. (1999) *J. Biol. Chem.* 274:20587-20596).

Kynurenine aminotransferase catalyzes the irreversible transamination of the L-tryptophan metabolite L-kynurenine to form kynurenic acid. The enzyme may also catalyze the reversible transamination reaction between L-2-aminoadipate and 2-oxoglutarate to produce 2-oxoadipate and L-glutamate. Kynurenic acid is a putative modulator of glutamatergic neurotransmission, thus a deficiency in kynurenine aminotransferase may be associated with pleiotropic effects (Buchli, R. et al. (1995) *J. Biol. Chem.* 270:29330-29335).

Copper-zinc superoxide dismutases

Copper-zinc superoxide dismutases are compact homodimeric metalloenzymes involved in cellular defenses against oxidative damage. The enzymes contain one atom of zinc and one atom of copper per subunit and catalyze the dismutation of superoxide anions into O₂ and H₂O₂. The rate of

WO 02/12467

PCT/US01/24382

dismutation is diffusion-limited and consequently enhanced by the presence of favorable electrostatic interactions between the substrate and enzyme active site. Examples of this class of enzyme have been identified in the cytoplasm of all the eukaryotic cells as well as in the periplasm of several bacterial species. Copper-zinc superoxide dismutases are robust enzymes that are highly resistant to proteolytic digestion and denaturing by urea and SDS. In addition to the compact structure of the enzymes, the presence of the metal ions and intrasubunit disulfide bonds is believed to be responsible for enzyme stability. The enzymes undergo reversible denaturation at temperatures as high as 70 °C (Battistoni, A. et al. (1998) J. Biol. Chem. 273:5655-5661).

Overexpression of superoxide dismutase has been implicated in enhancing freezing tolerance of transgenic *Alfalfa* as well as providing resistance to environmental toxins such as the diphenyl ether herbicide, acifluorfen (McKersie, B.D. et al. (1993) Plant Physiol. 103:1155-1163). In addition, yeast cells become more resistant to freeze-thaw damage following exposure to hydrogen peroxide which causes the yeast cells to adapt to further peroxide stress by upregulating expression of superoxide dismutases. In this study, mutations to yeast superoxide dismutase genes had a more detrimental effect on freeze-thaw resistance than mutations which affected the regulation of glutathione metabolism, long suspected of being important in determining an organisms survival through the process of cryopreservation (Jong-In Park, J-I. et al. (1998) J. Biol. Chem. 273:22921-22928).

Expression of superoxide dismutase is also associated with *Mycobacterium tuberculosis*, the organism that causes tuberculosis. Superoxide dismutase is one of the ten major proteins secreted by *M. tuberculosis* and its expression is upregulated approximately 5-fold in response to oxidative stress. *M. tuberculosis* expresses almost two orders of magnitude more superoxide dismutase than the nonpathogenic mycobacterium *M. smegmatis*, and secretes a much higher proportion of the expressed enzyme. The result is the secretion of ~350-fold more enzyme by *M. tuberculosis* than *M. smegmatis*, providing substantial resistance to oxidative stress (Harth, G. and Horwitz, M.A. (1999) J. Biol. Chem. 274:4281-4292).

The reduced expression of copper-zinc superoxide dismutases, as well as other enzymes with anti-oxidant capabilities, has been implicated in the early stages of cancer. The expression of copper-zinc superoxide dismutases has been shown to be lower in prostatic intraepithelial neoplasia and prostate carcinomas, compared to normal prostate tissue (Bostwick, D.G. (2000) Cancer 89:123-134).

30 Phosphodiesterases

Phosphodiesterases make up a class of enzymes which catalyze the hydrolysis of one of the two ester bonds in a phosphodiester compound. Phosphodiesterases are therefore crucial to a variety of cellular processes. Phosphodiesterases include DNA and RNA endonucleases and exonucleases.

WO 02/12467

PCT/US01/24382

which are essential for cell growth and replication, and topoisomerases, which break and rejoin nucleic acid strands during topological rearrangement of DNA. A Tyr-DNA phosphodiesterase functions in DNA repair by hydrolyzing dead-end covalent intermediates formed between topoisomerase I and DNA (Pouliot, J.J. et al. (1999) *Science* 286:552-555; Yang, S.-W. (1996) *Proc. Natl. Acad. Sci. USA* 93:11534-11539).

5 Acid sphingomyelinase is a phosphodiesterase which hydrolyzes the membrane phospholipid sphingomyelin to produce ceramide and phosphorylcholine. Phosphorylcholine is used in the synthesis of phosphatidylcholine, which is involved in numerous intracellular signaling pathways, while ceramide is an essential precursor for the generation of gangliosides, membrane lipids found in high concentration in neural tissue. Defective acid sphingomyelinase leads to a build-up of sphingomyelin molecules in lysosomes, resulting in Niemann-Pick disease (Schuchman, E.H. and S.R. Miranda (1997) *Genet. Test.* 1:13-19).

10 Glycerophosphoryl diester phosphodiesterase (also known as glycerophosphodiester phosphodiesterase) is a phosphodiesterase which hydrolyzes deacetylated phospholipid glycerophosphodiester to produce sn-glycerol-3-phosphate and an alcohol. Glycerophosphocholine, glycerophosphoethanolamine, glycerophosphoglycerol, and glycerophosphoinositol are examples of substrates for glycerophosphoryl diester phosphodiesterases. A glycerophosphoryl diester phosphodiesterase from *E. coli* has broad specificity for glycerophosphodiester substrates (Larson, T.J. et al. (1983) *J. Biol. Chem.* 248:5428-5432).

20 Cyclic nucleotide phosphodiesterases (PDEs) are crucial enzymes in the regulation of the cyclic nucleotides cAMP and cGMP. cAMP and cGMP function as intracellular second messengers to transduce a variety of extracellular signals including hormones, light, and neurotransmitters. PDEs degrade cyclic nucleotides to their corresponding monophosphates, thereby regulating the intracellular concentrations of cyclic nucleotides and their effects on signal transduction. Due to their roles as regulators of signal transduction, PDEs have been extensively studied as chemotherapeutic targets (Perry, M.J. and G.A. Higgs (1998) *Curr. Opin. Chem. Biol.* 2:472-481; Torphy, J.T. (1998) *Ann. J. Resp. Crit. Care Med.* 157:351-370).

25 Families of mammalian PDEs have been classified based on their substrate specificity and affinity, sensitivity to cofactors, and sensitivity to inhibitory agents (Beavo, J.A. (1995) *Physiol. Rev.* 75:725-748; Conti, M. et al. (1995) *Endocrine Rev.* 16:370-389). Several of these families contain distinct genes, many of which are expressed in different tissues as splice variants. Within PDE families, there are multiple isozymes and multiple splice variants of these isozymes (Conti, M. and S.-L.C. Jin (1999) *Frog. Nucleic Acid Res. Mol. Biol.* 63:1-38). The existence of multiple PDE families,

WO 02/12467

PCT/US01/24382

isozymes, and splice variants is an indication of the variety and complexity of the regulatory pathways involving cyclic nucleotides (Houslay, M.D. and G. Milligan (1997) Trends Biochem. Sci. 22:217-224).

Type 1 PDEs (PDE1s) are Ca^{2+} /calmodulin-dependent and appear to be encoded by at least three different genes, each having at least two different splice variants (Kakkar, R. et al. (1999) Cell Mol. Life Sci. 55:1164-1186). PDE1s have been found in the lung, heart, and brain. Some PDE1 isozymes are regulated *in vitro* by phosphorylation/dephosphorylation. Phosphorylation of these PDE1 isozymes decreases the affinity of the enzyme for calmodulin, decreases PDE activity, and increases steady state levels of cAMP (Kakkar, *supra*). PDE1s may provide useful therapeutic targets for disorders of the central nervous system, and the cardiovascular and immune systems due to the involvement of PDE1s in both cyclic nucleotide and calcium signaling (Perry, M.J. and G.A. Higgs (1998) Curr. Opin. Chem. Biol. 2:472-481).

PDE2s are cGMP-stimulated PDEs that have been found in the cerebellum, neocortex, heart, kidney, lung, pulmonary artery, and skeletal muscle (Sadhu, K. et al. (1999) J. Histochem. Cytochem. 47:895-906). PDE2s are thought to mediate the effects of cAMP on catecholamine secretion, participate in the regulation of aldosterone (Beavo, *supra*), and play a role in olfactory signal transduction (Juifis, D.M. et al. (1997) Proc. Natl. Acad. Sci. USA 94:3388-3395).

PDE3s have high affinity for both cGMP and cAMP, and so these cyclic nucleotides act as competitive substrates for PDE3s. PDE3s play roles in stimulating myocardial contractility, inhibiting platelet aggregation, relaxing vascular and airway smooth muscle, inhibiting proliferation of T-lymphocytes and cultured vascular smooth muscle cells, and regulating catecholamine-induced release of free fatty acids from adipose tissue. The PDE3 family of phosphodiesterases are sensitive to specific inhibitors such as cilostamide, enoximone, and lixazinone. Isozymes of PDE3 can be regulated by cAMP-dependent protein kinase, or by insulin-dependent kinases (Degerman, E. et al. (1997) J. Biol. Chem. 272:6823-6826).

PDE4s are specific for cAMP; are localized to airway smooth muscle, the vascular endothelium, and all inflammatory cells; and can be activated by cAMP-dependent phosphorylation. Since elevation of cAMP levels can lead to suppression of inflammatory cell activation and to relaxation of bronchial smooth muscle, PDE4s have been studied extensively as possible targets for novel anti-inflammatory agents, with special emphasis placed on the discovery of asthma treatments. PDE4 inhibitors are currently undergoing clinical trials as treatments for asthma, chronic obstructive pulmonary disease, and atopic eczema. All four known isozymes of PDE4 are susceptible to the inhibitor rolipram, a compound which has been shown to improve behavioral memory in mice (Barad, M. et al. (1998) Proc. Natl. Acad. Sci. USA 95:15020-15025). PDE4 inhibitors have also been

WO 02/12467

PCT/US01/24382

studied as possible therapeutic agents against acute lung injury, endotoxemia, rheumatoid arthritis, multiple sclerosis, and various neurological and gastrointestinal indications (Doherty, A.M. (1999) *Curr. Opin. Chem. Biol.* 3:466-473).

PDE5 is highly selective for cGMP as a substrate (Turko, I.V. et al. (1998) *Biochemistry* 37:4200-4205), and has two allosteric cGMP-specific binding sites (McAllister-Lucas, L.M. et al. (1995) *J. Biol. Chem.* 270:30671-30679). Binding of cGMP to these allosteric binding sites seems to be important for phosphorylation of PDE5 by cGMP-dependent protein kinase rather than for direct regulation of catalytic activity. High levels of PDE5 are found in vascular smooth muscle, platelets, lung, and kidney. The inhibitor zaprinast is effective against PDE5 and PDE1s. Modification of zaprinast to provide specificity against PDE5 has resulted in sildenafil (VIAGRA; Pfizer, Inc., New York NY), a treatment for male erectile dysfunction (Terrett, N. et al. (1996) *Bioorg. Med. Chem. Lett.* 6:1819-1824). Inhibitors of PDE5 are currently being studied as agents for cardiovascular therapy (Perry, M.J. and G.A. Higgs (1998) *Curr. Opin. Chem. Biol.* 2:472-481).

PDE6s, the photoreceptor cyclic nucleotide phosphodiesterases, are crucial components of the phototransduction cascade. In association with the G-protein transducin, PDE6s hydrolyze cGMP to regulate cGMP-gated cation channels in photoreceptor membranes. In addition to the cGMP-binding active site, PDE6s also have two high-affinity cGMP-binding sites which are thought to play a regulatory role in PDE6 function (Artemyev, N.O. et al. (1998) *Methods* 14:93-104). Defects in PDE6s have been associated with retinal disease. Retinal degeneration in the rd mouse (Yan, W. et al. (1998) *Invest. Ophthalmol. Vis. Sci.* 39:2529-2536), autosomal recessive retinitis pigmentosa in humans (Danciger, M. et al. (1995) *Genomics* 30:1-7), and rod/cone dysplasia 1 in Irish Setter dogs (Suber, M.L. et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:3968-3972) have been attributed to mutations in the PDE6B gene.

The PDE7 family of PDEs consists of only one known member having multiple splice variants (Bloom, T.J. and J.A. Beavo (1996) *Proc. Natl. Acad. Sci. USA* 93:14188-14192). PDE7s are cAMP specific, but little else is known about their physiological function. Although mRNAs encoding PDE7s are found in skeletal muscle, heart, brain, lung, kidney, and pancreas, expression of PDE7 proteins is restricted to specific tissue types (Han, P. et al. (1997) *J. Biol. Chem.* 272:16152-16157; Perry, M.J. and G.A. Higgs (1998) *Curr. Opin. Chem. Biol.* 2:472-481). PDE7s are very closely related to the PDE4 family; however, PDE7s are not inhibited by rolipram, a specific inhibitor of PDE4s (Beavo, supra).

PDE8s are cAMP specific, and are closely related to the PDE4 family. PDE8s are expressed in thyroid gland, testis, eye, liver, skeletal muscle, heart, kidney, ovary, and brain. The cAMP-

WO 02/12467

PCT/US01/24382

- hydrolyzing activity of PDE8s is not inhibited by the PDE inhibitors rolipram, vinpocetine, milrinone, IBMX (3-isobutyl-1-methylxanthine), or zaprinast, but PDE8s are inhibited by dipyrindamole (Fisher, D.A. et al. (1998) *Biochem. Biophys. Res. Commun.* 246:570-577; Hayashi, M. et al. (1998) *Biochem. Biophys. Res. Commun.* 250:751-756; Soderling, S.H. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:8991-8996).
- 5 PDE9s are cGMP specific and most closely resemble the PDE8 family of PDEs. PDE9s are expressed in kidney, liver, lung, brain, spleen, and small intestine. PDE9s are not inhibited by sildenafil (VIAGRA; Pfizer, Inc., New York NY), rolipram, vinpocetine, dipyrindamole, or IBMX (3-isobutyl-1-methylxanthine), but they are sensitive to the PDE5 inhibitor zaprinast (Fisher, D.A. et al. (1998) *J. Biol. Chem.* 273:15559-15564; Soderling, S.H. et al. (1998) *J. Biol. Chem.* 273:15553-15558).
- 10 PDE10s are dual-substrate PDEs, hydrolyzing both cAMP and cGMP. PDE10s are expressed in brain, thyroid, and testis. (Soderling, S.H. et al. (1999) *Proc. Natl. Acad. Sci. USA* 96:7071-7076; Fujishige, K. et al. (1999) *J. Biol. Chem.* 274:18438-18445; Loughney, K. et al. (1999) *Gene* 234:109-117).
- 15 PDEs are composed of a catalytic domain of about 270-300 amino acids, an N-terminal regulatory domain responsible for binding cofactors, and, in some cases, a hydrophilic C-terminal domain of unknown function (Conti, M. and S.-L.C. Jin (1999) *Prog. Nucleic Acid Res. Mol. Biol.* 63:1-38). A conserved, putative zinc-binding motif, HDXXHXGXXN, has been identified in the catalytic domain of all PDEs. N-terminal regulatory domains include non-catalytic cGMP-binding domains in PDE2s, PDE5s, and PDE6s; calmodulin-binding domains in PDE1s; and domains containing phosphorylation sites in PDE3s and PDE4s. In PDE5, the N-terminal cGMP-binding domain spans about 380 amino acid residues and comprises tandem repeats of the conserved sequence motif N(R/K)XnFX₂DE (McAllister-Lucas, L.M. et al. (1993) *J. Biol. Chem.* 268:22863-22873). The NKXnD motif has been shown by mutagenesis to be important for cGMP binding
- 20 (Turko, I.V. et al. (1996) *J. Biol. Chem.* 271:22240-22244). PDE families display approximately 30% amino acid identity within the catalytic domain; however, isozymes within the same family typically display about 85-95% identity in this region (e.g. PDE4A vs PDE4B). Furthermore, within a family there is extensive similarity (>60%) outside the catalytic domain; while across families, there is little or no sequence similarity outside this domain.
- 25 Many of the constituent functions of immune and inflammatory responses are inhibited by agents that increase intracellular levels of cAMP (Verghese, M.W. et al. (1995) *Mol. Pharmacol.* 47:1164-1171). A variety of diseases have been attributed to increased PDE activity and associated with decreased levels of cyclic nucleotides. For example, a form of diabetes insipidus in mice has
- 30

WO 02/12467

PCT/US01/24382

been associated with increased PDE4 activity, an increase in low- K_m cAMP PDE activity has been reported in leukocytes of atopic patients, and PDE3 has been associated with cardiac disease.

Many inhibitors of PDEs have been identified and have undergone clinical evaluation (Perry, M.J. and G.A. Higgs (1998) *Curr. Opin. Chem. Biol.* 2:472-481; Torphy, T.J. (1998) *Am. J. Respir. Crit. Care Med.* 157:351-370). PDE3 inhibitors are being developed as antithrombotic agents, antihypertensive agents, and as cardiotoxic agents useful in the treatment of congestive heart failure. Rolipram, a PDE4 inhibitor, has been used in the treatment of depression, and other inhibitors of PDE4 are undergoing evaluation as anti-inflammatory agents. Rolipram has also been shown to inhibit lipopolysaccharide (LPS) induced TNF- α which has been shown to enhance HIV-1 replication *in vitro*. Therefore, rolipram may inhibit HIV-1 replication (Angel, J.B. et al. (1995) *AIDS* 9:1137-1144). Additionally, rolipram, based on its ability to suppress the production of cytokines such as TNF- α and β and interferon γ , has been shown to be effective in the treatment of encephalomyelitis. Rolipram may also be effective in treating tardive dyskinesia and was effective in treating multiple sclerosis in an experimental animal model (Sommer, N. et al. (1995) *Nat. Med.* 1:244-248; Sasaki, H. et al. (1995) *Eur. J. Pharmacol.* 282:71-76).

Theophylline is a nonspecific PDE inhibitor used in the treatment of bronchial asthma and other respiratory diseases. Theophylline is believed to act on airway smooth muscle function and in an anti-inflammatory or immunomodulatory capacity in the treatment of respiratory diseases (Banner, K.H. and C.P. Page (1995) *Eur. Respir. J.* 8:996-1000). Pentoxifylline is another nonspecific PDE inhibitor used in the treatment of intermittent claudication and diabetes-induced peripheral vascular disease. Pentoxifylline is also known to block TNF- α production and may inhibit HIV-1 replication (Angel et al., *supra*).

PDEs have been reported to affect cellular proliferation of a variety of cell types (Conti et al. (1995) *Endocrine Rev.* 16:370-389) and have been implicated in various cancers. Growth of prostate carcinoma cell lines DU145 and LNCaP was inhibited by delivery of cAMP derivatives and PDE inhibitors (Bang, Y.J. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:5330-5334). These cells also showed a permanent conversion in phenotype from epithelial to neuronal morphology. It has also been suggested that PDE inhibitors have the potential to regulate mesangial cell proliferation (Matousovich, K. et al. (1995) *J. Clin. Invest.* 96:401-410) and lymphocyte proliferation (Joulain, C. et al. (1995) *J. Lipid Mediat. Cell Signal.* 11:63-79). A cancer treatment has been described that involves intracellular delivery of PDEs to particular cellular compartments of tumors, resulting in cell death (Deonarain, M.P. and A.A. Epenetos (1994) *Br. J. Cancer* 70:786-794).

Phosphodiesterases

WO 02/12467

PCT/US01/24382

Phosphotriesterases (PTE, paraoxonases) are enzymes that hydrolyze toxic organophosphorus compounds and have been isolated from a variety of tissues. The enzymes appear to be lacking in birds and insects and abundant in mammals, explain the reduced tolerance of birds and insects to organophosphorus compound (Vilanova, E. and Sogorb, M.A. (1999) Crit. Rev. Toxicol. 29:21-57).

5 Phosphotriesterases play a central role in the detoxification of insecticides by mammals. Phosphotriesterase activity varies among individuals and is lower in infants than adults. Knockout mice are markedly more sensitive to the organophosphate-based toxins diazoxon and chlorpyrifos oxon (Furlong, C.E., et al. (2000) Neurotoxicology 21:91-100). PTEs have attracted interest as enzymes capable of the detoxification of organophosphate-containing chemical waste and warfare reagents (e.g., parathion), in addition to pesticides and insecticides. Some studies have also implicated

10 phosphotriesterase in atherosclerosis and diseases involving lipoprotein metabolism.

Thioesterases

Two soluble thioesterases involved in fatty acid biosynthesis have been isolated from mammalian tissues, one which is active only toward long-chain fatty-acyl thioesters and one which is

15 active toward thioesters with a wide range of fatty-acyl chain-lengths. These thioesterases catalyze the chain-terminating step in the *de novo* biosynthesis of fatty acids. Chain termination involves the hydrolysis of the thioester bond which links the fatty acyl chain to the 4'-phosphopantetheine prosthetic group of the acyl carrier protein (ACP) subunit of the fatty acid synthase (Smith, S. (1981a) Methods Enzymol. 71:181-188; Smith, S. (1981b) Methods Enzymol. 71:188-200).

20 *E. coli* contains two soluble thioesterases, thioesterase I which is active only toward long-chain acyl thioesters, and thioesterase II (TEII) which has a broad chain-length specificity (Naggert, J. et al. (1991) J. Biol. Chem. 266:11044-11050). *E. coli* TEII does not exhibit sequence similarity with either of the two types of mammalian thioesterases which function as chain-terminating enzymes in *de novo* fatty acid biosynthesis. Unlike the mammalian thioesterases, *E. coli* TEII lacks the characteristic

25 serine active site gly-X-ser-X-gly sequence motif and is not inactivated by the serine modifying agent diisopropyl fluorophosphate. However, modification of histidine 58 by iodoacetamide and diethylpyrocarbonate abolished TEII activity. Overexpression of TEII did not alter fatty acid content in *E. coli*, which suggests that it does not function as a chain-terminating enzyme in fatty acid biosynthesis (Naggert et al., *supra*). For that reason, Naggert et al. (*supra*) proposed that the

30 physiological substrates for *E. coli* TEII may be coenzyme A (CoA)-fatty acid esters instead of ACP-phosphopantetheine-fatty acid esters.

Carboxylesterases

Mammalian carboxylesterases constitute a multigene family expressed in a variety of tissues

WO 02/12467

PCT/US01/24382

and cell types. Isozymes have significant sequence homology and are classified primarily on the basis of amino acid sequence. Acetylcholinesterase, butyrylcholinesterase, and carboxylesterase are grouped into the serine super family of esterases (B-esterases). Other carboxylesterases included thyroglobulin, thrombin, Factor IX, gliotactin, and plasminogen. Carboxylesterases catalyze the hydrolysis of ester- and amide- groups from molecules and are involved in detoxification of drugs, environmental toxins, and carcinogens. Substrates for carboxylesterases include short- and long-chain acyl-glycerols, acylcarnitine, carbonates, dipiveftin hydrochloride, cocaine, salicylates, capsaicin, palmitoyl-coenzyme A, imidapril, haloperidol, pyrrolizidine alkaloids, steroids, p-nitrophenyl acetate, malathion, butanilcaine, and isocarboxazide. The enzymes often demonstrate low substrate specificity. Carboxylesterases are also important for the conversion of prodrugs to their respective free acids, which may be the active form of the drug (e.g., lovastatin, used to lower blood cholesterol) (reviewed in Satoh, T. and Hosokawa, M. (1998) *Annu. Rev. Pharmacol. Toxicol.* 38:257-288).

Neuroigins are a class of molecules that (i) have N-terminal signal sequences, (ii) resemble cell-surface receptors, (iii) contain carboxylesterase domains, (iv) are highly expressed in the brain, and (v) bind to neurexins in a calcium-dependent manner. Despite the homology to carboxylesterases, neuroigins lack the active site serine residue, implying a role in substrate binding rather than catalysis (Ichtchenko, K. et al. (1996) *J. Biol. Chem.* 271:2676-2682).

Squalene epoxidase

Squalene epoxidase (squalene monooxygenase, SE) is a microsomal membrane-bound, FAD-dependent oxidoreductase that catalyzes the first oxygenation step in the sterol biosynthetic pathway of eukaryotic cells. Cholesterol is an essential structural component of cytoplasmic membranes acquired via the LDL receptor-mediated pathway or the biosynthetic pathway. In the latter case, all 27 carbon atoms in the cholesterol molecule are derived from acetyl-CoA (Stryer, L., *supra*). SE converts squalene to 2,3(S)-oxidosqualene, which is then converted to lanosterol and then cholesterol. The steps involved in cholesterol biosynthesis are summarized below (Stryer, L. (1988) Biochemistry. W.H Freeman and Co., Inc. New York. pp. 554-560 and Sakakibara, J. et al. (1995) 270:17-20): acetate (from Acetyl-CoA) - 3-hydroxy-3-methyl-glutaryl CoA - mevalonate - 5-phosphomevalonate - 5-pyrophosphomevalonate - isopentenyl pyrophosphate - dimethylallyl pyrophosphate - geranyl pyrophosphate - farnesyl pyrophosphate - squalene - squalene epoxide - lanosterol - cholesterol

While cholesterol is essential for the viability of eukaryotic cells, inordinately high serum cholesterol levels results in the formation of atherosclerotic plaques in the arteries of higher organisms. This deposition of highly insoluble lipid material onto the walls of essential blood vessels (e.g., coronary arteries) results in decreased blood flow and potential necrosis of the tissues deprived of adequate

WO 02/12467

PCT/US01/24382

blood flow. HMG-CoA reductase is responsible for the conversion of 3-hydroxy-3-methyl-glutaryl CoA (HMG-CoA) to mevalonate, which represents the first committed step in cholesterol biosynthesis. HMG-CoA is the target of a number of pharmaceutical compounds designed to lower plasma cholesterol levels. However, inhibition of HMG-CoA also results in the reduced synthesis of non-sterol intermediates (e.g., mevalonate) required for other biochemical pathways. SE catalyzes a rate-limiting reaction that occurs later in the sterol synthesis pathway and cholesterol is the only end product of the pathway following the step catalyzed by SE. As a result, SE is the ideal target for the design of anti-hyperlipidemic drugs that do not cause a reduction in other necessary intermediates (Nakamura, Y. et al. (1996) 271:8053-8056).

10 Epoxide hydrolases

Epoxide hydrolases catalyze the addition of water to epoxide-containing compounds, thereby hydrolyzing epoxides to their corresponding 1,2-diols. They are related to bacterial haloalkane dehalogenases and show sequence similarity to other members of the α/β hydrolase fold family of enzymes (e.g., bromoperoxidase A2 from Streptomyces aureofaciens, hydroxymuconic semialdehyde hydrolases from Pseudomonas putida, and haloalkane dehalogenase from Xanthobacter autotrophicus). Epoxide hydrolases are ubiquitous in nature and have been found in mammals, invertebrates, plants, fungi, and bacteria. This family of enzymes is important for the detoxification of xenobiotic epoxide compounds which are often highly electrophilic and destructive when introduced into an organism. Examples of epoxide hydrolase reactions include the hydrolysis of cis-9,10-epoxyoctadec-9(Z)-enoic acid (leukotoxin) to form its corresponding diol, three-9,10-dihydroxyoctadec-12(Z)-enoic acid (leukotoxin diol), and the hydrolysis of cis-12,13-epoxyoctadec-9(Z)-enoic acid (isoleukotoxin) to form its corresponding diol three-12,13-dihydroxyoctadec-9(Z)-enoic acid (isoleukotoxin diol). Leukotoxins alter membrane permeability and ion transport and cause inflammatory responses. In addition, epoxide carcinogens are known to be produced by cytochrome P450 as intermediates in the detoxification of drugs and environmental toxins.

The enzymes possess a catalytic triad composed of Asp (the nucleophile), Asp (the histidine-supporting acid), and His (the water-activating histidine). The reaction mechanism of epoxide hydrolase proceeds via a covalently bound ester intermediate initiated by the nucleophilic attack of one of the Asp residues on the primary carbon atom of the epoxide ring of the target molecule, leading to a covalently bound ester intermediate (Michael Arand, M. et al. (1996) J. Biol. Chem. 271:4223-4229; Rink, R. et al. (1997) J. Biol. Chem. 272:14650-14657; Argiriadi, M.A. et al. (2000) J. Biol. Chem. 275:15265-15270).

WO 02/12467

PCT/US01/24382

Catechol-O-methyltransferase:

Catechol-O-methyltransferase (COMT) catalyzes the transfer of the methyl group of S-adenosyl-L-methionine (AdoMet; SAM) donor to one of the hydroxyl groups of the catechol substrate (e.g., L-dopa, dopamine, or DBA). Methylation of the 3'-hydroxyl group is favored over methylation of the 4'-hydroxyl group and the membrane bound isoform of COMT is more regiospecific than the soluble form. Translation of the soluble form of this enzyme results from utilization of an internal start codon in a full-length mRNA (1.5 kb) or from the translation of a shorter mRNA (1.3 kb), transcribed from an internal promoter. The proposed S₉2-like methylation reaction requires Mg⁺⁺ and is inhibited by Ca⁺⁺. The binding of the donor and substrate to COMT occurs sequentially. AdoMet first binds to COMT in a Mg⁺⁺-independent manner, followed by the binding of Mg⁺⁺ and the binding of the catechol substrate.

The amount of COMT in tissues is relatively high compared to the amount of activity normally required, thus inhibition is problematic. Nonetheless, inhibitors have been developed for in vitro use (e.g., gallates, tropolone, U-0521, and 3',4'-dihydroxy-2-methyl-propiohetroplone) and for clinical use (e.g., nitrocatechol-based compounds and tolcapone). Administration of these inhibitors results in the increased half-life of L-dopa and the consequent formation of dopamine. Inhibition of COMT is also likely to increase the half-life of various other catechol-structure compounds, including but not limited to epinephrine/norepinephrine, isoprenaline, rimeterol, dobutamine, fenoldopam, apomorphine, and α-methyl-dopa. A deficiency in norepinephrine has been linked to clinical depression, hence the use of COMT inhibitors could be useful in the treatment of depression. COMT inhibitors are generally well tolerated with minimal side effects and are ultimately metabolized in the liver with only minor accumulation of metabolites in the body (Männistö, P.T. and Kaakkola, S. (1999) Pharmacological Reviews 51:593-628).

The discovery of new drug metabolizing enzymes, 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 autoimmune/inflammatory, cell proliferative, developmental, endocrine, eye, metabolic, and gastrointestinal disorders, including liver disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of drug metabolizing enzymes.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, drug metabolizing enzymes, referred to

WO 02/12467

PCT/US01/24382

collectively as "DME" and individually as "DME-1," "DME-2," "DME-3," "DME-4," "DME-5,"
"DME-6," "DME-7," "DME-8," "DME-9," "DME-10," "DME-11," "DME-12," "DME-13," "DME-
14," "DME-15," "DME-16," "DME-17," "DME-18," and "DME-19." In one aspect, the invention
provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an
5 amino acid sequence selected from the group consisting of SEQ ID NO:1-19, 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:1-19, c) a biologically active fragment of
a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19,
and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the
10 group consisting of SEQ ID NO:1-19. In one alternative, the invention provides an isolated
polypeptide comprising the amino acid sequence of SEQ ID NO:1-19.

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:1-19, b) a polypeptide comprising a naturally occurring amino acid
15 sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ
ID NO:1-19, c) a biologically active fragment of a polypeptide having an amino acid sequence
selected from the group consisting of SEQ ID NO:1-19, and d) an immunogenic fragment of a
polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19.
In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ
20 ID NO:1-19. In another alternative, the polynucleotide is selected from the group consisting of SEQ
ID NO:20-38.

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
25 of SEQ ID NO:1-19, 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:1-19, c) a
biologically active fragment of a polypeptide having an amino acid sequence selected from the group
consisting of SEQ ID NO:1-19, and d) an immunogenic fragment of a polypeptide having an amino
acid sequence selected from the group consisting of SEQ ID NO:1-19. In one alternative, the
30 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

WO 02/12467

PCT/US01/24382

of SEQ ID NO:1-19 , 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:1-19 , c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 , and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 . 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:1-19 , 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:1-19 , c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 , and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 .

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:20-38, 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:20-38, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:20-38, 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:20-38, 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

WO 02/12467

PCT/US01/24382

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
5 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:20-38, 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:20-38, 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
10 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
15 selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-19, 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:1-19, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19, and d) an immunogenic
20 fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-19. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional DME, comprising administering to a patient in need of such treatment the
25 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:1-19, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected
30 from the group consisting of SEQ ID NO:1-19, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19. The method comprises a) exposing a sample comprising the

WO 02/12467

PCT/US01/24382

polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional DME, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-19, 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:1-19, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional DME, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-19, 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:1-19, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19. 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:1-19, 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:1-19, c) a biologically active fragment of a polypeptide

WO 02/12467

PCT/US01/24382

having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 , and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 . 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:20-38, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:20-38, 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:20-38, 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:20-38, 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:20-38, 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

WO 02/12467

PCT/US01/24382

toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

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

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

Table 3 shows structural features of polypeptide sequences of the invention, including
10 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.

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

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

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

20

DESCRIPTION OF THE INVENTION

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

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

Unless defined otherwise, all technical and scientific terms used herein have the same
meanings as commonly understood by one of ordinary skill in the art to which this invention belongs.

WO 02/12467

PCT/US01/24382

Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

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

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of DME. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of DME either by directly interacting with DME or by acting on components of the biological pathway in which DME participates.

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

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

WO 02/12467

PCT/US01/24382

arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

5 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.

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

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant.
20 Antibodies that bind DME 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
25 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
30 the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA;

WO 02/12467

PCT/US01/24382

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

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

"Conservative amino acid substitutions" are those substitutions that are predicted to least

WO 02/12467

PCT/US01/24382

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

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

30 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
 35 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.

40 "Differential expression" refers to increased or upregulated; or decreased, downregulated, or

WO 02/12467

PCT/US01/24382

absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

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

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

A fragment of SEQ ID NO:20-38 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:20-38, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:20-38 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:20-38 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:20-38 and the region of SEQ ID NO:20-38 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:1-19 is encoded by a fragment of SEQ ID NO:20-38. A fragment of SEQ ID NO:1-19 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-19. For example, a fragment of SEQ ID NO:1-19 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-19. The precise length of a fragment of SEQ ID NO:1-19 and the region of SEQ ID NO:1-19 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

WO 02/12467

PCT/US01/24382

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

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

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

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

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

Matrix: BLOSUM62

WO 02/12467

PCT/US01/24382

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

5 *Expect: 10*

Word Size: 11

Filter: on

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

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

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

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

30 Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version

WO 02/12467

PCT/US01/24382

2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

5 *Gap x drop-off: 50*

Expect: 10

Word Size: 3

Filter: on

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

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

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

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

WO 02/12467

PCT/US01/24382

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

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

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_{0t} or R_{0t} analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

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

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

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of DME which is capable of eliciting an immune response when introduced into a living organism, for example, a

WO 02/12467

PCT/US01/24382

mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of DME 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.

5 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 DME. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of DME.

10 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.

20 "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.

25 "Post-translational modification" of an DME 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 DME.

30 "Probe" refers to nucleic acid sequences encoding DME, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target

WO 02/12467

PCT/US01/24382

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
5 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
10 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
15 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
20 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
25 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
30 PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved

WO 02/12467

PCT/US01/24382

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 DME, nucleic acids encoding DME, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a

WO 02/12467

PCT/US01/24382

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

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

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

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

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

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

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro*

WO 02/12467

PCT/US01/24382

fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation.

5 Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), *supra*.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-10 1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have 15 significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each 20 other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

25 A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 30 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

WO 02/12467

PCT/US01/24382

The invention is based on the discovery of new human drug metabolizing enzymes (DME), the polynucleotides encoding DME, and the use of these compositions for the diagnosis, treatment, or prevention of autoimmune/inflammatory, cell proliferative, developmental, endocrine, eye, metabolic, and gastrointestinal disorders, including liver disorders.

5 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
10 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
15 polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank homolog. Column 5 shows the annotation of the GenBank homolog along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

20 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
25 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
30 properties establish that the claimed polypeptides are drug metabolizing enzymes. For example, SEQ ID NO:3 is 40% identical to a mouse cytochrome P450 monooxygenase (GenBank ID g2653663) as determined by the Basic Local Alignment Search Tool (BLAST, see Table 2). The BLAST probability score is 5.3e-91, which indicates the probability of obtaining the observed polypeptide

WO 02/12467

PCT/US01/24382

sequence alignment by chance. SEQ ID NO:3 also contains cytochrome P450 signature sequences as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains (see Table 3). Data from BLIMPS and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:3 is a member of
5 the cytochrome P450 family.

In an alternative example, SEQ ID NO:1 is 58% identical to a lysyl oxidase from the yellow perch (*Perca flavescens*; GenBank ID g4929199) as determined by BLAST analysis. The BLAST probability score is $1.9e-248$. SEQ ID NO:1 also contains cytochrome P450 signature sequences as determined by searching for statistically significant matches in the HMM-based PFAM database of conserved protein family domains and by BLIMPS analyses.
10

In an alternative example, SEQ ID NO:2 is 61% identical to human flavin-containing monooxygenase 5 (GenBank ID g559046) as determined by BLAST analysis, with a probability score of $4.5e-181$. SEQ ID NO:2 also contains flavin-containing monooxygenase signature sequences as determined by searching for statistically significant matches in the HMM-based PFAM database of conserved protein family domains and by BLIMPS and PROFILESCAN analyses.
15

In an alternative example, SEQ ID NO:4 is 39% identical to a *Pseudomonas* 2,3-butanediol dehydrogenase (GenBank ID g529564) as determined by BLAST analysis, with a probability score of $2.0e-61$. SEQ ID NO:4 also contains dehydrogenase signature sequences as determined by searching for statistically significant matches in the HMM-based PFAM database of conserved protein family domains. Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:4 is a dehydrogenase.
20

SEQ ID NO:5 is 54% identical to a *Bacillus* quinone oxidase (GenBank ID g2633069) as determined by BLAST analysis, with a probability score of $7.1e-96$. Data obtained by searching the HMM-based PFAM database of conserved protein family domains and by BLIMPS analyses provide
25 further corroborative evidence that SEQ ID NO:5 is a quinone oxidase.

In an alternative example, SEQ ID NO:6 is 92% identical to mouse heparan sulfate 6-sulfotransferase 2 (GenBank ID g6683558) as determined by BLAST analysis, with a probability score of $2.3e-255$.

In an alternative example, SEQ ID NO:7 is 90% identical to a human glutathione S-transferase subunit (GenBank ID g242749) as determined by BLAST analysis, with a probability score of $1.3e-101$. SEQ ID NO:7 also contains glutathione S-transferase signature sequences as determined by searching for statistically significant matches in the HMM-based PFAM database of conserved protein family domains and by BLIMPS analyses.
30

WO 02/12467

PCT/US01/24382

In an alternative example, SEQ ID NO:8 is 40% identical to a human steroid dehydrogenase (GenBank ID g5531815) as determined by BLAST analysis, with a probability score of 1.9e-56. SEQ ID NO:8 also contains dehydrogenase signature sequences as determined by searching for statistically significant matches in the HMM-based PFAM database of conserved protein family domains.

5 SEQ ID NO:9 is 47% identical to a rabbit liver carboxylesterase (GenBank ID g3219695) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 6.3e-72, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:9 also contains carboxylesterase domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:9 is a
10 carboxylesterase.

SEQ ID NO:10 is 45% identical to human carboxylesterase (GenBank ID g180950) as determined by the Basic Local Alignment Search Tool (BLAST, see Table 2). The BLAST
15 probability score is 8.7e-130, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:10 also contains carboxylesterase domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains (see Table 3). Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:10
20 is a carboxylesterase.

In an alternative example, SEQ ID NO:11 is 89% identical to murine heparan sulfate 6-sulfotransferase 2 (GenBank ID g6683558) as determined by the Basic Local Alignment Search Tool (BLAST, see Table 2). The BLAST probability score is 1.8e-236, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance, and provides evidence that SEQ ID
25 NO:11 is a DME, and specifically that SEQ ID NO:11 is a sulfotransferase.

In an alternative example, SEQ ID NO:12 is 25% identical to a *Bacillus subtilis* epoxide hydrolase (GenBank ID g2633182) as determined by the Basic Local Alignment Search Tool (BLAST, see Table 2). The BLAST probability score is 1.3e-11, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:12 also contains
30 hydrolase domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains (see Table 3). Data from BLIMPS analyses provide further corroborative evidence that SEQ ID NO:12 is a hydrolase.

WO 02/12467

PCT/US01/24382

In an alternative example, SEQ ID NO:13 is 83% identical to a rat beta-alanine-pyruvate aminotransferase (GenBank ID g1944136) as determined by the BLAST analysis (see Table 2). The BLAST probability score is $1.1e-234$. SEQ ID NO:13 also contains aminotransferase domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains (see Table 3). Data from BLIMPS and PROFLESCAN analyses provide further corroborative evidence that SEQ ID NO:13 is an aminotransferase.

In an alternative example, SEQ ID NO:14 is 50% identical to a guinea pig hydroxysteroid sulfotransferase (GenBank ID g1151081) as determined by the BLAST analysis (see Table 2). The BLAST probability score is $5.4e-34$, and provides evidence that SEQ ID NO:14 is a sulfotransferase.

In an alternative example, SEQ ID NO:15 is 52% identical to a guinea pig copper/zinc superoxide dismutase (GenBank ID g1066120) as determined by the BLAST analysis (see Table 2). The BLAST probability score is $2.1e-25$. SEQ ID NO:15 also contains copper/zinc superoxide dismutase domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains (see Table 3). Data from BLIMPS analyses provide further corroborative evidence that SEQ ID NO:15 is a copper/zinc superoxide dismutase.

SEQ ID NO:16 is 37% identical to human 3'-phosphoadenylylsulfate-galactosylceramide 3'-sulfotransferase (cerebroside sulfotransferase, GenBank ID g1871141) as determined by the Basic Local Alignment Search Tool (BLAST, see Table 2). The BLAST probability score is $2.8e-60$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance.

In an alternative example, SEQ ID NO:17 is 38% identical to a putative *C. elegans* monoamine oxidase (GenBank ID g6782275) as determined by BLAST analysis with a probability score of $3.0e-99$. SEQ ID NO:17 also contains a monoamine oxidase 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 analysis provide further corroborative evidence that SEQ ID NO:17 is a monoamine oxidase.

In an alternative example, SEQ ID NO:18 is 36% identical to human catechol-*O*-methyltransferase (GenBank ID g179955) as determined by BLAST analysis with a probability score of $9.5e-41$. SEQ ID NO:18 is also 36% identical to murine catechol-*O*-methyltransferase (GenBank ID g3493253) as determined by BLAST analysis with a probability score of $1.3e-41$.

In an alternative example, SEQ ID NO:19 is 44% identical to *Fundulus heteroclitus* cytochrome P450 2N1 (GenBank ID g5852342) as determined by the Basic Local Alignment Search

WO 02/12467

PCT/US01/24382

Tool (BLAST, see Table 2). The BLAST probability score is $4.6e-99$, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:19 also contains cytochrome P450 domains as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains (see Table 5). 3). Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:19 is a cytochrome P450.

The algorithms and parameters for the analysis of SEQ ID NO:1-19 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:20-38 or that distinguish between SEQ ID NO:20-38 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5') and stop (3') positions of the cDNA and/or genomic sequences in column 5 relative to their respective full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 7690384J1 is the identification number of an Incyte cDNA sequence, and PROSTME06 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 55017748J1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g1203094) which contributed to the assembly of the full length polynucleotide sequences. In addition, the identification numbers in column 5 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences including the designation "ENST"). Alternatively, the identification numbers in column 5 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records

WO 02/12467

PCT/US01/24382

(i.e., those sequences including the designation "NP"). Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, FL_XXXXXX_N₁N₂YYYY_N₃N₄ 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 identification numbers in column 5 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, FLXXXXXX_gAAAAA_gBBBBB_1_N is the identification number of 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).
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).
INCY	Full length transcript and exon prediction from mapping of EST sequences to the genome. Genomic location and EST composition data are combined to predict the exons and resulting transcript.

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

WO 02/12467

PCT/US01/24382

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

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

10 The invention also encompasses polynucleotides which encode DME. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:20-38, which encodes DME. The polynucleotide sequences of SEQ ID NO:20-38, 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
15 backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding DME. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding DME. A particular aspect of the invention encompasses a variant of a
20 polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:20-38 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:20-38. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of DME.

25 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 DME, 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
30 accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring DME, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode DME and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring DME under appropriately selected

WO 02/12467

PCT/US01/24382

conditions of stringency, it may be advantageous to produce nucleotide sequences encoding DME 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 DME 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 DME and DME derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding DME or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:20-38 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

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

The nucleic acid sequences encoding DME may be extended utilizing a partial nucleotide

WO 02/12467

PCT/US01/24382

sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.)

5 Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-3060).

10 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

20 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

25 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

30

WO 02/12467

PCT/US01/24382

which may be present in limited amounts in a particular sample.

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

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

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) *Nat. Biotechnol.* 17:793-797; Christians, F.C. et al. (1999) *Nat. Biotechnol.* 17:259-264; and Cramer, A. et al. (1996) *Nat. Biotechnol.* 14:315-319) to alter or improve the biological properties of DME, 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 DME 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, DME 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.,

WO 02/12467

PCT/US01/24382

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 DME, or any part thereof, may be altered during direct synthesis and/or combined with sequences
5 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.
10 (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active DME, the nucleotide sequences encoding DME 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
15 inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding DME. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding DME. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding DME and its initiation codon and upstream regulatory sequences are inserted into
20 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
25 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 DME and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in
30 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

WO 02/12467

PCT/US01/24382

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

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding DME. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding DME 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 DME 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 DME are needed, e.g. for the production of antibodies, vectors which direct high level expression of DME 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 DME. 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

WO 02/12467

PCT/US01/24382

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 DME. Transcription of sequences encoding
5 DME 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
10 constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., *The McGraw Hill Yearbook of Science and Technology* (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases
15 where an adenovirus is used as an expression vector, sequences encoding DME 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 DME 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-
20 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-
25 355.)

For long term production of recombinant proteins in mammalian systems, stable expression of
DME in cell lines is preferred. For example, sequences encoding DME 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
30 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

WO 02/12467

PCT/US01/24382

culture techniques appropriate to the cell type.

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

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding DME is inserted within a marker gene sequence, transformed cells containing sequences encoding DME can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding DME 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 DME and that express DME 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 membranes, 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 DME using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on DME is preferred, but a competitive

WO 02/12467

PCT/US01/24382

binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, 5 Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polymucleotides encoding DME include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the 10 sequences encoding DME, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and 15 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 DME may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein 20 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 polymucleotides which encode DME may be designed to contain signal sequences which direct secretion of DME through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the 25 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 30 (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

WO 02/12467

PCT/US01/24382

sequences encoding DME may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric DME protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of DME activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the DME encoding sequence and the heterologous protein sequence, so that DME 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 DME 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.

DME of the present invention or fragments thereof may be used to screen for compounds that specifically bind to DME. At least one and up to a plurality of test compounds may be screened for specific binding to DME. 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 DME, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coigan, 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 DME 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 DME, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, *Drosophila*, or *E.*

WO 02/12467

PCT/US01/24382

coli. Cells expressing DME or cell membrane fractions which contain DME are then contacted with a test compound and binding, stimulation, or inhibition of activity of either DME or the compound is analyzed.

5 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 DME, either in solution or affixed to a solid support, and detecting the binding of DME to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor.

10 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.

DME of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of DME. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for DME activity, wherein DME is combined with at least one test compound, and the activity of DME in the presence of a test compound is compared with the activity of DME in the absence of the test compound. A change in the activity of DME in the presence of the test compound is indicative of a compound that modulates the activity of DME. Alternatively, a test compound is combined with an *in vitro* or cell-free system comprising DME under conditions suitable for DME activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of DME 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 DME or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (*neo*; Capecchi, M.R. (1989) *Science* 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) *Clin. Invest.* 97:1999-2002; Wagner, K.U. et al. (1997) *Nucleic Acids Res.* 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such

WO 02/12467

PCT/US01/24382

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.

5 Polynucleotides encoding DME 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).

10 Polynucleotides encoding DME can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding DME 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
15 potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress DME, e.g., by secreting DME in its milk, may also serve as a convenient source of that protein (Jaime, 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
20 regions of DME and drug metabolizing enzymes. In addition, the expression of DME is closely associated with a variety of diseased tissues, including that of the brain, prostate, bone, intestine, and breast. Therefore, DME appears to play a role in autoimmune/inflammatory, cell proliferative, developmental, endocrine, eye, metabolic, and gastrointestinal disorders, including liver disorders. In the treatment of disorders associated with increased DME expression or activity, it is desirable to
25 decrease the expression or activity of DME. In the treatment of disorders associated with decreased DME expression or activity, it is desirable to increase the expression or activity of DME.

Therefore, in one embodiment, DME 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 DME. Examples of such disorders include, but are not limited to, an autoimmune/inflammatory disorder, such
30 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,

WO 02/12467

PCT/US01/24382

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,

5 osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cell proliferative disorder, such as actinic keratosis, arteriosclerosis,

10 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,

15 pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; 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

20 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; an endocrine disorder, such as disorders of the hypothalamus and pituitary resulting from lesions such as primary brain tumors, adenomas, infarction associated with pregnancy, hypophysectomy, aneurysms,

25 vascular malformations, thrombosis, infections, immunological disorders, and complications due to head trauma; disorders associated with hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism; disorders associated with hyperpituitarism including acromegaly, gigantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH)

30 often caused by benign adenoma; disorders associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism; disorders associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema,

WO 02/12467

PCT/US01/24382

toxic multinodular goiter, thyroid carcinoma, and Plummer's disease; disorders associated with hyperparathyroidism including Conn disease (chronic hypercalcemia); pancreatic disorders such as Type I or Type II diabetes mellitus and associated complications; disorders associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's disease, Liddle's syndrome, and Arnold-Healy-Gurdon syndrome, pheochromocytoma tumors, and Addison's disease; disorders associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbations of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis; and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, hypergonadal disorders associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α -reductase, and gynecomastia; an eye disorder, such as conjunctivitis, keratoconjunctivitis sicca, keratitis, episcleritis, iritis, posterior uveitis, glaucoma, amaurosis fugax, ischemic optic neuropathy, optic neuritis, Leber's hereditary optic neuropathy, toxic optic neuropathy, vitreous detachment, retinal detachment, cataract, macular degeneration, central serous chorioretinopathy, retinitis pigmentosa, melanoma of the choroid, retinoblastoma, and chiasmal tumor; a metabolic disorder, such as Addison's disease, cerebrotendinous xanthomatosis, congenital adrenal hyperplasia, coumarin resistance, cystic fibrosis, diabetes, fatty hepatocirrhosis, fructose-1,6-diphosphatase deficiency, galactosemia, goiter, glucagonoma, glycogen storage diseases, hereditary fructose intolerance, hyperadrenalism, hypoadrenalism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia, hyperthyroidism, hypoglycemia, hypothyroidism, hyperlipidemia, hyperlipemia, lipid myopathies, lipodystrophies, lysosomal storage diseases, Menkes syndrome, occipital horn syndrome, mannosidosis, neuraminidase deficiency, obesity, pentosuria phenylketonuria, pseudovitamin D-deficiency rickets; and a gastrointestinal disorder, such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, hereditary hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice,

WO 02/12467

PCT/US01/24382

hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha₁-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas.

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

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

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

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

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

20 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 DME may be produced using methods which are generally known in the art.

WO 02/12467

PCT/US01/24382

In particular, purified DME may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind DME. Antibodies to DME 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
5 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 DME or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to
10 increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

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

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

WO 02/12467

PCT/US01/24382

D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

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

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

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between DME and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering DME 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 DME. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of DME-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 DME epitopes, represents the average affinity, or avidity, of the antibodies for DME. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular DME 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 DME-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 DME, preferably in active form, from the antibody (Catty, D. (1988) *Antibodies, Volume I: A Practical Approach*, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) *A Practical Guide to Monoclonal Antibodies*, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine

WO 02/12467

PCT/US01/24382

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

In another embodiment of the invention, the polynucleotides encoding DME, 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,
10 RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding DME. 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 DME. (See, e.g., Agrawal, S., ed. (1996) *Antisense Therapeutics*, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense
15 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
20 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 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding DME 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
30 immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaise, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassamias, familial

WO 02/12467

PCT/US01/24382

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.,

5 against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA*. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as *Candida albicans* and *Paracoccidioides brasiliensis*; and protozoan parasites such as *Plasmodium falciparum* and *Trypanosoma cruzi*). In the case where a genetic deficiency in DME expression or regulation causes disease, the expression of

10 DME 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 DME are treated by constructing mammalian expression vectors encoding DME and introducing these vectors by mechanical means into DME-deficient cells. Mechanical transfer technologies for use with

15 cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J.-L. and H. Récipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

20 Expression vectors that may be effective for the expression of DME include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, FVAX, PCR2-TOPOTA vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSE/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). DME may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous

25 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

30 FK506/rapamycin inducible promoter, or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. supra), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding DME from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID

WO 02/12467

PCT/US01/24382

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 DME expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding DME under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) *J. Virol.* 61:1647-1650; Bender, M.A. et al. (1987) *J. Virol.* 61:1639-1646; Adam, M.A. and A.D. Miller (1988) *J. Virol.* 62:3802-3806; Dull, T. et al. (1998) *J. Virol.* 72:8463-8471; Zufferey, R. et al. (1998) *J. Virol.* 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) *J. Virol.* 71:7020-7029; Bauer, G. et al. (1997) *Blood* 89:2259-2267; Bonyhadi, M.L. (1997) *J. Virol.* 71:4707-4716; Ranga, U. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95:1201-1206; Su, L. (1997) *Blood* 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding DME to cells which have one or more genetic abnormalities with respect to the expression of DME. 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 Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"),

WO 02/12467

PCT/US01/24382

hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) *Annu. Rev. Nutr.* 19:511-544 and Verma, I.M. and N. Somia (1997) *Nature* 18:389-239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver
5 polynucleotides encoding DME to target cells which have one or more genetic abnormalities with respect to the expression of DME. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing DME 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
10 been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) *Exp. Eye Res.* 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV δ 92 which consists of a genome containing at least one exogenous gene to be transferred to a
15 cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of
20 multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

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

WO 02/12467

PCT/US01/24382

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 DME into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

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

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

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

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding DME. Such DNA sequences may be incorporated into a wide variety of vectors

WO 02/12467

PCT/US01/24382

with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

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

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding DME. 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 DME expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding DME may be therapeutically useful, and in the treatment of disorders associated with decreased DME expression or activity, a compound which specifically promotes expression of the polynucleotide encoding DME 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 DME 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 DME are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected

WO 02/12467

PCT/US01/24382

by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding DME. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13).

5
10

A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

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

15

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.

20

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 DME, antibodies to DME, and mimetics, agonists, antagonists, or inhibitors of DME.

25

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.

30

Compositions for pulmonary administration may be prepared in liquid or dry powder form.

WO 02/12467

PCT/US01/24382

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
5 have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination
10 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 DME or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, DME or a fragment thereof may be joined to a short cationic N-
15 terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

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

A therapeutically effective dose refers to that amount of active ingredient, for example DME or fragments thereof, antibodies of DME, and agonists, antagonists or inhibitors of DME, which
25 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 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
30 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. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the

WO 02/12467

PCT/US01/24382

patient, and the route of administration.

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

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

15 DIAGNOSTICS

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

WO 02/12467

PCT/US01/24382

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 DME may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of DME, and to monitor regulation of DME levels during therapeutic intervention.

5 In one aspect, hybridization with FCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding DME or closely related molecules may be used to identify nucleic acid sequences which encode DME. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the
10 probe identifies only naturally occurring sequences encoding DME, 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 DME 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:20-38 or from
15 genomic sequences including promoters, enhancers, and introns of the DME gene.

Means for producing specific hybridization probes for DNAs encoding DME include the cloning of polynucleotide sequences encoding DME or DME 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
20 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 DME may be used for the diagnosis of disorders associated with expression of DME. Examples of such disorders include, but are not limited to, an
25 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
30 lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hyper eosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's

WO 02/12467

PCT/US01/24382

syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a cell proliferative disorder, such as

5 actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocytopenia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone,

10 bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a 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,

15 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; an endocrine disorder, such as disorders of the hypothalamus and pituitary resulting from lesions such as primary brain tumors, adenomas, infarction associated with pregnancy,

20 hypophysectomy, aneurysms, vascular malformations, thrombosis, infections, immunological disorders, and complications due to head trauma; disorders associated with hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism; disorders associated with hyperpituitarism including acromegaly, gigantism, and syndrome of inappropriate antidiuretic

25 hormone (ADH) secretion (SIADH) often caused by benign adenoma; disorders associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism; disorders associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's

30 disease; disorders associated with hyperparathyroidism including Conn disease (chronic hypercalcemia); pancreatic disorders such as Type I or Type II diabetes mellitus and associated complications; disorders associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's

WO 02/12467

PCT/US01/24382

disease, Liddle's syndrome, and Arnold-Healy-Gordon syndrome, pheochromocytoma tumors, and Addison's disease; disorders associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbations of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea,

5 hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis; and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, hypergonadal disorders associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α -reductase, and gynecomastia; an eye disorder, such as conjunctivitis, keratoconjunctivitis sicca, keratitis, episcleritis, iritis, posterior uveitis, glaucoma,

10 amaurosis fugax, ischemic optic neuropathy, optic neuritis, Leber's hereditary optic neuropathy, toxic optic neuropathy, vitreous detachment, retinal detachment, cataract, macular degeneration, central serous chorioretinopathy, retinitis pigmentosa, melanoma of the choroid, retrobulbar tumor, and chiasmal tumor; a metabolic disorder, such as Addison's disease, cerebrotendinous xanthomatosis, congenital adrenal hyperplasia, coumarin resistance, cystic fibrosis, diabetes, fatty hepatocirrhosis,

15 fructose-1,6-diphosphatase deficiency, galactosemia, goiter, glucagonoma, glycogen storage diseases, hereditary fructose intolerance, hyperadrenalism, hypoadrenalism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia, hyperthyroidism, hypoglycemia, hypothyroidism, hyperlipidemia, hyperlipemia, lipid myopathies, lipodystrophies, lysosomal storage diseases, Menkes syndrome, occipital horn syndrome, mannosidosis, neuraminidase deficiency, obesity, pentosuria

20 phenylketonuria, pseudovitamin D-deficiency rickets; and a gastrointestinal disorder, such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma,

25 biliary tract disease, hepatitis, hyperbilirubinemia, hereditary hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal

30 syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha₁-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors

WO 02/12467

PCT/US01/24382

including nodular hyperplasias, adenomas, and carcinomas. The polynucleotide sequences encoding DME 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 DME expression. Such qualitative or quantitative
5 methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding DME may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding DME 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
10 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 DME 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
15 the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of DME, 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 DME, under conditions suitable for hybridization or amplification.
20 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,
25 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.

30 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

WO 02/12467

PCT/US01/24382

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 DME 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 DME, or a fragment of a polynucleotide complementary to the polynucleotide encoding DME, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

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

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

WO 02/12467

PCT/US01/24382

quantitation.

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

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

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

WO 02/12467

PCT/US01/24382

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

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

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl

WO 02/12467

PCT/US01/24382

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

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

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

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

WO 02/12467

PCT/US01/24382

residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

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

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

In another embodiment of the invention, nucleic acid sequences encoding DME 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 plasmid 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 DME on a

WO 02/12467

PCT/US01/24382

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.

5 Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any
10 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, DME, its catalytic or immunogenic fragments, or
15 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 DME and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds
20 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 DME, or fragments thereof, and washed. Bound DME is then detected by methods well known in the art. Purified DME can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively,
25 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 DME specifically compete with a test compound for binding DME. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with DME.

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

WO 02/12467

PCT/US01/24382

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

5 The disclosures of all patents, applications and publications, mentioned above and below, including U.S. Ser. No. 60/223,055, U.S. Ser. No. 60/224,728, U.S. Ser. No. 60/226,440, U.S. Ser. No. 60/228,067, U.S. Ser. No. 60/230,063, U.S. Ser. No. 60/232,244, and U.S. Ser. No. 60/234,269, are expressly incorporated by reference herein.

10

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were
15 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.

20 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
25 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 SUPERSCRIP^T plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra,
30 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

WO 02/12467

PCT/US01/24382

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., EBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), PCR2-TOPOTA (Invitrogen), or 5 pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* 10 excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 15 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using 20 PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such 25 as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). 30 Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading

WO 02/12467

PCT/US01/24382

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
5 vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family
10 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
15 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
20 begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering,
25 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
30 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

WO 02/12467

PCT/US01/24382

applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:20-38. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

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

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene

WO 02/12467

PCT/US01/24382

identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence. Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbprj public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

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

VI. Chromosomal Mapping of DME Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:20-38 were compared with

WO 02/12467

PCT/US01/24382

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:20-38 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Génethon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

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

VII. Analysis of Polynucleotide Expression

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

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

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

The product score takes into account both the degree of similarity between two sequences and the

WO 02/12467

PCT/US01/24382

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

VIII. Extension of DME Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30

WO 02/12467

PCT/US01/24382

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
5 extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg²⁺, (NH₄)₂SO₄, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE
10 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
15 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

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

The extended nucleotides were desalted and concentrated, transferred to 384-well plates,
25 digested with CviII cholera virus endonuclease (Molecular Biology Research, Madison WI), and
sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For
shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%)
agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended
clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector
30 (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.

WO 02/12467

PCT/US01/24382

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

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

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:20-38 are employed to screen cDNAs,
15 genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 µCi of [γ -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase
20 (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10⁷ counts per minute of the labeled probe is used in a typical membrane-based 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).

25 The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and
30 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

WO 02/12467

PCT/US01/24382

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.

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 (2 lmer), 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 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

WO 02/12467

PCT/US01/24382

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

5 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).

10 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

15 oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 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.

20 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.

25 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

30 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.

WO 02/12467

PCT/US01/24382

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
5 focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477,
10 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,
15 although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

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

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital
25 (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission
30 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

WO 02/12467

PCT/US01/24382

for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

XI. Complementary Polynucleotides

Sequences complementary to the DME-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring DME. 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 DME. 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 DME-encoding transcript.

XII. Expression of DME

Expression and purification of DME is achieved using bacterial or virus-based expression systems. For expression of DME 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 DME upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of DME 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 DME 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, DME is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from *Schistosoma japonicum*, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from DME at

WO 02/12467

PCT/US01/24382

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 DME obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, and XVIII, where applicable.

XIII. Functional Assays

DME function is assessed by expressing the sequences encoding DME 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.

The influence of DME on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding DME 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

WO 02/12467

PCT/US01/24382

NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding DME and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIV. Production of DME Specific Antibodies

5 DME 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 DME 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-15 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-DME activity by, for example, binding the peptide or DME to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit 20 IgG.

XV. Purification of Naturally Occurring DME Using Specific Antibodies

Naturally occurring or recombinant DME is substantially purified by immunoaffinity chromatography using antibodies specific for DME. An immunoaffinity column is constructed by covalently coupling anti-DME antibody to an activated chromatographic resin, such as CNBr-activated 25 SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing DME are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of DME (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/DME binding 30 (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and DME is collected.

XVI. Identification of Molecules Which Interact with DME

WO 02/12467

PCT/US01/24382

DME, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled DME, washed, and any wells with labeled DME complex are assayed. Data obtained using different concentrations of DME are used to calculate values for the number, affinity, and association of DME with the candidate molecules.

Alternatively, molecules interacting with DME 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).

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

XVII. Demonstration of DME Activity

Cytochrome P450 activity of DME is measured using the 4-hydroxylation of aniline. Aniline is converted to 4-aminophenol by the enzyme, and has an absorption maximum at 630 nm (Gibson and Skett, *supra*). This assay is a convenient measure, but underestimates the total hydroxylation, which also occurs at the 2- and 3- positions. Assays are performed at 37 °C and contain an aliquot of the enzyme and a suitable amount of aniline (approximately 2 mM) in reaction buffer. For this reaction, the buffer must contain NADPH or an NADPH-generating cofactor system. One formulation for this reaction buffer includes 85 mM Tris pH 7.4, 15 mM MgCl_2 , 50 mM nicotinamide, 40 mg trisodium isocitrate, and 2 units isocitrate dehydrogenase, with 8 mg NADP^+ added to a 10 mL reaction buffer stock just prior to assay. Reactions are carried out in an optical cuvette, and the absorbance at 630 nm is measured. The rate of increase in absorbance is proportional to the enzyme activity in the assay. A standard curve can be constructed using known concentrations of 4-aminophenol.

$1\alpha,25$ -dihydroxyvitamin D 24-hydroxylase activity of DME is determined by monitoring the conversion of ^3H -labeled $1\alpha,25$ -dihydroxyvitamin D ($1\alpha,25(\text{OH})_2\text{D}$) to $24,25$ -dihydroxyvitamin D ($24,25(\text{OH})_2\text{D}$) in transgenic rats expressing DME. 1 μg of $1\alpha,25(\text{OH})_2\text{D}$ dissolved in ethanol (or ethanol alone as a control) is administered intravenously to approximately 6-week-old male transgenic rats expressing DME or otherwise identical control rats expressing either a defective variant of DME or not expressing DME. The rats are killed by decapitation after 8 hrs, and the kidneys are rapidly removed, rinsed, and homogenized in 9 volumes of ice-cold buffer (15 mM Tris-acetate (pH 7.4), 0.19 M sucrose, 2 mM magnesium acetate, and 5 mM sodium succinate). A portion (e.g., 3 ml) of each

WO 02/12467

PCT/US01/24382

homogenate is then incubated with 0.25 nM $1\alpha,25(\text{OH})_2[1\text{-}^3\text{H}]\text{D}$, with a specific activity of approximately 3.5 GBq/mmol, for 15 min at 37 °C under oxygen with constant shaking. Total lipids are extracted as described (Bligh, E.G. and W.J. Dyer (1959) *Can. J. Biochem. Physiol.* 37: 911-917) and the chloroform phase is analyzed by HPLC using a FINNPAK SIL column (JASCO, Tokyo, Japan) with a *n*-hexane/chloroform/methanol (10:2.5:1.5) solvent system at a flow rate of 1 ml/min. In the alternative, the chloroform phase is analyzed by reverse phase HPLC using a J SPHERE ODS-AM column (YMC Co. Ltd., Kyoto, Japan) with an acetonitrile buffer system (40 to 100%, in water, in 30 min) at a flow rate of 1 ml/min. The eluates are collected in fractions of 30 seconds (or less) and the amount of ^3H present in each fraction is measured using a scintillation counter. By comparing the chromatograms of control samples (i.e., samples comprising $1\alpha,25$ -dihydroxyvitamin D or $24,25$ -dihydroxyvitamin D ($24,25(\text{OH})_2\text{D}$), with the chromatograms of the reaction products, the relative mobilities of the substrate ($1\alpha,25(\text{OH})_2[1\text{-}^3\text{H}]\text{D}$) and product ($24,25(\text{OH})_2[1\text{-}^3\text{H}]\text{D}$) are determined and correlated with the fractions collected. The amount of $24,25(\text{OH})_2[1\text{-}^3\text{H}]\text{D}$ produced in control rats is subtracted from that of transgenic rats expressing DME. The difference in the production of $24,25(\text{OH})_2[1\text{-}^3\text{H}]\text{D}$ in the transgenic and control animals is proportional to the amount of 25-hydroxylase activity of DME present in the sample. Confirmation of the identity of the substrate and product(s) is confirmed by means of mass spectroscopy (Miyamoto, Y. et al. (1997) *J. Biol. Chem.* 272:14115-14119).

Flavin-containing monooxygenase activity of DME is measured by chromatographic analysis of metabolic products. For example, Ring, B. J. et al. (1999; *Drug Metab. Dis.* 27:1099-1103) incubated FMO in 0.1 M sodium phosphate buffer (pH 7.4 or 8.3) and 1 mM NADPH at 37 °C, stopped the reaction with an organic solvent, and determined product formation by HPLC. Alternatively, activity is measured by monitoring oxygen uptake using a Clark-type electrode. For example, Ziegler, D. M. and Poulsen, L. L. (1978; *Methods Enzymol.* 52:142-151) incubated the enzyme at 37 °C in an NADPH-generating cofactor system (similar to the one described above) containing the substrate methimazole. The rate of oxygen uptake is proportional to enzyme activity.

UDP glucuronyltransferase activity of DME is measured using a colorimetric determination of free amine groups (Gibson and Skett, *supra*). An amine-containing substrate, such as 2-aminophenol, is incubated at 37 °C with an aliquot of the enzyme in a reaction buffer containing the necessary cofactors (40 mM Tris pH 8.0, 7.5 mM MgCl_2 , 0.025% Triton X-100, 1 mM ascorbic acid, 0.75 mM UDP-glucuronic acid). After sufficient time, the reaction is stopped by addition of ice-cold 20% trichloroacetic acid in 0.1 M phosphate buffer pH 2.7, incubated on ice, and centrifuged to clarify the supernatant. Any unreacted 2-aminophenol is destroyed in this step. Sufficient freshly-prepared

WO 02/12467

PCT/US01/24382

sodium nitrite is then added; this step allows formation of the diazonium salt of the glucuronidated product. Excess nitrite is removed by addition of sufficient ammonium sulfamate, and the diazonium salt is reacted with an aromatic amine (for example, N-naphthylethylenediamine) to produce a colored azo compound which can be assayed spectrophotometrically (at 540 nm for the example). A standard curve can be constructed using known concentrations of aniline, which will form a chromophore with similar properties to 2-aminophenol glucuronide.

Sulfotransferase activity of DME is measured using the incorporation of ^{35}S from [^{35}S]PAPS into a model substrate such as phenol (Folds, A. and Meek, J. L. (1973) *Biochim. Biophys. Acta* 327:365-374). An aliquot of enzyme is incubated at 37 °C with 1 mL of 10 mM phosphate buffer pH 6.4, 50 μM phenol, 0.4-4.0 μM [^{35}S]PAPS. After sufficient time for 5-20% of the radiolabel to be transferred to the substrate, 0.2 mL of 0.1 M barium acetate is added to precipitate protein and phosphate buffer. Then 0.2 mL of 0.1 M $\text{Ba}(\text{OH})_2$ is added, followed by 0.2 mL ZnSO_4 . The supernatant is cleared by centrifugation, which removes proteins as well as unreacted [^{35}S]PAPS. Radioactivity in the supernatant is measured by scintillation. The enzyme activity is determined from the number of moles of radioactivity in the reaction product.

Glutathione S-transferase activity of DME is measured using a model substrate, such as 2,4-dinitro-1-chlorobenzene, which reacts with glutathione to form a product, 2,4-dinitrophenyl-glutathione, that has an absorbance maximum at 340 nm. It is important to note that GSTs have differing substrate specificities, and the model substrate should be selected based on the substrate preferences of the GST of interest. Assays are performed at ambient temperature and contain an aliquot of the enzyme in a suitable reaction buffer (for example, 1 mM glutathione, 1 mM dinitrochlorobenzene, 90 mM potassium phosphate buffer pH 6.5). Reactions are carried out in an optical cuvette, and the absorbance at 340 nm is measured. The rate of increase in absorbance is proportional to the enzyme activity in the assay.

N-acyltransferase activity of DME is measured using radiolabeled amino acid substrates and measuring radiolabel incorporation into conjugated products. Enzyme is incubated in a reaction buffer containing an unlabeled acyl-CoA compound and radiolabeled amino acid, and the radiolabeled acyl-conjugates are separated from the unreacted amino acid by extraction into n-butanol or other appropriate organic solvent. For example, Johnson, M. R. et al. (1990; *J. Biol. Chem.* 266:10227-10233) measured bile acid-CoA:amino acid N-acyltransferase activity by incubating the enzyme with cholesteryl-CoA and ^3H -glycine or ^3H -taurine, separating the tritiated cholate conjugate by extraction into n-butanol, and measuring the radioactivity in the extracted product by scintillation. Alternatively, N-acyltransferase activity is measured using the spectrophotometric determination of reduced CoA

WO 02/12467

PCT/US01/24382

(CoASH) described below.

N-acetyltransferase activity of DME is measured using the transfer of radiolabel from [¹⁴C]acetyl-CoA to a substrate molecule (for example, see Deguchi, T. (1975) *J. Neurochem.* 24:1083-5). Alternatively, a newer spectrophotometric assay based on DTNB (5,5'-dithio-bis(2-nitrobenzoic acid; Ellman's reagent) reaction with CoASH may be used. Free thiol-containing CoASH is formed during N-acetyltransferase catalyzed transfer of an acetyl group to a substrate. CoASH is detected using the absorbance of DTNB conjugate at 412 nm (De Angelis, J. et al. (1997) *J. Biol. Chem.* 273:3045-3050). Enzyme activity is proportional to the rate of radioactivity incorporation into substrate, or the rate of absorbance increase in the spectrophotometric assay.

Protein arginine methyltransferase activity of DME is measured at 37 °C for various periods of time. S-adenosyl-L-[methyl-³H]methionine ([³H]AdoMet; specific activity = 75 Ci/mmol; NEN Life Science Products) is used as the methyl-donor substrate. Useful methyl-accepting substrates include glutathione S-transferase fibrillarlin glycine-arginine domain fusion protein (GST-GAR), heterogeneous nuclear ribonucleoprotein (hnRNP), or hypomethylated proteins present in lysates from adenosine dialdehyde-treated cells. Methylation reactions are stopped by adding SDS-PAGE sample buffer. The products of the reactions are resolved by SDS-PAGE and visualized by fluorography. The presence of ³H-labeled methyl-donor substrates is indicative of protein arginine methyltransferase activity of DME (Tang, J. et al. (2000) *J. Biol. Chem.* 275:7723-7730 and Tang, J. et al. (2000) *J. Biol. Chem.* 275:19866-19876).

Aldo/keto reductase activity of DME is measured using the decrease in absorbance at 340 nm as NADPH is consumed. A standard reaction mixture is 135 mM sodium phosphate buffer (pH 6.2-7.2 depending on enzyme), 0.2 mM NADPH, 0.3 M lithium sulfate, 0.5-2.5 µg enzyme and an appropriate level of substrate. The reaction is incubated at 30 °C and the reaction is monitored continuously with a spectrophotometer. Enzyme activity is calculated as mol NADPH consumed / µg of enzyme.

Alcohol dehydrogenase activity of DME is measured using the increase in absorbance at 340 nm as NAD⁺ is reduced to NADH. A standard reaction mixture is 50 mM sodium phosphate, pH 7.5, and 0.25 mM EDTA. The reaction is incubated at 25 °C and monitored using a spectrophotometer. Enzyme activity is calculated as mol NADH produced / µg of enzyme.

DME activity is determined using 4-methylumbelliferyl acetate as a substrate. The enzymatic reaction is initiated by adding approximately 10 µl of DME-containing sample to 1 ml of reaction buffer (90 mM KH₂PO₄, 40 mM KCl, pH 7.3) with 0.5 mM 4-methylumbelliferyl acetate at 37 °C. The production of 4-methylumbelliferone is monitored with a spectrophotometer ($\epsilon_{330} =$

WO 02/12467

PCT/US01/24382

12.2 mM⁻¹ cm⁻¹) for 1.5 min. Specific activity is expressed as micromoles of product formed per minute per milligram of protein and corresponds to the activity of DME in the sample (Evgenia, V. et al. (1997) *J. Biol. Chem.* 272:14769-14775).

In the alternative, the cocaine benzoyl ester hydrolase activity of DME is measured by
5 incubating approximately 0.1 ml of enzyme 3.3 mM cocaine in reaction buffer (50 mM NaH₂PO₄,
pH 7.4) with 1 mM benzamidine, 1 mM EDTA, and 1 mM dithiothreitol at 37 °C. The reaction is
incubated for 1 h in a total volume of 0.4 ml then terminated with an equal volume of 5%
trichloroacetic acid. 0.1 ml of the internal standard 3,4-dimethylbenzoic acid (10 µg/ml) is added.
Precipitated protein is separated by centrifugation at 12,000 × g for 10 min. The supernatant is
10 transferred to a clean tube and extracted twice with 0.4 ml of methylene chloride. The two extracts
are combined and dried under a stream of nitrogen. The residue is resuspended in 14% acetonitrile,
250 mM KH₂PO₄, pH 4.0, with 8 µl of diethylamine per 100 ml and) and injected onto a C18
reverse-phase HPLC column for separation. The column eluate was monitored at 235 nm. DME
activity is quantified by comparing peak area ratios of the analyte to the internal standard. A
15 standard curve was generated with benzoic acid standards prepared in a trichloroacetic acid-treated
protein matrix (Evgenia, V. et al. (1997) *J. Biol. Chem.* 272:14769-14775).

In another alternative, DME carboxyl esterase activity against the water-soluble substrate
para-nitrophenyl butyric acid is determined by spectrophotometric methods well known to those
skilled in the art. In this procedure, the DME-containing samples are diluted with 0.5 M Tris-HCl
20 (pH 7.4 or 8.0) or sodium acetate (pH 5.0) in the presence of 6 mM taurocholate. The assay is
initiated by adding a freshly prepared para-nitrophenyl butyric acid solution (100 µg/ml in sodium
acetate, pH 5.0). Carboxyl esterase activity was then monitored and compared with control
autohydrolysis of the substrate using an spectrophotometer set at 405 nm (Wan, L. et al. (2000) *J.*
Biol. Chem. 275:10041-10046).

25 Heparan sulfate 6-sulfotransferase activity of DME is measured *in vitro* by incubating a
sample containing DME along with 2.5 µmol imidazole HCl (pH 6.8), 3.75 µg of protamine chloride, 25
nmol (as hexosamine) of completely desulfated and N-resulfated heparin, and 50 pmol (about 5 × 10⁵
cpm) of [³⁵S] adenosine 3'-phosphate 5'-phosphosulfate (PAPS) in a final reaction volume of 50 µl at
37 °C for 20 min. The reaction is stopped by immersing the reaction tubes in a boiling water bath for
30 1 min. 0.1 µmol (as glucuronic acid) of chondroitin sulfate A is added to the reaction mixture as a
carrier. ³⁵S-labeled polysaccharides are precipitated with 3 volumes of cold ethanol containing 1.3%

WO 02/12467

PCT/US01/24382

potassium acetate and separated completely from unincorporated [³⁵S]PAPS and its degradation products by gel chromatography using desalting columns. One unit of enzyme activity is defined as the amount required to transfer 1 pmol of sulfate/min. as determined by the amount of [³⁵S]PAPS incorporated into the precipitated polysaccharides (Habuchi, H. et al. (1995) J. Biol. Chem.

5 270:4172-4179).

In the alternative, heparan sulfate 6-sulfotransferase activity of DME is measured by extraction and renaturation of enzyme from gels following separation by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Following separation, the gel is washed with 0.05 M Tris-HCl, pH 8.0, cut into 3-5 mm segments and subjected to agitation at 4 °C with 100 μl of 0.05 M Tris-HCl, pH 8.0 containing 0.15 M NaCl for 48 h. The eluted enzyme is collected by centrifugation and assayed for the sulfotransferase activity as above (Habuchi, H. et al. (1995) J. Biol. Chem. 270:4172-4179).

In another alternative, DME sulfotransferase activity is determined by measuring the transfer of [³⁵S]sulfate from [³⁵S]PAPS to an immobilized peptide. In one example, the peptide (QATEYBYLDYDFLPEC) represents the N-terminal 15 residues of the mature P-selectin glycoprotein ligand-1 polypeptide to which is added C-terminal cysteine residue. The peptide spans three potential tyrosine sulfation sites. The peptide is linked via the cysteine residue to iodoacetamide-activated resin at a density of 1.5-3.0 μmol peptide/ml of resin. The enzyme assay is performed by combining 10 μl of peptide-derivitized beads with 2-20 μl of DME-containing sample in 40 mM Pipes (pH 6.8), 0.3 M NaCl, 20 mM MnCl₂, 50 mM NaF, 1% Triton X-100, and 1 mM 5'-AMP in a final volume of 130 μl. The assay is initiated by addition of 0.5 μCi of [³⁵S]PAPS (1.7 μM; 1 Ci = 37 GBq). After 30 min at 37°C, the reaction beads are washed with 6 M guanidine at 65°C and the radioactivity incorporated into the beads is determined by liquid scintillation counting. Transfer of [³⁵S]sulfate to the bead-associated peptide is measured to determine the DME activity in the sample. One unit of activity is defined as 1 pmol of product formed per min (Ouyang, Y.-B. et al. (1998) Biochemistry 95:2896-2901).

In another alternative, DME sulfotransferase assays are performed using [³⁵S]PAPS as the sulfate donor in a final volume of 30 μl, contains 50 mM Hepes-NaOH (pH 7.0), 250 mM sucrose, 1 mM dithiothreitol, 14 μM [³⁵S]PAPS (15 Ci/mmol), and dopamine (25 μM), *p*-nitrophenol (5 μM), or other candidate substrates. Assay reactions are started by the addition of a purified DME enzyme preparation or a sample containing DME activity, allowed to proceed for 15 min at 37 °C, and terminated by heating at 100 °C for 3 min. The precipitates formed are cleared by

WO 02/12467

PCT/US01/24382

centrifugation. The supernatants are then subjected to the analysis of ³⁵S-sulfated product by either thin-layer chromatography or a two-dimensional thin layer separation procedure. Appropriate standards are run in parallel with the supernatants to allow the identification of the ³⁵S-sulfated products and determine the enzyme specificity of the DME-containing samples based on relative migration of reaction products (Sakakibara, Y. et al. (1998) J. Biol. Chem. 273:6242-6247).

Squalene epoxidase activity of DME is assayed in a mixture comprising purified DME (or a crude mixture comprising DME), 20 mM Tris-HCl (pH 7.5), 0.01 mM FAD, 0.2 unit of NADPH-cytochrome C (P-450) reductase, 0.01 mM [¹⁴C]squalene (dispersed with the aid of 20 μl of Tween 80), and 0.2% Triton X-100. 1 mM NADPH is added to initiate the reaction followed by incubation at 37 °C for 30 min. The nonsaponifiable lipids are analyzed by silica gel TLC developed with ethyl acetate/benzene (0.5:99.5, v/v). The reaction products are compared to those from a reaction mixture without DME. The presence of 2,3(S)-oxidosqualene is confirmed using appropriate lipid standards (Sakakibara, J. et al. (1995) 270:17-20).

Epoxide hydrolase activity of DME is determined by following substrate depletion using gas chromatographic (GC) analysis of ethereal extracts or by following substrate depletion and diol production by GC analysis of reaction mixtures quenched in acetone. A sample containing DME or an epoxide hydrolase control sample is incubated in 10 mM Tris-HCl (pH 8.0), 1 mM ethylenediaminetetraacetate (EDTA), and 5 mM epoxide substrate (e.g., ethylene oxide, styrene oxide, propylene oxide, isoprene monoxide, epichlorohydrin, epibromohydrin, epifluorohydrin, glycidol, 1,2-epoxybutane, 1,2-epoxyhexane, or 1,2-epoxyoctane). A portion of the sample is withdrawn from the reaction mixture at various time points, and added to 1 ml of ice-cold acetone containing an internal standard for GC analysis (e.g., 1-nonanol). Protein and salts are removed by centrifugation (15 min, 4000 × g) and the extract is analyzed by GC using a 0.2 mm × 25-m CP-Wax57-CB column (CHROMPACK, Middelburg, The Netherlands) and a flame-ionization detector. The identification of GC products is performed using appropriate standards and controls well known to those skilled in the art. 1 Unit of DME activity is defined as the amount of enzyme that catalyzes the production of 1 μmol of diol/min (Rink, R. et al. (1997) J. Biol. Chem. 272:14650-14657).

Aminotransferase activity of DME is assayed by incubating samples containing DME for 1 hour at 37 °C in the presence of 1 mM L-kynurenine and 1 mM 2-oxoglutarate in a final volume of 200 μl of 150 mM Tris acetate buffer (pH 8.0) containing 70 μM PLP. The formation of kynurenic acid is quantified by HPLC with spectrophotometric detection at 330 nm using the appropriate standards and controls well known to those skilled in the art. In the alternative,

WO 02/12467

PCT/US01/24382

L-3-hydroxykynurenine is used as substrate and the production of xanthurenic acid is determined by HPLC analysis of the products with UV detection at 340 nm. The production of kynurenic acid xanthurenic acid, respectively, is indicative of aminotransferase activity (Buchli, R. et al. (1995) *J. Biol. Chem.* 270:29330-29335).

5 In another alternative, aminotransferase activity of DME is measured by determining the activity of purified DME or crude samples containing DME toward various amino and oxo acid substrates under single turnover conditions by monitoring the changes in the UV/VIS absorption spectrum of the enzyme-bound cofactor, PLP. The reactions are performed at 25 °C in 50 mM 4-methylmorpholine (pH 7.5) containing 9 μM purified DME or DME containing samples and
10 substrate to be tested (amino and oxo acid substrates). The half-reaction from amino acid to oxo acid is followed by measuring the decrease in absorbance at 360 nm and the increase in absorbance at 330 nm due to the conversion of enzyme-bound PLP to PMP. The specificity and relative activity of DME is determined by the activity of the enzyme preparation against specific substrates (Vacca, R.A. et al. (1997) *J. Biol. Chem.* 272:21932-21937).

15 Superoxide dismutase activity of DME is assayed from cell pellets, culture supernatants, or purified protein preparations. Samples or lysates are resolved by electrophoresis on 15% non-denaturing polyacrylamide gels. The gels are incubated for 30 min in 2.5 mM nitro blue tetrazolium, followed by incubation for 20 min in 30 mM potassium phosphate, 30 mM TEMED, and 30 μM riboflavin (pH 7.8). Superoxide dismutase activity is visualized as white bands against a blue
20 background, following illumination of the gels on a lightbox. Quantitation of superoxide dismutase activity is performed by densitometric scanning of the activity gels using the appropriate superoxide dismutase positive and negative controls (e.g., various amounts of commercially available *E. coli* superoxide dismutase (Harth, G. and Horwitz, M.A. (1999) *J. Biol. Chem.* 274:4281-4292).

Catechol-*O*-methyltransferase activity of DME is measured in a reaction mixture consisting
25 of 50 mM Tris-HCl (pH 7.4), 1.2 mM MgCl₂, 200 μM SAM (*S*-adenosyl-L-methionine) iodide (containing 0.5 μCi of [methyl-³H]SAM), 1 mM dithiothreitol, and varying concentrations of catechol substrate (e.g., L-dopa, dopamine, or DBA) in a final volume of 1.0 ml. The reaction is initiated by the addition of 250-500 μg of purified DME or crude DME-containing sample and performed at 37 °C for 30 min. The reaction is arrested by rapidly cooling on ice and immediately
30 extracting with 7 ml of ice-cold n-heptane. Following centrifugation at 1000 x g for 10 min, 3-ml aliquots of the organic extracts are analyzed for radioactivity content by liquid scintillation counting. The level of catechol-associated radioactivity in the organic phase is proportional to the activity

WO 02/12467

PCT/US01/24382

catechol-*O*-methyltransferase activity of DME (Zhu, B.T. Liebr, J.G. (1996) 271:1357-1363).

DHFR activity of DME is determined spectrophotometrically at 15 °C by following the disappearance of NADPH at 340 nm ($\epsilon_{340} = 11,800 \text{ M}^{-1}\text{cm}^{-1}$). The standard assay mixture contains 100 μM NADPH, 14 mM 2-mercaptoethanol, MTEN buffer (50 mM 2-morpholinoethanesulfonic acid, 25 mM tris(hydroxymethyl)aminomethane, 25 mM ethanolamine, and 100 mM NaCl, pH 7.0), and DME in a final volume of 2.0 ml. The reaction is started by the addition of 50 μM dihydrofolate (as substrate). The oxidation of NADPH to NADP⁺ corresponds to the reduction of dihydrofolate in the reaction and is proportional to the amount of DHFR activity in the sample (Nakamura, T. and Iwakura, M. (1999) J. Biol. Chem. 274:19041-19047).

10 Sulfotransferase activity of DME is measured using the incorporation of ³⁵S from [³⁵S]PAPS into a model substrate such as phenol (Folds, A. and Meek, J. L. (1973) Biochim. Biophys. Acta 327:365-374). An aliquot of enzyme is incubated at 37°C with 1 mL of 10 mM phosphate buffer pH 6.4, 50 μM phenol, 0.4-4.0 μM [³⁵S]PAPS. After sufficient time for 5-20% of the radiolabel to be transferred to the substrate, 0.2 mL of 0.1 M barium acetate is added to precipitate protein and 15 phosphate buffer. Then 0.2 mL of 0.1 M Ba(OH)₂ is added, followed by 0.2 mL ZnSO₄. The supernatant is cleared by centrifugation, which removes proteins as well as unreacted [³⁵S]PAPS. Radioactivity in the supernatant is measured by scintillation. The enzyme activity is determined from the number of moles of radioactivity in the reaction product.

XVIII. Identification of DME Inhibitors

20 Compounds to be tested are arrayed in the wells of a multi-well plate in varying concentrations along with an appropriate buffer and substrate, as described in the assays in Example XVII. DME activity is measured for each well and the ability of each compound to inhibit DME activity can be determined, as well as the dose-response profiles. This assay could also be used to identify molecules which enhance DME activity.

25 Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. 30 Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Incyte Patent ID	Polypeptide SEQ ID NO:	Incyte Patent ID	Polymuc.ectide SEQ ID NO:	Incyte Patent ID	Polymuc.ectide SEQ ID NO:	Incyte Patent ID
7248285	1	7248285	20	7248285	20	7248285
7472835	2	7472835	21	7472835	21	7472835
7476203	3	7476203	22	7476203	22	7476203
7478583	4	7478583	23	7478583	23	7478583
7478585	5	7478585	24	7478585	24	7478585
7479904	6	7479904	25	7479904	25	7479904
7480367	7	7480367	26	7480367	26	7480367
7473869	8	7473869	27	7473869	27	7473869
7478588	9	7478588	28	7478588	28	7478588
55046125	10	55046125	29	55046125	29	55046125
3538709	11	3538709	30	3538709	30	3538709
7156310	12	7156310	31	7156310	31	7156310
7472027	13	7472027	32	7472027	32	7472027
1451928	14	1451928	33	1451928	33	1451928
3387823	15	3387823	34	3387823	34	3387823
55142051	16	55142051	35	55142051	35	55142051
7395274	17	7395274	36	7395274	36	7395274
	18		37		37	
	19		38		38	

WO 02/12467

PCT/US01/24382

Table 2

Polypeptide SEQ. ID. NO.	Enzyme Accession ID	GenBank Accession ID	Probability Range	GenBank Homolog
1	7478583CD1	g4929139	1.90E-248	[<i>Berca flavescens</i>] lysyl oxidase related protein homolog Langenau, D.M. et al. (1999) J. Mol. Endocrinol. 23:137-152.
2	7472835CD1	g559046	4.50E-181	[<i>Homo sapiens</i>] Flavin-containing monooxygenase O'Connor, H. et al. (1995) Arch. Biochem. Biophys. 317:275-284.
3	7476203CD1	g2653663	5.30E-91	[<i>Mus musculus</i>] cytochrome P450 monooxygenase
4	7478583CD1	g14475602	3.00E-72	[<i>Fl</i>] [<i>Bacillus cereus</i>] acetylacetoin reductase
4	7478583CD1	g529564	2.00E-61	[<i>Stenotrophomonas putida</i>] 2,3-butanediol dehydrogenase
5	7478585CD1	g14582559	1.00E-125	Ruang, M. et al. (1994) FEMS Microbiol. Lett. 124:141-150.
6	7479904CD1	g6683556	2.30E-255	[<i>Fl</i>] [<i>Staphylococcus aureus</i>] (AF282920) guanine oxidoreductase
7	7480367CD1	g243749	1.30E-101	[<i>Mus musculus</i>] heparan sulfate 6-O-sulfatase [<i>Homo sapiens</i>] glutathione S-transferase Hal subunit (BC2.5.1.18) Rosen, F. et al. (1992) Arch. Biochem. Biophys. 292:1589-1593
8	8069390CD1	g5531815	1.90E-56	[<i>Homo sapiens</i>] steroid dehydrogenase
8	8069390CD1	g13183088	1.00E-179	[<i>Fl</i>] [<i>Homo sapiens</i>] steroid dehydrogenase-like protein

WO 02/12467

PCT/US01/24382

Table 2 (cont.)

Polypeptide Seq. ID No.	Incyte Accession ID	GenBank Accession ID	Probability	GenBank Homology
18	55142951.0D1	U179355	9.50E-41	[Homo sapiens] catechol-O- methyltransferase Bertocci, B. et al. (1991) Human catechol-O-methyltransferase: cloning and expression of the cDNA sequence. Mol. Cell. Biol. 11:367-371. Walt. Acad. Sci. U.S.A. 86:1416- 1420.
19	7395274.0D1	55852342	4.60E-99	[Fundulus heteroclitus] cytochrome P450 2M1 Oleksiak, M.F. et al. (2000) Cloning, characterization and regulation of a new cytochrome P450 subfamily, the CYP2M8. J. Biol. Chem. 275:2312- 2321.

Table 3 (cont.)

SEQ ID	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Programs
6	747990CDL	439	S5, S59, S55, T65, T120, T129, S156, S187, S199, T205, S209, T220, Y277, S314, T355, S359, S464, S471, S475	S5, S59, S55, T65, T120, N296, N354, N438, N450, N458, N483, N486	HEPARANSULFATE 6-SULFOTRANSFERASE TRANSFERASE: PD139786; K207-E412, M1-K182 Signal peptide (signal_peptide): M1-Q24 Transmembrane domain (transmem_domain): I8-K25 Signal cleavage: M1-Q24	BLAST-PRODOM HMMER SFS-CAN
7	7480367CDL	222			GLUTAMICONE TRANSFERASE EM00127 P08263 71-112 GLUTAMICONE TRANSFERASE MULTIGENE FAMILY PROTEIN CLASS-ALPHA S-CRYSTALLIN GST LYASE: PD000312; L61-K204, K6-V102 (p = 8.0e-10) Glutathione S-transferase: PF00043; Q51-Q81 Glutathione S-transferase: E7-E92 Glutathione S-transferase family EM00034 E37058 43-272; I63-T297 Short-chain dehydrogenase: BL00061; G198-E235 Alcohol dehydrogenase: PR00080; F218-A237, K143-V154, G198-G206 Glucose/fructose dehydrogenase: E000051; K66-106, K143-V154, K152-C206, F210-A237, K239-T256 Short chain dehydrogenase (adh_short): R68-S254 Short-chain alcohol dehydrogenase family signature (adh_short.prf): G198-F248 Signal cleavage: M1-K36	BLAST-PRODOM BLAST-PRODOM HMMER-PFAM BLAST-PRODOM BLIMPS-BLOCKS BLIMPS-PRINTS BLIMPS-PRINTS HMMER-PFAM PROFILES-CAN SFS-CAN
8	8069390CDL	330	T33, S58, S98, T114, S178, S271, S286, T290, S320	N318		

Table 3 (cont.)

SEQ ID	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Programs
11	S5046125CD1	508	S23, S54, S99, T104, T114, T169, T178, S205, T246, T260, S269, T266, T324, S378, S473, S480, S485	N112, N107, N163, N447, N459, N467, N492, N495	HEPARANSULFATE 5-SULFOTRANSFERASE TRANSFERASE: PD19786:H50-H421 Transmembrane domain (transmem_domain): L57-Y74	BLAST-PRODOM HMMER
12	3538709CD1	439	T34, T83, T167, S188, T281, S423	N108, N332	Alpha/beta hydrolase fold: PRO0111: G202-S217, A262-L275, V362-E376 Phosphatase fold: PRO0412: H162-P200, G202-S217, L249-L251, I393-F415 Alpha/beta hydrolase fold (abhydrolase): Y203-L416 AMINO TRANSFERASES CLASS-III PYRIDOXAL-5-PHOSPHATE ATTACHMENT DRO0188 PI6922 2-426: D271-A304, K94-R233 Phosphatase fold: PRO0412: H162-P200, G202-S217, L249-L251, I393-F415 Precursor PK 2.5.1.44 SACT BETALAMINEPYRUVATE BETAALAAAT II TRANSFERASE PYRIDOXAL PHOSPHATE MITOCHONDRION TRANSIT PEPTIDE: PD032383: M1-S94 Amino transferases class-3: E00500: F303, G345, I377-L399, D84-P123 L305, I381-G394, F302-G330 Transmembrane domain: Y104-N124 Amino transferases class-III pyridoxal-phosphate: aminotran.3: W95-T208 I281-N487 Amino transferases class-III pyridoxal-phosphate attachment site (3c-transferases-class.3.pepf1): E308-C374	BLIMPS-PRIMFS BLIMPS-PRIMFS HMMER-PPAM BLAST-DMO BLAST-PRODOM
13	7156310CD1	514	T32, S169, S152, T248, S253, T383, T430, T480		AMINO TRANSFERASES CLASS-III PYRIDOXAL-5-PHOSPHATE ATTACHMENT DRO0188 PI6922 2-426: D271-A304, K94-R233 Phosphatase fold: PRO0412: H162-P200, G202-S217, L249-L251, I393-F415 Precursor PK 2.5.1.44 SACT BETALAMINEPYRUVATE BETAALAAAT II TRANSFERASE PYRIDOXAL PHOSPHATE MITOCHONDRION TRANSIT PEPTIDE: PD032383: M1-S94 Amino transferases class-3: E00500: F303, G345, I377-L399, D84-P123 L305, I381-G394, F302-G330 Transmembrane domain: Y104-N124 Amino transferases class-III pyridoxal-phosphate: aminotran.3: W95-T208 I281-N487 Amino transferases class-III pyridoxal-phosphate attachment site (3c-transferases-class.3.pepf1): E308-C374	BLIMPS-BLOCKS HMMER HMMER-PPAM PROFILESSCAN

Table 3 (cont.)

SEQ NO.	Incyte ID	Amino Residues	Potential Site	Potential Site	Potential Site	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
14	7472037CD1	226	S17, T23, T28, S106, S116, T195	M21, M65	M69	PHOSPHADENOSINE-PHOSPHOSULFATE SYNTHETASE (PAPS) BINDING SITE BINDING (DM00991 JC45311-287; M50-E19) TRANSFERASE SULFOTRANSFERASE STEROID METABOLISM KINON STEROID ACYLATION PROTEIN STEROID BINDING P001218; L73-V165	BLAST-DOMO BLAST-PRODOM
15	7480358CD1	121	T38, S96, T119	M69	M69	COPPER/ZINC SUPEROXIDE DISMUTASE (DM00227 F0044213-147; K3-V116) SUPEROXIDE DISMUTASE CUZN OXIDOREDUCTASE COPPER ZINC PEROXIDASE SIGNAL PEPTIDASE (DM00087; K3-022, R40-W63) Cu-Zn-superoxide dismutase proteclins: P55, V97-T119	BLAST-DOMO BLAST-PRODOM
16	1618256CD1	486	T9, S55, T74, T45, S296, T349, T356, T420, S431, S444, T480	M374	M374	Copper/zinc superoxide dismutase (SODC) (scodcu); M2-P73, S96-V116 SIGNAL PEPTIDE: M1-G36 CEREBROSIDE SULFOTRANSFERASE: PD140372; M13-G215, P279-Y461	HMMER-PPM SIGREPT SUSCAN BLAST-PRODOM

WO 02/12467

PCT/US01/24382

Table 3 (cont.)

SEQ ID NO.	Enzyme ID	Amino Residues	Potential Phosphorylation Sites	Potential Ubiquitination Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
17	3387823CD1	649	T64, T184, T236, S253, S368, S382, S441, T438, S456, S516, S547, T550, T563, T624	M309	PROTEIN SIMILARITY KINASE SLIME MOLD RL3610.2 CHROMOSOME III OXIDOREDUCTASE FLAVOPROTEIN: PD041935: I358-D433, D250-G137 Signal peptides: M1-E72 Amino oxidase: P219-S467 Adrenodoxin reductase family (flavin-containing component of mitochondrial monooxygenase system): PR00419A: S211-I233, PR00419E: R234-R247, PR00419D: V212-R226 monooxygenase (cytochrome P-450) PR00420C: OG14-P629 Transmembrane domain: M1-V18 C-MET/PTP/TRANSFERASE TRANSFERASE ME/PY/PTP/TRANSFERASE: PD173699: I65-D162 (E9-S8-ID) C-MET/PTP/TRANSFERASE CHAIN: DM02481 P22734141-138: E46-S137	BLAST-PRODOM SIGPEPT HMMER-PPM ELIMES-PRINTS ELIMES-PRINTS HMMER BLAST-PRODOM BLAST-DOMO
18	55142051CD1	258	S35, T67, S73, S74, S159, S174			

Table 3 (cont.)

SEQ ID NO	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Categories
13	739574C1	544	S35, S88, S135, S148, S161, T184, Y205, T244, S335, S336, S393, T399, S416, T479, S519, T525, S542	S135, N333, N334, N338	Signature Sequences, Domains and Motifs L110-F328 CYTOCHROME P450 MONOOXYGENASE OXIDOREDUCTASE HEME ELECTRON TRANSPORT MEMBRANE MICROSOME ENDOPLASMIC R297-S432, S484-Q471, S490-I539, I436-I539, S433, S434, S435 Cytochrome P450 cysteine BL00086: F480-F511 Mitochondrial P450 sigma PR00408: I370-E381, A398-V416, F462-D470, I481-C490, C490-F501, M50-P65, H169-K179, A352-S369 P450 sigma PR00411: I370-E381, A398-V416, F462-D470, I481-C490, C490-F501, M50-P65, H169-K179, A352-S369 S135-T152, A279-R245, Y341-Q348, L361-C387, S404-P422, M445-P469, F480-C490, C490-F513 E-class P450 group IV PR00465: V105-V128, H450-L468, L474-C490, C490-R508 Cytochrome P450: L107-R543 Cytochrome P450: F381-S492 Cytochrome P450 sigma PR00411: I370-E381, A398-V416, F462-D470, I481-C490, C490-F501, M50-P65, H169-K179, A352-S369 Signal_cleavage: M1-G47	BLAST-PRODOM BLIMPS-BLOCKS BLIMPS-PRINTS BLIMPS-PRINTS BLIMPS-PRINTS HMMER-PPAE MOTIFS PROFILSCAN SPSCAN

Table 4

Polynucleotide ID	Incyte polynucleotide ID	Sequence length	Selected sequence length (s)	Sequence Fragments	5' Position	3' Position
20	724828CB1	2893	526-572	7590384AT (PROSTRAC6) 6460333BT (GSTRAC01) 8035282HH (SACRAC01) 2796172F6 (INPOLAC01) 7590384HH (PROSTRAC6) PL7248285_g7939436_000006 G1850108_L7-8 G1850108_L7-8 G1850108_L7-8	652 1144 276 2072 1650 1106 1428	1320 1999 832 2303 2185 1428
21	7472835CB1	1745	1-57, 412-1745	G52-g7107731_0001.edt.2 comp	1 1445	456 1745
22	7476203CB1	1587	1-209, 380-1587	G3L-g7107731_0001.edt.1 G3L-g8567322_000003_00000 4.edt 4.edt.2 G3L-g8567322_000003_00000 4.edt.2	1 1 960	1674 1139 1587
23	7478583CB1	1038	1-1038	GNN-g8576427_000068_002	1	1038
24	7478585CB1	1584	1-1584	GNN-g8576427_CDS_1.edt GNN_7478585CTL	1 596	1467 1584
25	7479904CB1	1500	1-149	G1655588_g319191972_g856913 3.edt	1	1500
26	7480367CB1	669		G5332848/076.edt.1 2020035HH (PERAC034)	1 508	669 669

Table 4 (cont.)

Polynucleotide SEQ. ID NO.	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
27	8069390CB1	3551	1609-1783, 201-3101-3551 241, 3101-3551	7360250HL [BRA1FER05]	607	1056
				74605974H [OVAR1NG02]	2157	575
				6460195HL [BRA1NOC01]	330	989
				17594708G [PIT1NOC03]	2966	3512
				1619129TG [BRA1TUT12]	2859	3484
				6249653R6 [BRA1D1R01]	2357	2781
				91203094	3095	3551
				1576067G [LIND1NOC03]	2736	3390
				4365248HL [OVAR1NOC01]	2920	3180
				6249653R6 [BRA1D1R01]	2357	2781
				7100733HL [BRA1D1R02]	1597	2148
				3333887F6 [BRA1FER01]	1068	1673
28	7473869CB1	2178	1-231, 1549- 1579, 1949- 2178, 654-732	5502367L1	1360	2178
				5505144L1	1	1032
29	7478368CB1	2081	1-97, 1256- 1459	55071187J1	1045	1662
				55071187J1	997	1606
				5507278F1	436	1709
				6025171D3	1333	1837
				55058627J1	157	714
				6205079HL [PIT1NOC01]	1562	2081
				43653190HL [SK1NOC01]	1129	1391
				55111596H1	1	540
				55071187J1	792	1365
				21823248T1	1694	2245
				55071237J1	2124	2482
				55046165J2	720	1572
30	55046125CB1	2642	2250-2642, 485-466	55046125H1	314	1143
				56246385H1 [E1ACFER01]	1512	2108

Table 4 (cont.)

SEQ ID NO:	Polynucleotide Invcye Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
31	3538709CB1	2080	863-1020, 2027-2090, 380-635	707524HL (BRANDR01)	624	1173
				225514E5 (SENGT01)	177	390
				398374HL (BRANDR01)	174	390
				7784424HL (BRANDR01)	1166	1461
				68834350J (BRANDR01)	1	453
32	71563101CB1	2219	1-44, 541- 1063	7191721H2 (BRANDR01)	1315	1932
				2560864H1 (ADRETT01)	253	855
				62895189	1656	1947
				7156352AV1	1734	2212
				5626568AV1	684	1306
				5974568AV1	1	184
				293718H1 (LIVRNR04)	1828	2066
				71563101V1	529	1354
				FL7472027CT1_00001	1	681
				GNL_99211248_000026_002	1	399
33	7472027CB1	681	148-681			
34	7480358CB1	399	1-35			
35	1618256CB1	2302	2241-2302, 1- 260, 1172- 1194	7010510H1 (CCJNFEC01)	1	604
				2618258F5 (BRATTU02)	1829	1866
				3407444H3 (STNTNR01)	3475	2112
				4922615F9 (WESTNR01)	646	1374
				2605131F6 (LUNGTTU07)	1835	2302

Table 4 (cont.)

Polynucleotide SEQ ID NO.	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
36	3387823CB1	3341	1-447, 8270-1445	774129SH (THYRGR01)	1	563
				724129SH	1819	341
				7085632TVI	1919	2849
				70857153VI	2459	2872
				70856101VI	1381	1902
				657765HI (ROSINOR03)	2322	2590
				70002313DI	1551	2004
				995959HI (KIDNFTF01)	2584	2902
				6741554JI (HBAERGR01)	768	1428
				7087665HI	436	1024
37	55142051CB1	777	1-257	55142051L1	72	337
				GEI_48567879_e41t	1	777
38	7395274CB1	3600	1-1330, 2414-3076, 2342-2372, 3576-3600	6779264HI (OVARDR01)	2278	2867
				62064300	3001	3596
				55052831HI	787	1597
				673714R8 (B64BDL02)	454	531
				673714R9 (B64BDL02)	454	531
				2300407R6 (BBSNOC05)	3046	2645
				2300407R6 (BBSNOC05)	3046	3568
				7996595HI (ADPRTTC01)	1493	2221
				62079829	3243	3600
				7456912HI (LLVFTUE01)	1677	2855
				1259632T6 (KIDNFTF03)	2869	3567
				6544665HI (LARDONOG2)	3124	3590

Table 5

Polymer Peptide SEQ. ID. NO.	Incubate Primer ID	Representative Library
26	7448285CB1	PROSTM06
26	7480367CB1	SINSTM02
27	8069390CB1	HN12A001
28	7473869CB1	BRAOTDR02
29	7478588CB1	OVARL01
30	5004612CB1	BRSSTM013
31	3538709CB1	BRAHDR03
32	71563101CB1	SINSTM01
35	1618256CB1	BRA1WT12
36	3387823CB1	BRA1WT19
38	7395274CB1	BRA07E02

Table 6

Library	Vector	Library Description
BRAD10E02	pUNCY	This 5' biased random primed library was constructed using RNA isolated from diseased cerebellar tissue from a 55-year-old Caucasian male who died from a cerebrovascular accident. Serologies were negative. Patient history included Huntington's disease, emphysema, and tobacco abuse (3-4 packs per day, for 40 years).
BRAD10R03	pCDNA2.1	This random primed library was constructed using RNA isolated from archaocortex anterior hippocampus tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningeal fibrosis predominantly over the convoluted sulcus, and acute spheroidal fibrils with matrix over 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. Serologies were negative.
BRAD10T19	pUNCY	Library constructed using RNA isolated from diseased and healthy tissue from the left frontal lobe of a 27-year-old Caucasian male during a brain lobectomy. Pathology indicated a focal deep white matter lesion, characterized by marked gliosis, calcifications, and hemosiderin-laden macrophages, consistent with a remote perinatal injury. This tissue also showed mild to moderate generalized gliosis, predominantly subpial and subcortical, consistent with chronic seizure disorder. The left temporal lobe showed moderate to severe gliosis, consistent with mesial temporal sclerosis. Loss and gliosis in hippocampal sector CA1 consistent with intractable epilepsy. GFAP was positive for astrocytes. The patient presented with intractable epilepsy, focal epilepsy, hemiplegia, and an unspecified brain injury. Patient history included cerebral palsy, abnormality of gait, and depressive disorder. Family history included brain cancer.

WO 02/12467

PCT/US01/24382

Table 6 (cont.)

Library	Vector	Library Description
BRALTWT12	P1N1CY	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe of a 40-year-old Caucasian female during excision of a cerebral meningioma lesion. Pathology indicated grade 4 gemistocytic astrocytoma.
BRAGTTR02	PCDMA2.1	This random primed library was constructed using RNA isolated from mixed midbrain from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated meningeal fibrosis predominately 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 mixed differentiated glioblastoma. Associated pathology included malnutrition, post-operative Budd-Chiari syndrome, biliary ascites, cholelithiasis, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 8% of the liver.
BRSTMT13	P1N1CY	Library was constructed using RNA isolated from breast tissue removed from a 36-year-old Caucasian female during bilateral simple mastectomy. Patient history included a breast cancer diagnosis, bilateral mastectomy, and a history of breast cancer. Patient history included a cardiovascular and cerebrovascular disease, hyperlipidemia, skin, breast, esophageal, bladder, and bone cancer, and Hodgkin's lymphoma.
HNPTAST01	PELUSCRIP1	Library was constructed at Stratagene (STR037233), using RNA isolated from the HNT2 cell line derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor. Cells were treated with retinoic acid for 5 weeks and with 1,25(OH) ₂ D ₃ for two weeks and allowed to mature for an additional 4 weeks in conditioned medium.
OVARDIR01	PCDMA2.1	This random primed library was constructed using RNA isolated from right ovary tissue removed from a 45-year-old Caucasian female during total abdominal hysterectomy, bilateral salpingo-oophorectomy, vaginal suspension and fixation, and incidental appendectomy. Pathology indicated atypical hyperthecosis of the right and left ovaries. Pathology indicated atypical hyperthecosis of the right and left ovaries (teratoma) in the left ovary. Multiple (3) intramural leiomyomata were identified in the uterine wall. Patient history included metrorrhagia, female stress incontinence, alopecia, depressive disorder, pneumonia, normal delivery, and deficiency anemia. Family history included benign hypertension, atherosclerotic coronary artery disease, hyperlipidemia, and primary tuberculous complex.

Table 6 (cont.)

Library	Vector	Library Description
PROSTW06	FCDMA2.1	This 5' biased random primed library was constructed using RNA isolated from diseased prostate tissue removed from a 57-year-old Caucasian male during closed prostatic biopsy, radical prostatectomy, and regional lymph node excision. Pathology indicated adenocarcinoma of the prostate gland. The specimen was snap frozen and stored at -80°C centrally. The patient presented with elevated prostate specific antigen and prostate cancer. Patient history included tobacco abuse in remission. Previous surgeries included cholecystectomy, repair of diaphragm hernia, and repair of vertebral fracture. Patient medications included Pepsid, Omipren, and Bulexin. Family history included benign prostatic hyperplasia, prostate cancer in the mother, prostate cancer in the father, drug abuse, cancer, and type II diabetes in the mother, prostate cancer in the father, drug abuse, prostate cancer, and breast cancer in the sibling(s).
SINTW01	FCDMA2.1	This random primed library was constructed using RNA isolated from ileum tissue removed from a 70-year-old Caucasian female during right hemicolectomy, open liver biopsy, flexible sigmoidoscopy, colonoscopy, and permanent colostomy. Pathology for the matched ileum tissue indicated invasive grade adenocarcinoma forming a conglomerate mass, breast neoplasm, type II diabetes, hypelipidemia, viral hepatitis, an unspecified thyroid disorder, osteoarthritis, a malignant skin neoplasm, deficiency anemia, and normal delivery. Family history included breast cancer, atherosclerotic coronary artery disease, benign hypertension, cerebrovascular disease, ovarian cancer, and hyperlipidemia.
SINTW02	PELUSCRIPT	This library was constructed using RNA isolated from the small intestines of a 55-year-old Caucasian female, who died from a subarachnoid hemorrhage. Serologies were positive for Cytomegalovirus (CMV). Previous surgeries included a hysterectomy.

WO 02/12467

PCT/US01/24382

Table 7

Program	Description	Reference	Parameter Threshold
ABIFACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABJ/ARACEL.FDF	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 <5%
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) <i>J. Mol. Biol.</i> 215:403-410; Altschul, S.F. et al. (1997) <i>Nucleic Acids Res.</i> 25:3389-3402.	ESTs: Probability value= 1.0E-3 or less Full Length sequence: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, ffasta, fastx, fasty, and search.	Pearson, W.R. and D.J. Lipman (1988) <i>Proc. Natl. Acad. Sci. USA</i> 85:2444-2448; Pearson, W.R. (1990) <i>Methods Enzymol.</i> 183:63-98; and Smith, T.F. and M.S. Waterman (1981) <i>Adv. Appl. Math.</i> 2:482-489.	ESTs: fasta E value= 1.0E-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 sequence: fasta score=100 or greater
BUMPS	A Blocks Improved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) <i>Nucleic Acids Res.</i> 19:6665-6672; Henikoff, J.G. and S. Henikoff (1996) <i>Methods Enzymol.</i> 266:88-105; and Atwood, T.K. et al. (1997) <i>J. Chem. Inf. Comput. Sci.</i> 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) <i>J. Mol. Biol.</i> 235:1401-1531; Sonnhammer, E.L.J. et al. (1988) <i>Nucleic Acids Res.</i> 16:220-222; Durbin, R. et al. (1988) <i>Oxford World View</i> ; in a Nutshell, Cambridge Univ. Press, pp. 1-250.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater

WO 02/12467

PCT/US01/24382

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1987) Nucleic Acids Res. 25:217-221.	Normalized quality scores CCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	Score=120 or greater; Match length=56 or greater
Phrap	A Phis Revised Assembly Program including SWAT and Crossmatch programs based on efficient implementation of the Smith-Waterman algorithm, used in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:462-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	
Conseq	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nelson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Scores=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth 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 02/12467

PCT/US01/24382

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
 - a) a polypeptide comprising an amino acid sequence selected from the group consisting of
5 SEQ ID NO:1-19 ,
 - 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:1-19 ,
 - c) a biologically active fragment of a polypeptide having an amino acid sequence selected
from the group consisting of SEQ ID NO:1-19 , and
 - 10 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from
the group consisting of SEQ ID NO:1-19 .
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-19 .
- 15 3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID
20 NO:20-38.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a
polynucleotide of claim 3.
- 25 7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
9. A method of producing a polypeptide of claim 1, the method comprising:
 - 30 a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell
is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a
promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and
 - b) recovering the polypeptide so expressed.

WO 02/12467

PCT/US01/24382

10. An isolated antibody which specifically binds to a polypeptide of claim 1.
11. An isolated polynucleotide selected from the group consisting of:
- a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:20-38,
 - 5 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:20-38,
 - c) a polynucleotide complementary to a polynucleotide of a),
 - d) a polynucleotide complementary to a polynucleotide of b), and
 - 10 e) an RNA equivalent of a)-d).
12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.
13. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex
 - 15 is formed between said probe and said target polynucleotide or fragments thereof, and
 - b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.
14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.
15. A method of detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:
- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
 - 20 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.
16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable

WO 02/12467

PCT/US01/24382

excipient.

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 .

5

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

10

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

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

15

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

20

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

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

25

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

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

30

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

25. A method of screening for a compound that specifically binds to the polypeptide of claim

WO 02/12467

PCT/US01/24382

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

5 compound that specifically binds to the polypeptide of claim 1.

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

- a) combining the polypeptide of claim 1 with at least one test compound under conditions

10 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

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

27. A method 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

20 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

25 of the compound and in the absence of the compound.

28. 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

30 least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof,

WO 02/12467

PCT/US01/24382

- 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.
- 5
29. A diagnostic test for a condition or disease associated with the expression of DME in a biological sample, the method comprising:
- a) combining the biological sample with an antibody of claim 10, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex, and
10 b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.
30. The antibody of claim 10, wherein the antibody is:
- a) a chimeric antibody,
15 b) a single chain antibody,
c) a Fab fragment,
d) a F(ab')₂ fragment, or
e) a humanized antibody.
- 20 31. A composition comprising an antibody of claim 10 and an acceptable excipient.
32. A method of diagnosing a condition or disease associated with the expression of DME in a subject, comprising administering to said subject an effective amount of the composition of claim 31.
- 25 33. A composition of claim 31, wherein the antibody is labeled.
34. A method of diagnosing a condition or disease associated with the expression of DME in a subject, comprising administering to said subject an effective amount of the composition of claim 33.
- 30 35. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 10, the method comprising:
- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19, or an immunogenic fragment thereof, under conditions to elicit

WO 02/12467

PCT/US01/24382

- an antibody response,
- b) isolating antibodies from said animal, and
 - c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 .
- 5
36. An antibody produced by a method of claim 35.
37. A composition comprising the antibody of claim 36 and a suitable carrier.
- 10
38. A method of making a monoclonal antibody with the specificity of the antibody of claim 10, the method comprising:
- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 , or an immunogenic fragment thereof, under conditions to elicit an antibody response,
 - 15 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
 - 20 e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 .
39. A monoclonal antibody produced by a method of claim 38.
- 25
40. A composition comprising the antibody of claim 39 and a suitable carrier.
41. The antibody of claim 10, wherein the antibody is produced by screening a Fab expression library.
- 30
42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant immunoglobulin library.
43. A method of detecting a polypeptide having an amino acid sequence selected from the

WO 02/12467

PCT/US01/24382

group consisting of SEQ ID NO:1-19 in a sample, the method comprising:

- a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide, and
 - b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide
- 5 having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 in the sample.

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19 from a sample, the method comprising:

- a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding
- 10 of the antibody and the polypeptide, and
- b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-19.

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

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

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

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

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

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

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

52. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

30 53. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

54. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

WO 02/12467

PCT/US01/24382

55. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.
56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.
- 5 57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.
58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.
59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.
- 10 60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.
61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.
- 15 62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.
63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.
64. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:20.
- 20 65. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:21.
66. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:22.
- 25 67. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:23.
68. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:24.
69. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:25.
- 30 70. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:26.
71. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:27.

WO 02/12467

PCT/US01/24382

72. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:28.
73. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:29.
- 5 74. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:30.
75. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:31.
76. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:32.
- 10 77. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:33.
78. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:34.
- 15 79. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:35.
80. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:36.
81. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:37.
- 20 82. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:38.

WO 02/12467

PCT/US01/24382

<110> INCYTE GENOMICS, INC.
 BAUGHN, Mariah R.
 BRUNS, Christopher M.
 DAS, Debopriya
 DELEGEMME, Angelo M.
 DINI, Li
 ELLIOTT, Vicki S.
 GANDHI, Ameena R.
 GRIPFIN, Jennifer A.
 HAFALIA, April J.A.
 KHAN, Farrah A.
 LAL, Preeti
 LEE, Sally
 LU, Dying Aina M.
 LU, Yan
 PATTERSON, Chandra
 RAMKUMAR, Jayalaxmi
 RING, Huijun Z.
 SANJANWALA, Madhu S.
 TANG, Y. Tom
 THANGAVELU, Kavitha
 THORNTON, Michael
 TRIBOULEY, Catherine M.
 WALIA, Harinder K.
 WARREN, Bridget A.
 YANG, Junning
 YAO, Monique G.
 YUE, Henry

<120> DRUG METABOLIZING ENZYMES

<130> FI-0185 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/223,055; 60/224,728; 60/226,440; 60/228,057; 60/230,063; 60/232,244;
 60/234,269

<151> 2000-08-04; 2000-08-11; 2000-08-18; 2000-08-24; 2000-08-31; 2000-09-13;
 2000-09-20

<160> 38

<170> PERL Program

<210> 1

<211> 756

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7248285CD1

<400> 1
 Met Ala Trp Ser Pro Ala Thr Leu Phe Leu Leu Leu
 1 5 10 15
 Leu Gly Gln Pro Pro Ser Arg Pro Gln Ser Leu Gly Thr Thr
 20 25 30
 Lys Leu Arg Leu Val Gly Pro Glu Ser Lys Pro Glu Glu Gly Arg
 35 40 45
 Leu Glu Val Leu His Gln Gly Gln Trp Gly Thr Val Cys Asp Asp
 50 55 60

WO 02/12467

PCT/US01/24382

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

WO 02/12467

PCT/US01/24382

Val Ser Cys Met Asp Ser Ala Pro Asp Leu Val Met Asn Ala Gln
 530 535 540
 Leu Val Gln Glu Thr Ala Tyr Leu Glu Asp Arg Pro Leu Ser Gln
 545 550 555
 Leu Tyr Cys Ala His Glu Glu Asn Cys Leu Ser Lys Ser Ala Asp
 560 565 570
 His Met Asp Trp Pro Tyr Gly Tyr Arg Arg Leu Leu Arg Phe Ser
 575 580 585
 Thr Gln Ile Tyr Asn Leu Gly Arg Thr Asp Phe Arg Pro Lys Thr
 590 595 600
 Gly Arg Asp Ser Trp Val Trp His Gln Cys His Arg His Tyr His
 605 610 615
 Ser Ile Glu Val Phe Thr His Tyr Asp Leu Leu Thr Leu Asn Gly
 620 625 630
 Ser Lys Val Ala Glu Gly His Lys Ala Ser Phe Cys Leu Glu Asp
 635 640 645
 Thr Asn Cys Pro Thr Gly Leu Gln Arg Arg Tyr Ala Cys Ala Asn
 650 655 660
 Phe Gly Glu Gln Gly Val Thr Val Gly Cys Trp Asp Thr Tyr Arg
 665 670 675
 His Asp Ile Asp Cys Gln Trp Val Asp Ile Thr Asp Val Gly Pro
 680 685 690
 Gly Asn Tyr Ile Phe Gln Val Ile Val Asn Pro His Tyr Glu Val
 695 700 705
 Ala Glu Ser Asp Phe Ser Asn Asn Met Leu Gln Cys Arg Cys Lys
 710 715 720
 Tyr Asp Gly His Arg Val Trp Leu His Asn Cys His Thr Gly Asn
 725 730 735
 Ser Tyr Pro Ala Asn Ala Glu Leu Ser Leu Glu Gln Glu Gln Arg
 740 745 750
 Leu Arg Asn Asn Leu Ile
 755

<210> 2

<211> 544

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7472835CD1

<400> 2

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

WO 02/12467

PCT/US01/24382

```

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

```

<210> 3
<211> 501
<212> PRT
<213> Homo sapiens
<220>

WO 02/12467

PCT/US01/24382

```

<221> misc_feature
<223> Incyte ID No: 7476203CD1

<400> 3
Met Leu Ser Leu Leu Ser Gly Leu Ala Leu Leu Ala Ile Ser Phe
 1      5      10      15
Leu Leu Leu Lys Leu Gly Thr Phe Cys Trp Asp Arg Ser Cys Leu
 20     25     30
Pro Pro Gly Pro Leu Pro Phe Pro Ile Leu Gly Asn Leu Trp Gln
 35     40     45
Leu Cys Phe Gln Gln Pro His Leu Ser Leu Lys Asn Phe Gln Lys
 50     55     60
Lys Ile Gly Asn Ile Phe Met Asn Leu Gly Ser Ser Val Val Pro
 65     70     75
Leu Ala Leu Pro Leu Leu Pro Val Thr Phe His Pro Leu Asn Gln
 80     85     90
Gly Val Leu Cys Lys Pro Leu Ile Thr Phe Pro Lys Pro Phe Pro
 95     100    105
Thr Arg Asn Pro Gly Ile Ile Cys Ser Ser Gly His Thr Trp Arg
 110    115    120
Gln Lys Arg Arg Phe Cys Leu Val Met Ile Arg Gly Leu Gly Leu
 125    130    135
Gly Lys Leu Ala Leu Glu Val Gln Leu Gln Lys Glu Ala Ala Glu
 140    145    150
Leu Ala Glu Ala Phe Arg Gln Glu Gln Gly Lys Arg Pro Phe Asp
 155    160    165
Pro Gln Val Ser Ile Val Arg Ser Thr Val Arg Val Ile Gly Ala
 170    175    180
Leu Val Phe Gly His His Phe Leu Leu Glu Asp Pro Ile Phe Gln
 185    190    195
Glu Leu Thr Gln Ala Ile Asp Phe Gly Leu Ala Phe Val Ser Thr
 200    205    210
Val Trp Arg Gln Leu Tyr Asp Val Phe Pro Trp Ala Leu Cys His
 215    220    225
Leu Pro Gly Pro His Gln Glu Ile Phe Arg Tyr Gln Glu Val Val
 230    235    240
Leu Ser Leu Ile His Gln Glu Ile Thr Arg His Lys Leu Arg Ala
 245    250    255
Pro Glu Ala Pro Arg Asp Phe Ile Ser Cys Tyr Leu Ala Gln Ile
 260    265    270
Ser Lys Ala Met Asp Asp Pro Val Ser Thr Phe Asn Gln Glu Asn
 275    280    285
Leu Val Gln Val Val Ile Asp Leu Phe Leu Gly Gly Thr Asp Thr
 290    295    300
Thr Ala Thr Thr Leu Cys Trp Ala Leu Ile His Met Ile Gln His
 305    310    315
Gly Ala Val Gln Gln Thr Val Gln Leu Glu Leu Asp Glu Val Leu
 320    325    330
Gly Ala Ala Pro Val Val Cys Tyr Glu Asp Arg Lys Arg Leu Pro
 335    340    345
Tyr Thr Met Ala Val Leu His Asp Val Gln Arg Leu Ser Ser Val
 350    355    360
Met Ala Met Gly Ala Val Arg Gln Cys Val Thr Ser Thr Arg Val
 365    370    375
Cys Ser Tyr Pro Val Ser Lys Gly Thr Ile Ile Leu Pro Asn Leu
 380    385    390
Ala Ser Val Leu Tyr Asp Pro Glu Cys Trp Glu Thr Pro Arg Gln
 395    400    405
Phe Asn Pro Gly His Phe Ser Asp Lys Asp Gly Asn Phe Val Ala
 410    415    420
Asn Glu Ala Phe Leu Pro Phe Ser Ala Gly Thr Arg Val Tyr Pro
 425    430    435

```

WO 02/12467

PCT/US01/24382

Ala Asp Gln Leu Ala Gln Met Glu Leu Phe Leu Met Phe Ala Thr
 440 445 450
 Leu Leu Arg Thr Phe Arg Phe Gln Leu Pro Glu Gly Ser Pro Gly
 455 460 465
 Leu Lys Leu Glu Tyr Ile Phe Gly Gly Thr Trp Gln Pro Gln Pro
 470 475 480
 Gln Glu Ile Cys Ala Val Pro Arg Leu Ser Ser Pro Ser Pro Gly
 485 490 495
 Pro Arg Glu Asp Gly Leu
 500

<210> 4
 <211> 345
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7478583CD1

<400> 4
 Met Lys Ala Ala Val Trp Tyr Gly Gln Lys Asp Val Arg Val Glu
 1 5 10 15
 Glu Arg Glu Pro Lys Glu Leu Gln Asp Asn Glu Val Lys Val Lys
 20 25 30
 Val Ser Trp Ala Gly Ile Cys Gly Thr Asp Leu His Glu Tyr Leu
 35 40 45
 Glu Gly Pro Ile Phe Ile Ser Thr Glu Lys Pro Asp Pro Phe Leu
 50 55 60
 Gly Gln Lys Ala Pro Val Thr Leu Gly His Glu Phe Ala Gly Val
 65 70 75
 Val Glu Glu Thr Gly Ser Gln Val Thr Lys Phe Asn Lys Gly Asp
 80 85 90
 Arg Val Val Val Asn Pro Thr Val Ser Lys Arg Glu Lys Glu Glu
 95 100 105
 Asn Ile Asp Leu Tyr Asp Gly Tyr Ser Phe Ile Gly Leu Gly Ser
 110 115 120
 Asp Gly Gly Phe Ala Glu Phe Thr Asn Ala Pro Glu Glu Asn Val
 125 130 135
 Tyr Lys Leu Pro Asp Asn Val Ser Asp Lys Glu Gly Ala Leu Val
 140 145 150
 Glu Pro Thr Ala Val Ala Val Gln Ala Ile Lys Glu Gly Glu Val
 155 160 165
 Leu Phe Gly Asp Thr Val Ala Ile Phe Gly Ala Gly Pro Ile Gly
 170 175 180
 Leu Leu Thr Val Val Ala Ala Lys Ala Ala Gly Ala Ser Lys Ile
 185 190 195
 Phe Val Phe Asp Leu Ser Glu Glu Arg Leu Ser Lys Ala Lys Ala
 200 205 210
 Leu Gly Ala Thr His Ala Ile Asn Ser Gly Lys Thr Asp Pro Val
 215 220 225
 Asp Val Ile Asn Glu Tyr Thr Glu Asn Gly Val Asp Val Ser Phe
 230 235 240
 Glu Val Ala Gly Val Ala Pro Thr Leu Lys Ser Ser Ile Asp Val
 245 250 255
 Thr Lys Ala Arg Gly Thr Val Val Ile Val Ser Ile Phe Gly His
 260 265 270
 Pro Ile Glu Trp Asn Pro Met Gln Leu Thr Asn Thr Gly Val Lys
 275 280 285
 Leu Thr Ser Thr Ile Ala Tyr Thr Pro Thr Thr Phe Gln Gln Thr
 290 295 300
 Ile Asp Leu Ile Asn Glu Gly Asn Leu Asn Val Lys Asp Val Val

WO 02/12467

PCT/US01/24382

```

          305          310          315
Thr Asp Glu Ile Glu Leu Glu Asn Ile Val Glu Ser Gly Phe Glu
          320          325          330
Gln Leu Val Asn Asp Lys Ser Gln Ala Lys Ile Leu Ile Lys Leu
          335          340          345

```

```

<210> 5
<211> 361
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7478585CD1

```

```

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

```

WO 02/12467

PCT/US01/24382

	335		340		345
Phe Thr Gly Asp	Asn Phe Gly Lys Gln	Val Val Lys Ile Lys	Glu		
	350		355		360
Glu					

<210> 6
 <211> 499
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7479904CD1

<400> 6
 Met Asp Glu Lys Ser Asn Lys Leu Leu Leu Ala Leu Val Met Leu
 1 5 10 15
 Phe Leu Phe Ala Val Ile Val Leu Gln Tyr Val Cys Pro Gly Thr
 20 25 30
 Glu Cys Gln Leu Leu Arg Leu Gln Ala Phe Ser Ser Pro Val Pro
 35 40 45
 Asp Pro Tyr Arg Ser Glu Asp Glu Ser Ser Ala Arg Phe Val Pro
 50 55 60
 Arg Tyr Asn Phe Thr Arg Gly Asp Leu Leu Arg Lys Val Asp Phe
 65 70 75
 Asp Ile Lys Gly Asp Asp Leu Ile Val Phe Leu His Ile Gln Lys
 80 85 90
 Thr Gly Gly Thr Thr Phe Gly Arg His Leu Val Arg Asn Ile Gln
 95 100 105
 Leu Glu Gln Pro Cys Glu Cys Arg Val Gly Gln Lys Lys Cys Thr
 110 115 120
 Cys His Arg Pro Gly Lys Arg Glu Thr Trp Leu Phe Ser Arg Phe
 125 130 135
 Ser Thr Gly Trp Ser Cys Gly Leu His Ala Asp Trp Thr Glu Leu
 140 145 150
 Thr Ser Cys Val Pro Ser Val Val Asp Gly Lys Arg Asp Ala Arg
 155 160 165
 Leu Arg Pro Ser Arg Trp Arg Ile Phe Gln Ile Leu Asp Ala Ala
 170 175 180
 Ser Lys Asp Lys Arg Gly Ser Pro Asn Thr Asn Ala Gly Ala Asn
 185 190 195
 Ser Pro Ser Ser Thr Lys Thr Arg Asn Thr Ser Lys Ser Gly Lys
 200 205 210
 Asn Phe His Tyr Ile Thr Ile Leu Arg Asp Pro Val Ser Arg Tyr
 215 220 225
 Leu Ser Glu Trp Arg His Val Gln Arg Gly Ala Thr Trp Lys Ala
 230 235 240
 Ser Leu His Val Cys Asp Gly Arg Pro Thr Ser Glu Glu Leu
 245 250 255
 Pro Ser Cys Tyr Thr Gly Asp Asp Trp Ser Gly Cys Pro Leu Lys
 260 265 270
 Glu Phe Met Asp Cys Pro Tyr Asn Leu Ala Asn Asn Arg Gln Val
 275 280 285
 Arg Met Leu Ser Asp Leu Thr Leu Val Gly Cys Tyr Asn Leu Ser
 290 295 300
 Val Met Pro Glu Lys Gln Arg Asn Lys Val Leu Leu Glu Ser Ala
 305 310 315
 Lys Ser Asn Leu Lys His Met Ala Phe Phe Gly Leu Thr Glu Phe
 320 325 330
 Gln Arg Lys Thr Gln Tyr Leu Phe Glu Lys Thr Phe Asn Met Asn
 335 340 345

WO 02/12467

PCT/US01/24382

```

Phe Ile Ser Pro Phe Thr Gln Tyr Asn Thr Thr Arg Ala Ser Ser
350 355 360
Val Glu Ile Asn Glu Glu Ile Gln Lys Arg Ile Glu Gly Leu Asn
365 370 375
Phe Leu Asp Met Glu Leu Tyr Ser Tyr Ala Lys Asp Leu Phe Leu
380 385 390
Gln Arg Tyr Gln Phe Met Arg Gln Lys Glu His Gln Glu Ala Arg
395 400 405
Arg Lys Arg Gln Glu Gln Arg Lys Phe Leu Lys Gly Arg Leu Leu
410 415 420
Gln Thr His Phe Gln Ser Gln Gly Gln Gly Gln Ser Gln Asn Pro
425 430 435
Asn Gln Asn Gln Ser Gln Asn Pro Asn Pro Asn Ala Asn Gln Asn
440 445 450
Leu Thr Gln Asn Leu Met Gln Asn Leu Thr Gln Ser Leu Ser Gln
455 460 465
Lys Glu Asn Arg Glu Ser Pro Lys Gln Asn Ser Gly Lys Glu Gln
470 475 480
Asn Asp Asn Thr Ser Asn Gly Thr Asn Asp Tyr Ile Gly Ser Val
485 490 495
Glu Lys Trp Arg

```

```

<210> 7
<211> 222
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7480367CD1

```

```

<400> 7
Met Ala Glu Lys Pro Lys Leu His Tyr Ser Asn Ala Arg Gly Ser
1 5 10 15
Met Glu Ser Ile Arg Trp Leu Leu Ala Ala Ala Gly Val Glu Leu
20 25 30
Glu Glu Lys Phe Leu Glu Ser Ala Glu Asp Leu Asp Lys Leu Arg
35 40 45
Asn Asp Gly Ser Leu Leu Phe Gln Gln Val Pro Met Val Glu Ile
50 55 60
Asp Gly Met Lys Leu Val Gln Thr Arg Ala Ile Leu Asn Tyr Ile
65 70 75
Ala Ser Lys Tyr Asn Leu Tyr Gly Lys Asp Met Lys Glu Arg Ala
80 85 90
Leu Ile Asp Met Tyr Thr Glu Gly Ile Val Asp Leu Thr Glu Met
95 100 105
Ile Leu Leu Leu Leu Ile Cys Gln Pro Glu Glu Arg Asp Ala Lys
110 115 120
Thr Ala Leu Val Lys Glu Lys Ile Lys Asn Arg Tyr Phe Pro Ala
125 130 135
Phe Glu Lys Val Leu Lys Ser His Arg Gln Asp Tyr Leu Val Gly
140 145 150
Asn Lys Leu Ser Trp Ala Asp Ile His Leu Val Glu Leu Phe Tyr
155 160 165
Tyr Val Glu Glu Leu Asp Ser Ser Leu Ile Ser Ser Phe Pro Leu
170 175 180
Leu Lys Ala Leu Lys Thr Arg Ile Ser Asn Leu Pro Thr Val Lys
185 190 195
Lys Phe Leu Gln Pro Gly Ser Gln Arg Lys Pro Pro Met Asp Glu
200 205 210
Lys Ser Leu Glu Glu Ala Arg Lys Ile Phe Arg Phe

```

WO 02/12467

PCT/US01/24382

215 220

<210> 8
<211> 330
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 8069390CD1

```

<400> 8
Met Ala Ala Val Asp Ser Phe Tyr Leu Leu Tyr Arg Glu Ile Ala
1 5 10 15
Arg Ser Cys Asn Cys Tyr Met Glu Ala Leu Ala Leu Val Gly Ala
20 25 30
Trp Tyr Thr Ala Arg Lys Ser Ile Thr Val Ile Cys Asp Phe Tyr
35 40 45
Ser Leu Ile Arg Leu His Phe Ile Pro Arg Leu Gly Ser Arg Ala
50 55 60
Asp Leu Ile Lys Gln Tyr Gly Arg Trp Ala Val Val Ser Gly Ala
65 70 75
Thr Asp Gly Ile Gly Lys Ala Tyr Ala Glu Glu Leu Ala Ser Arg
80 85 90
Gly Leu Asn Ile Ile Leu Ile Ser Arg Asn Glu Glu Lys Leu Gln
95 100 105
Val Val Ala Lys Asp Ile Ala Asp Thr Tyr Lys Val Glu Thr Asp
110 115 120
Ile Ile Val Ala Asp Phe Ser Ser Gly Arg Glu Ile Tyr Leu Pro
125 130 135
Ile Arg Glu Ala Leu Lys Asp Lys Asp Val Gly Ile Leu Val Asn
140 145 150
Asn Val Gly Val Phe Tyr Pro Tyr Pro Gln Tyr Phe Thr Gln Leu
155 160 165
Ser Glu Asp Lys Leu Trp Asp Ile Ile Asn Val Asn Ile Ala Ala
170 175 180
Ala Ser Leu Met Val His Val Val Leu Pro Gly Met Val Glu Arg
185 190 195
Lys Lys Gly Ala Ile Val Thr Ile Ser Ser Gly Ser Cys Cys Lys
200 205 210
Pro Thr Pro Gln Leu Ala Ala Phe Ser Ala Ser Lys Ala Tyr Leu
215 220 225
Asp His Phe Ser Arg Ala Leu Gln Tyr Glu Tyr Ala Ser Lys Gly
230 235 240
Ile Phe Val Gln Ser Leu Ile Pro Phe Tyr Val Ala Thr Ser Met
245 250 255
Thr Ala Pro Ser Asn Phe Leu His Arg Cys Ser Trp Leu Val Pro
260 265 270
Ser Pro Lys Val Tyr Ala His His Ala Val Ser Thr Leu Gly Ile
275 280 285
Ser Lys Arg Thr Thr Gly Tyr Trp Ser His Ser Ile Gln Phe Leu
290 295 300
Phe Ala Gln Tyr Met Pro Glu Trp Leu Trp Val Trp Gly Ala Asn
305 310 315
Ile Leu Asn Arg Ser Leu Arg Lys Glu Ala Leu Ser Cys Thr Ala
320 325 330

```

<210> 9
<211> 303
<212> PRT
<213> Homo sapiens

WO 02/12467

PCT/US01/24382

```

<220>
<221> misc_feature
<223> Incyte ID No: 7473869CD1

<400> 9
Met Tyr Val Ser Thr Arg Glu Arg Tyr Lys Trp Leu Arg Phe Ser
1      5      10      15
Glu Asp Cys Leu Tyr Leu Asn Val Tyr Ala Pro Ala Arg Ala Pro
20     25     30
Gly Asp Pro Gln Leu Pro Val Met Val Trp Phe Pro Gly Gly Ala
35     40     45
Phe Ile Val Gly Ala Ala Ser Ser Tyr Glu Gly Ser Asp Leu Ala
50     55     60
Ala Arg Glu Lys Val Val Leu Val Phe Leu Gln His Arg Leu Gly
65     70     75
Ile Phe Gly Phe Leu Ser Thr Asp Asp Ser His Ala Arg Gly Asn
80     85     90
Trp Gly Leu Leu Asp Gln Met Ala Ala Leu Arg Trp Val Gln Glu
95     100    105
Asn Ile Ala Ala Phe Gly Gly Asp Pro Gly Asn Val Thr Leu Phe
110    115    120
Gly Gln Ser Ala Gly Ala Met Ser Ile Ser Gly Leu Met Met Ser
125    130    135
Pro Leu Ala Ser Gly Leu Phe His Arg Ala Ile Ser Gln Ser Gly
140    145    150
Thr Ala Leu Phe Arg Leu Phe Ile Thr Ser Asn Pro Leu Lys Val
155    160    165
Ala Lys Lys Val Ala His Leu Ala Gly Cys Asn His Asn Ser Thr
170    175    180
Gln Ile Leu Val Asn Cys Leu Arg Ala Leu Ser Gly Thr Lys Val
185    190    195
Met Arg Val Ser Asn Lys Met Arg Phe Leu Gln Leu Asn Phe Gln
200    205    210
Arg Asp Pro Glu Glu Ile Ile Trp Ser Met Ser Pro Val Val Asp
215    220    225
Gly Val Val Ile Pro Asp Asp Pro Leu Val Leu Leu Thr Gln Gly
230    235    240
Lys Val Ser Ser Val Pro Tyr Leu Leu Gly Val Asn Asn Leu Glu
245    250    255
Phe Asn Trp Leu Leu Pro Tyr Ile Met Lys Phe Pro Leu Asn Arg
260    265    270
Gln Ala Met Arg Lys Glu Thr Ile Thr Lys Met Leu Trp Ser Thr
275    280    285
Arg Thr Leu Leu Val Arg Asp Pro Ala Gly Arg Gly Ala Gln Phe
290    295    300
Gly Gln Gly

```

```

<210> 10
<211> 584
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7478588CD1

<400> 10
Met Pro Ser Thr Val Leu Pro Ser Thr Val Leu Pro Ser Leu Leu
1      5      10      15
Pro Thr Ala Gly Ala Gly Trp Ser Met Arg Trp Ile Leu Cys Trp
20     25     30

```

WO 02/12467

PCT/US01/24382

Ser Leu Thr Leu Cys Leu Met Ala Gln Thr Ala Leu Gly Ala Leu 35 40 45
His Thr Lys Arg Pro Gln Val Val Thr Lys Tyr Gly Thr Leu Gln 50 55 60
Gly Lys Gln Met His Val Gly Lys Thr Pro Ile Gln Val Phe Leu 65 70 75
Gly Val Pro Phe Ser Arg Pro Pro Leu Gly Ile Leu Arg Phe Ala 80 85 90
Pro Pro Glu Pro Pro Glu Pro Trp Lys Gly Ile Arg Asp Ala Thr 95 100 105
Thr Tyr Pro Pro Gly Cys Leu Gln Glu Ser Trp Gly Gln Leu Ala 110 115 120
Ser Met Tyr Val Ser Thr Arg Glu Arg Tyr Lys Trp Leu Arg Phe 125 130 135
Ser Glu Asp Cys Leu Tyr Leu Asn Val Tyr Ala Pro Ala Arg Ala 140 145 150
Pro Gly Asp Pro Gln Leu Pro Val Met Val Trp Phe Pro Gly Gly 155 160 165
Ala Phe Ile Val Gly Ala Ala Ser Ser Tyr Glu Gly Ser Asp Leu 170 175 180
Ala Ala Arg Glu Lys Val Val Leu Val Phe Leu Gln His Arg Leu 185 190 195
Gly Ile Phe Gly Phe Leu Ser Thr Asp Asp Ser His Ala Arg Gly 200 205 210
Asn Trp Gly Leu Leu Asp Gln Met Ala Ala Leu Arg Trp Val Gln 215 220 225
Glu Asn Ile Ala Ala Phe Gly Gly Asp Pro Gly Asn Val Thr Leu 230 235 240
Phe Gly Gln Ser Ala Gly Ala Met Ser Ile Ser Gly Leu Met Met 245 250 255
Ser Pro Leu Ala Ser Gly Leu Phe His Arg Ala Ile Ser Gln Ser 260 265 270
Gly Thr Ala Leu Phe Arg Leu Phe Ile Thr Ser Asn Pro Leu Lys 275 280 285
Val Ala Lys Lys Val Ala His Leu Ala Gly Cys Asn His Asn Ser 290 295 300
Thr Gln Ile Leu Val Asn Cys Leu Arg Ala Leu Ser Gly Thr Lys 305 310 315
Val Met Arg Val Ser Asn Lys Met Arg Phe Leu Gln Leu Asn Phe 320 325 330
Gln Arg Asp Pro Glu Glu Ile Ile Trp Ser Met Ser Pro Val Val 335 340 345
Asp Gly Val Val Ile Pro Asp Asp Pro Leu Val Leu Leu Thr Gln 350 355 360
Gly Lys Val Ser Ser Val Pro Tyr Leu Leu Gly Val Asn Asn Leu 365 370 375
Glu Phe Asn Trp Leu Leu Pro Tyr Ile Met Lys Phe Pro Leu Asn 380 385 390
Arg Gln Ala Met Arg Lys Glu Thr Ile Thr Lys Met Leu Trp Ser 395 400 405
Thr Arg Thr Leu Leu Asn Ile Thr Lys Glu Gln Val Pro Leu Val 410 415 420
Val Glu Glu Tyr Leu Asp Asn Val Asn Glu His Asp Trp Lys Met 425 430 435
Leu Arg Asn Arg Met Met Asp Ile Val Gln Asp Ala Thr Phe Val 440 445 450
Tyr Ala Thr Leu Gln Thr Ala His Tyr His Arg Asp Ala Gly Leu 455 460 465
Pro Val Tyr Leu Tyr Glu Phe Glu His His Ala Arg Gly Ile Ile 470 475 480
Val Lys Pro Arg Thr Asp Gly Ala Asp His Gly Asp Glu Met Tyr 485 490 495

WO 02/12467

PCT/US01/24382

```

Phe Leu Phe Gly Gly Pro Phe Ala Thr Gly Leu Ser Met Gly Lys
500 505 510
Glu Lys Ala Leu Ser Leu Gln Met Met Lys Tyr Trp Ala Asn Phe
515 520 525
Ala Arg Thr Gly Asn Pro Asn Asp Gly Asn Leu Pro Cys Trp Pro
530 535 540
Arg Tyr Asn Lys Asp Glu Lys Tyr Leu Gln Leu Asp Phe Thr Thr
545 550 555
Arg Val Gly Met Lys Leu Lys Glu Lys Lys Met Ala Phe Trp Met
560 565 570
Ser Leu Tyr Gln Ser Gln Arg Pro Glu Lys Gln Arg Gln Phe
575 580

```

```

<210> 11
<211> 508
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 55046125CD1

```

```

<400> 11
Met His Val Leu Arg Arg Arg Trp Asp Leu Gly Ser Leu Cys Arg
1 5 10 15
Ala Leu Leu Thr Arg Gly Leu Ala Ala Leu Gly His Ser Leu Lys
20 25 30
His Val Leu Gly Ala Ile Phe Ser Lys Ile Phe Gly Pro Met Ala
35 40 45
Ser Val Gly Asn Met Asp Glu Lys Ser Asn Lys Leu Leu Ala
50 55 60
Leu Val Met Leu Phe Leu Phe Ala Val Ile Val Leu Gln Tyr Val
65 70 75
Cys Pro Gly Thr Glu Cys Gln Leu Leu Arg Leu Gln Ala Phe Ser
80 85 90
Ser Pro Val Pro Asp Pro Tyr Arg Ser Glu Asp Glu Ser Ser Ala
95 100 105
Arg Phe Val Pro Arg Tyr Asn Phe Thr Arg Gly Asp Leu Leu Arg
110 115 120
Lys Val Asp Phe Asp Ile Lys Gly Asp Asp Leu Ile Val Phe Leu
125 130 135
His Ile Gln Lys Thr Gly Gly Thr Thr Phe Gly Arg His Leu Val
140 145 150
Arg Asn Ile Gln Leu Glu Gln Pro Cys Glu Cys Arg Val Gly Gln
155 160 165
Lys Lys Cys Thr Cys His Arg Pro Gly Lys Arg Glu Thr Trp Leu
170 175 180
Phe Ser Arg Phe Ser Thr Gly Trp Ser Cys Gly Leu His Ala Asp
185 190 195
Trp Thr Glu Leu Thr Ser Cys Val Pro Ser Val Val Asp Gly Lys
200 205 210
Arg Asp Ala Arg Leu Arg Pro Ser Arg Asn Phe His Tyr Ile Thr
215 220 225
Ile Leu Arg Asp Pro Val Ser Arg Tyr Leu Ser Glu Trp Arg His
230 235 240
Val Gln Arg Gly Ala Thr Trp Lys Ala Ser Leu His Val Cys Asp
245 250 255
Gly Arg Pro Pro Thr Ser Glu Glu Leu Pro Ser Cys Tyr Thr Gly
260 265 270
Asp Asp Trp Ser Gly Cys Pro Leu Lys Glu Phe Met Asp Cys Pro
275 280 285
Tyr Asn Leu Ala Asn Asn Arg Gln Val Arg Met Leu Ser Asp Leu

```

WO 02/12467

PCT/US01/24382

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

<210> 12
 <211> 439
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No.: 3538709CD1

<400> 12
 Met Leu Thr Gly Val Thr Asp Gly Ile Phe Cys Cys Leu Leu Gly
 1 5 10 15
 Thr Pro Pro Asn Ala Val Gly Pro Leu Glu Ser Val Glu Ser Ser
 20 25 30
 Asp Gly Tyr Thr Phe Val Glu Val Lys Pro Gly Arg Val Leu Arg
 35 40 45
 Val Lys His Ala Gly Pro Ala Pro Ala Ala Ala Pro Pro Pro Pro
 50 55 60
 Ser Ser Ala Ser Ser Asp Ala Ala Gln Gly Asp Leu Ser Gly Leu
 65 70 75
 Val Arg Cys Gln Arg Arg Ile Thr Val Tyr Arg Asn Gly Arg Leu
 80 85 90
 Leu Val Glu Asn Leu Gly Arg Ala Pro Arg Ala Asp Leu Leu His
 95 100 105
 Gly Gln Asn Gly Ser Gly Glu Pro Pro Ala Ala Leu Glu Val Glu
 110 115 120
 Leu Ala Asp Pro Ala Gly Ser Asp Gly Arg Leu Ala Pro Gly Ser
 125 130 135
 Ala Gly Ser Gly Ser Gly Ser Gly Ser Gly Gly Arg Arg Arg Arg
 140 145 150
 Ala Arg Arg Pro Lys Arg Thr Ile His Ile Asp Cys Glu Lys Arg
 155 160 165

WO 02/12467

PCT/US01/24382

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

<210> 13
 <211> 514
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 71563101CD1

<400> 13
 Met Thr Leu Ile Trp Arg His Leu Leu Arg Pro Leu Cys Leu Val
 1 5 10 15
 Thr Ser Ala Pro Arg Ile Leu Glu Met His Pro Phe Leu Ser Leu
 20 25 30
 Gly Thr Ser Arg Thr Ser Val Thr Lys Leu Ser Leu His Thr Lys
 35 40 45
 Pro Arg Met Pro Pro Cys Asp Phe Met Pro Glu Arg Tyr Gln Ser
 50 55 60
 Leu Gly Tyr Asn Arg Val Leu Glu Ile His Lys Glu His Leu Ser
 65 70 75
 Pro Val Val Thr Ala Tyr Phe Gln Lys Pro Leu Leu Leu His Gln
 80 85 90
 Gly His Met Glu Trp Leu Phe Asp Ala Glu Gly Asn Arg Tyr Leu

WO 02/12467

PCT/US01/24382

```

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

```

```

<210> 14
<211> 226
<212> PRT
<213> Homo sapiens

```

WO 02/12467

PCT/US01/24382

<220>
 <221> misc_feature
 <223> Inocyte ID No: 7472027CD1

<400> 14
 Met Arg Leu Cys Glu Lys Thr Glu Leu Gln Leu Ile Gly Val Pro
 1 5 10 15
 Glu Ser Asp Arg Glu Asn Gly Thr Lys Leu Glu Asn Thr Phe Gln
 20 25 30
 Asp Ile Ile Gln Glu Asn Phe Pro Asn Leu Ala Arg Gln Ala Asn
 35 40 45
 Ile Gln Ile Gln Met Ala Gly Gly Ser Ile Trp Ile Glu Gly Ile
 50 55 60
 Pro Phe Pro Ser Asn Asn Phe Thr Asp Leu Arg Arg Leu Gln Asp
 65 70 75
 Glu Ile Val Leu Arg Asp Glu Asp Val Ile Thr Leu Ser Tyr Pro
 80 85 90
 Lys Ser Gly Ser Phe Trp Ile Val Glu Ile Ile Ser Leu Ile His
 95 100 105
 Ser Lys Gly Asp Pro Ser Trp Val Gln Ser Val Val Pro Trp Asp
 110 115 120
 Arg Ser Pro Trp Ile Glu Val Lys Arg Lys Lys Ala Gly Leu Glu
 125 130 135
 Ser Gln Lys Gly Pro His Leu Tyr Thr Ser His Leu Pro Ile Gln
 140 145 150
 Leu Phe Pro Lys Ser Phe Leu Asn Ser Lys Ala Lys Cys Ile Tyr
 155 160 165
 Pro His Val Leu Met Leu Val Val Leu Ile Leu Gly His Lys Ser
 170 175 180
 Gln Trp Ser Ile Ala Ile Lys Ile Ser Glu Asn Ala Glu Ala Thr
 185 190 195
 Ser Lys Leu Gly Asn Gly Gln Arg Leu Glu Glu Phe Gly Gly Leu
 200 205 210
 Arg Arg Arg Gln Glu Asp Glu Arg Ser Leu Glu Phe Leu Arg Asp
 215 220 225
 Cys

<210> 15
 <211> 121
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Inocyte ID No: 7480358CD1

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

WO 02/12467

PCT/US01/24382

Gly	His	Tyr	Ser	Ile	Thr	Ala	His	Thr	Met	Val	Ser	Met	Thr	Thr
				110					115					120
Arg														

<210> 16
 <211> 486
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Enctype ID No. 1618256CD1

<400> 16
 Met Gly Pro Leu Ser Pro Ala Arg Thr Leu Arg Leu Trp Gly Pro
 1 5 10 15
 Arg Ser Leu Gly Val Ala Leu Gly Val Phe Met Thr Ile Gly Phe
 20 25 30
 Ala Leu Gln Leu Leu Gly Gly Pro Phe Gln Arg Arg Leu Pro Gly
 35 40 45
 Leu Gln Leu Arg Gln Pro Ser Ala Pro Ser Leu Arg Pro Ala Leu
 50 55 60
 Pro Ser Cys Pro Pro Arg Gln Arg Leu Val Phe Leu Lys Thr His
 65 70 75
 Lys Ser Gly Ser Ser Ser Val Leu Ser Leu Leu His Arg Tyr Gly
 80 85 90
 Asp Gln His Gly Leu Arg Phe Ala Leu Pro Ala Arg Tyr Gln Phe
 95 100 105
 Gly Tyr Pro Lys Leu Phe Gln Ala Ser Arg Val Lys Gly Tyr Arg
 110 115 120
 Pro Gln Gly Gly Gly Thr Gln Leu Pro Phe His Ile Leu Cys His
 125 130 135
 His Met Arg Phe Asn Leu Lys Glu Val Leu Gln Val Met Pro Ser
 140 145 150
 Asp Ser Phe Phe Phe Ser Ile Val Arg Asp Pro Ala Ala Leu Ala
 155 160 165
 Arg Ser Ala Phe Ser Tyr Tyr Lys Ser Thr Ser Ser Ala Phe Arg
 170 175 180
 Lys Ser Pro Ser Leu Ala Ala Phe Leu Ala Asn Pro Arg Gly Phe
 185 190 195
 Tyr Arg Pro Gly Ala Arg Gly Asp His Tyr Ala Arg Asn Leu Leu
 200 205 210
 Trp Phe Asp Phe Gly Leu Pro Phe Pro Pro Glu Lys Arg Ala Lys
 215 220 225
 Arg Gly Asn Ile His Pro Pro Arg Asp Pro Asn Pro Pro Gln Leu
 230 235 240
 Gln Val Leu Pro Ser Gly Ala Gly Pro Arg Ala Gln Thr Leu Asn
 245 250 255
 Pro Asn Ala Leu Ile His Pro Val Ser Thr Val Thr Asp His Arg
 260 265 270
 Ser Gln Ile Ser Ser Pro Ala Ser Phe Asp Leu Gly Ser Ser Ser
 275 280 285
 Phe Ile Gln Trp Gly Leu Ala Trp Leu Asp Ser Val Phe Asp Leu
 290 295 300
 Val Met Val Ala Glu Tyr Phe Asp Glu Ser Leu Val Leu Leu Ala
 305 310 315
 Asp Ala Leu Cys Trp Gly Leu Asp Asp Val Val Gly Phe Met His
 320 325 330
 Asn Ala Gln Ala Gly His Lys Gln Gly Leu Ser Thr Val Ser Asn
 335 340 345
 Ser Gly Leu Thr Ala Glu Asp Arg Gln Leu Thr Ala Arg Ala Arg

WO 02/12467

PCT/US01/24382

```

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

```

```

<210> 17
<211> 649
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 3387523CD1

```

```

<400> 17
Met Tyr Ile Ser Cys Leu Ser Leu Ser Leu Phe Phe Leu Ser Gly
1          5          10          15
Pro Leu Gln Arg Val Leu Glu Val Ser Asn His Trp Trp Tyr Ser
20          25          30
Met Leu Ile Leu Pro Pro Leu Leu Lys Asp Ser Val Ala Ala Pro
35          40          45
Leu Leu Ser Ala Tyr Tyr Pro Asp Cys Val Gly Met Ser Pro Ser
50          55          60
Cys Thr Ser Thr Asn Arg Ala Ala Thr Gly Asn Ala Ser Pro
65          70          75
Gly Lys Leu Glu His Ser Lys Ala Ala Leu Ser Val His Val Pro
80          85          90
Gly Met Asn Arg Tyr Phe Gln Pro Phe Tyr Gln Pro Asn Glu Cys
95          100          105
Gly Lys Ala Leu Cys Val Arg Pro Asp Val Met Glu Leu Asp Glu
110          115          120
Leu Tyr Glu Phe Pro Glu Tyr Ser Arg Asp Pro Thr Met Tyr Leu
125          130          135
Ala Leu Arg Asn Leu Ile Leu Ala Leu Trp Tyr Thr Asn Cys Lys
140          145          150
Glu Ala Leu Thr Pro Gln Lys Cys Ile Pro His Ile Ile Val Arg
155          160          165
Gly Leu Val Arg Ile Arg Cys Val Gln Glu Val Glu Arg Ile Leu
170          175          180
Tyr Phe Met Thr Arg Lys Gly Leu Ile Asn Thr Gly Val Leu Ser
185          190          195
Val Gly Ala Asp Gln Tyr Leu Leu Pro Lys Asp Tyr His Asn Lys
200          205          210
Ser Val Ile Ile Ile Gly Ala Gly Pro Ala Gly Leu Ala Ala Ala
215          220          225
Arg Gln Leu His Asn Phe Gly Ile Lys Val Thr Val Leu Glu Ala
230          235          240

```

WO 02/12467

PCT/US01/24382

Lys Asp Arg Ile Gly Gly Arg Val Trp Asp Asp Lys Ser Phe Lys 245
 245
 Gly Val Thr Val Gly Arg Gly Ala Gln Ile Val Asn Gly Cys Ile 255
 260
 Asn Asn Pro Val Ala Leu Met Cys Glu Gln Leu Gly Ile Ser Met 270
 275
 His Lys Phe Gly Glu Arg Cys Asp Leu Ile Gln Glu Gly Gly Arg 285
 290
 Ile Thr Asp Pro Thr Ile Asp Lys Arg Met Asp Phe His Phe Asn 300
 305
 Ala Leu Leu Asp Val Val Ser Glu Trp Arg Lys Asp Lys Thr Gln 315
 320
 Leu Gln Asp Val Pro Leu Gly Glu Lys Ile Glu Glu Ile Tyr Lys 330
 335
 Ala Phe Ile Lys Glu Ser Gly Ile Gln Phe Ser Glu Leu Glu Gly 345
 350
 Gln Val Leu Gln Phe His Leu Ser Asn Leu Glu Tyr Ala Cys Gly 360
 365
 Ser Asn Leu His Gln Val Ser Ala Arg Ser Trp Asp His Asn Glu 375
 380
 Phe Phe Ala Gln Phe Ala Gly Asp His Thr Leu Leu Thr Pro Gly 390
 395
 Tyr Ser Val Ile Ile Glu Lys Leu Ala Glu Gly Leu Asp Ile Gln 405
 410
 Leu Lys Ser Pro Val Gln Cys Ile Asp Tyr Ser Gly Asp Glu Val 420
 425
 Gln Val Thr Thr Thr Asp Gly Thr Gly Tyr Ser Ala Gln Lys Val 435
 440
 Leu Val Thr Val Pro Leu Ala Leu Leu Gln Lys Gly Ala Ile Gln 450
 455
 Phe Asn Pro Pro Leu Ser Glu Lys Lys Met Lys Ala Ile Asn Ser 465
 470
 Leu Gly Ala Gly Ile Ile Glu Lys Ile Ala Leu Gln Phe Pro Tyr 480
 485
 Arg Phe Trp Asp Ser Lys Val Gln Gly Ala Asp Phe Phe Gly His 495
 500
 Val Pro Pro Ser Ala Ser Lys Arg Gly Leu Phe Ala Val Phe Tyr 510
 515
 Asp Met Asp Pro Gln Lys Lys His Ser Val Leu Met Ser Val Ile 525
 530
 Ala Gly Glu Ala Val Ala Ser Val Arg Thr Leu Asp Asp Lys Gln 540
 545
 Val Leu Gln Gln Cys Met Ala Thr Leu Arg Glu Leu Phe Lys Glu 555
 560
 Gln Glu Val Pro Asp Pro Thr Lys Tyr Phe Val Thr Arg Trp Ser 570
 575
 Thr Asp Pro Trp Ile Gln Met Ala Tyr Ser Phe Val Lys Thr Gly 585
 590
 Gly Ser Gly Glu Ala Tyr Asp Ile Ile Ala Glu Asp Ile Gln Gly 600
 605
 Thr Val Phe Phe Ala Gly Glu Ala Thr Asn Arg His Phe Pro Gln 615
 620
 Thr Val Thr Gly Ala Tyr Leu Ser Gly Val Arg Glu Ala Ser Lys 630
 635
 Ile Ala Ala Phe 640
 645

<210> 18
 <211> 258
 <212> PRT
 <213> Homo sapiens

WO 02/12467

PCT/US01/24382

<220>
 <221> misc_feature
 <223> Incyte ID No: 55142051CD1

<400> 18
 Met Ser Pro Ala Ile Ala Leu Ala Phe Leu Pro Leu Val Val Thr
 1 5 10
 Leu Leu Val Arg Tyr Arg His Tyr Phe Arg Leu Leu Val Arg Thr
 20 25 30
 Val Leu Leu Arg Ser Leu Arg Asp Cys Leu Ser Gly Leu Arg Ile
 35 40 45
 Glu Glu Arg Ala Phe Ser Tyr Val Leu Thr His Ala Leu Pro Gly
 50 55 60
 Asp Pro Gly His Ile Leu Thr Thr Leu Asp His Trp Ser Ser Arg
 65 70 75
 Cys Glu Tyr Leu Ser His Met Gly Pro Val Lys Gly Gln Ile Leu
 80 85 90
 Met Arg Leu Val Glu Glu Lys Ala Pro Ala Cys Val Leu Glu Leu
 95 100 105
 Gly Thr Tyr Cys Gly Tyr Ser Thr Leu Leu Ile Ala Arg Ala Leu
 110 115 120
 Pro Pro Gly Gly Arg Leu Leu Thr Val Glu Arg Asp Pro Arg Thr
 125 130 135
 Ala Ala Val Ala Glu Lys Leu Ile Arg Leu Ala Gly Phe Asp Glu
 140 145 150
 His Met Val Glu Leu Ile Val Gly Ser Ser Glu Asp Val Ile Pro
 155 160 165
 Cys Leu Arg Thr Gln Tyr Gln Leu Ser Arg Ala Asp Leu Val Leu
 170 175 180
 Leu Ala His Arg Pro Arg Cys Tyr Leu Arg Asp Leu Gln Leu Leu
 185 190 195
 Glu Ala His Ala Leu Leu Pro Ala Gly Ala Thr Val Leu Ala Asp
 200 205 210
 His Val Leu Phe Pro Gly Ala Pro Arg Phe Leu Gln Tyr Ala Lys
 215 220 225
 Ser Cys Gly Arg Tyr Arg Cys Arg Leu His His Thr Gly Leu Pro
 230 235 240
 Asp Phe Pro Ala Ile Lys Asp Gly Ile Ala Gln Leu Thr Tyr Ala
 245 250 255
 Gly Pro Gly

<210> 19
 <211> 544
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7395274CD1

<400> 19
 Met Ser Ser Pro Gly Pro Ser Gln Pro Pro Ala Glu Asp Pro Pro
 1 5 10 15
 Trp Pro Ala Arg Leu Leu Arg Ala Pro Leu Gly Leu Leu Arg Leu
 20 25 30
 Asp Pro Ser Gly Gly Ala Leu Leu Leu Cys Gly Leu Val Ala Leu
 35 40 45
 Leu Gly Trp Ser Trp Leu Arg Arg Arg Ala Arg Gly Ile Pro
 50 55 60
 Pro Gly Pro Thr Pro Trp Pro Leu Val Gly Asn Phe Gly His Val
 65 70 75

WO 02/12467

PCT/US01/24382

Leu Leu Pro Pro Phe Leu Arg Arg Arg Ser Trp Leu Ser Ser Arg
 80 85 90
 Thr Arg Ala Ala Gly Ile Asp Pro Ser Val Ile Gly Pro Gln Val
 95 100 105
 Leu Leu Ala His Leu Ala Arg Val Tyr Gly Ser Ile Phe Ser Phe
 110 115 120
 Phe Ile Gly His Tyr Leu Val Val Val Leu Ser Asp Phe His Ser
 125 130 135
 Val Arg Glu Ala Leu Val Gln Gln Ala Glu Val Phe Ser Asp Arg
 140 145 150
 Pro Arg Val Pro Leu Ile Ser Ile Val Thr Lys Glu Lys Gly Val
 155 160 165
 Val Phe Ala His Tyr Gly Pro Val Trp Arg Gln Gln Arg Lys Phe
 170 175 180
 Ser His Ser Thr Leu Arg His Phe Gly Leu Gly Lys Leu Ser Leu
 185 190 195
 Glu Pro Lys Ile Ile Glu Glu Phe Lys Tyr Val Lys Ala Glu Met
 200 205 210
 Gln Lys His Gly Glu Asp Pro Phe Cys Pro Phe Ser Ile Ile Ser
 215 220 225
 Asn Ala Val Ser Asn Ile Ile Cys Ser Leu Cys Phe Gly Gln Arg
 230 235 240
 Phe Asp Tyr Thr Asn Ser Glu Phe Lys Lys Met Leu Gly Phe Met
 245 250 255
 Ser Arg Gly Leu Glu Ile Cys Leu Asn Ser Gln Val Leu Leu Val
 260 265 270
 Asn Ile Cys Pro Trp Leu Tyr Tyr Leu Pro Phe Gly Pro Phe Lys
 275 280 285
 Glu Leu Arg Gln Ile Glu Lys Asp Ile Thr Ser Phe Leu Lys Lys
 290 295 300
 Ile Ile Lys Asp His Gln Glu Ser Leu Asp Arg Glu Asn Pro Gln
 305 310 315
 Asp Phe Ile Asp Met Tyr Leu Leu His Met Glu Glu Glu Arg Lys
 320 325 330
 Asn Asn Ser Asn Ser Ser Phe Asp Glu Glu Tyr Leu Phe Tyr Ile
 335 340 345
 Ile Gly Asp Leu Phe Ile Ala Gly Thr Asp Thr Thr Asn Ser
 350 355 360
 Leu Leu Trp Cys Leu Leu Tyr Met Ser Leu Asn Pro Asp Val Gln
 365 370 375
 Glu Lys Val His Glu Glu Ile Glu Arg Val Ile Gly Ala Asn Arg
 380 385 390
 Ala Pro Ser Leu Thr Asp Lys Ala Gln Met Pro Tyr Thr Glu Ala
 395 400 405
 Thr Ile Met Glu Val Gln Arg Leu Thr Val Val Val Pro Leu Ala
 410 415 420
 Ile Pro His Met Thr Ser Glu Asn Thr Val Leu Gln Gly Tyr Thr
 425 430 435
 Ile Pro Lys Gly Thr Leu Ile Leu Pro Asn Leu Trp Ser Val His
 440 445 450
 Arg Asp Pro Ala Ile Trp Glu Lys Pro Glu Asp Phe Tyr Pro Asn
 455 460 465
 Arg Phe Leu Asp Asp Gln Gly Gln Leu Ile Lys Lys Glu Thr Phe
 470 475 480
 Ile Pro Phe Gly Ile Gly Lys Arg Val Cys Met Gly Glu Gln Leu
 485 490 495
 Ala Lys Met Glu Leu Phe Leu Met Phe Val Ser Leu Met Gln Ser
 500 505 510
 Phe Ala Phe Ala Leu Pro Glu Asp Ser Lys Lys Pro Leu Leu Thr
 515 520 525
 Gly Arg Phe Gly Leu Thr Leu Ala Pro His Pro Phe Asn Ile Thr
 530 535 540

WO 02/12467

PCT/US01/24382

Ile Ser Arg Arg

<210> 20

<211> 2603

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7248285CB1

<400> 20

```

atggcgtggg ccccaccagc caccctcttt ctgttctctg tgetgetagg ccagccccc 60
ccccaccagc cacagtcact gggcaccact aagctccggc tggtcggccc agcagcaca 120
ccagaggagg gccgcctgga ggtgctgcac cagggccagt ggggcaocgt gctgtatgc 180
aaccttgcta tccaggagcc cacagtgctc tgcgcacagc tgggcttcca agctgctctg 240
acctgggccc acagtgccaa gtacggccaa ggggaggagc caatctggct ggcacaatgt 300
cgtctgttgg gccacagagc ctccctggac cagtcggagt ctaatggctg ggcagtcagt 360
gactcagtc actcaqaaga cgttaggggtg atatgcccac ccgggcgcca togtggctac 420
ctttctgaaa ctgtctccaa tgcctctggg cccacgggac agcggctgga ggcggctgg 480
ctcaagccca tccctgcoag tgcacagcag catagcccag tgaccggagg agcctgtgag 540
gtcaagatag agggccactg ccggcaggtg tgtgaccagg gctggaccat gaaccaacgc 600
agggctgtgt gccggatgct gggcttcccc agcaggttgc ctgtgcagac caactactac 660
aggaagctct gggatctgaa gatgaggagc cctaagtcta gctggaagc cctgacgaa 720
aagaactcct tctgatatca ccaggtcacc tgcctgggga cagagcccca catggccaac 780
tgcacaggtc agtggctccc agcccggggc aagctggggc cagcctgccc aggtggcatg 840
catgctgtgg tcaagctgtgt ggcagggccct cacttccggc caccgaaagc aaagcccaaa 900
cgcaaaaggt ctgggcaaga ggaagccagc gtcgcctcgc gctccggggc ccaggctggc 960
gagggccggg tggaaagtct catgaaocgc cagtggggca cggctctgta ccaacagttg 1020
aacctcatct ctgcccaggt ctgtgtctct cagctgggct ttggctctgc tggggagccc 1080
ctctttgggg ccgcgctggg ccaaggccta gggcccaacc acctgagtga ggtgcgctgc 1140
aggggatgat agcggaccct cagcgaactgc cctgcccctg aagtgatcca gaattgtgtc 1200
caacatgaga atgatgctgc tgtcaggtgc aatgtcccta acatgggctt tcaagaatag 1260
gtgcgcttgg ctggtggcgc tatccctgag gaggggctat tggaggtgca ggtggaggtg 1320
aacggggtcc caccgtgggg gagcgtgtgc agtgaaaact gggggctcac cgaagccatg 1380
gtggcctgce gaacagctcg cctgggtttt gccatccatg cctacaagga aacctggctc 1440
tgcctgggga gcccagggc ccaggaaatg gtcgatgagt ggttggctg cccaggccca 1500
gagcctggcc tgcagcagtg ccaggagcac gggccggctg actgtctcca cgttggcgg 1560
cgttctctgg ctggagctct ctgcatggac agtgcaccag acctggtgat gaacgccagc 1620
ctagtgcaag agacggccta ctggaggac ccgccctcca gccagctgta ttgtgcccac 1680
gagggagaact gccctcccaa gctctgggat cacatggact gccctcagc ataccgctgc 1740
ctattgcctc tctccacaca gatctcaaat ctgggcccga ctgactttcg tccaaagact 1800
ggacgctgata gctgggtttg gcaccagctc cacaggcaat accacagcat tgaagctctc 1860
accactacg acctcctcac tctcaatggc tccaaggtgg ctgaggggca caaggccagc 1920
ttctgtctgg agaacacaaa ctgcccacaa ggcctgcagc ggcgctacgc atgtgccaac 1980
tttggagaac agggagtgac tgtagctctg tgggacacct accggctgga catgtattgc 2040
cagtggtgag atatccacga tggggccccc gggaaatata tcttccaggt gattgtgac 2100
ccccactatg aagtggcaga gtcagatctc tccaaacata tgcctcagct ccgctgcagc 2160
tatgatgggc accggctctg gctycacaac tgcacacagc ggaattcata cccagccaal 2220
gcagaactct ccttggagca ggaacagcgt ctcaagaca accctatctg aagctgtcac 2280
tgcacaactc tagctctgc cyatacaca gataacctag ctatttggag ccattgocctt 2340
cacagagtec caactcagag gaadaggccc agtgcacaag ggcaccaaga acctgtctag 2400
gaagcccttt gettgcaaga tcaaccaatc agatggtatt gctccctcag gatggctctg 2460
ggcctgcccc taaggccctg tggcctatgg aatatgtctt ccaggcttgg ctccagctgag 2520
ctccctctct gtaaggaaac ccagtcactc ctgaaatctg ccacagagat ccgggattca 2580
ggagctctca gtttctaaag gag 2603

```

<210> 21

<211> 1745

<212> DNA

<213> Homo sapiens

WO 02/12467

PCT/US01/24382

<220>
<221> misc_feature
<223> Incyte ID No: 7472835CB1

<400> 21
gtaatttttt ttaaatggga tggatgacag tgacagagct ctaaatttct actaaccagc 60
tgagacacaa tggcccaaaa agcagattgct gtgatggag ctggaattag cggactgggg 120
gcatcaagt gctgcctgga tgaagatctg gagcccaact gntttgaaag aaatgatgat 180
attggacatc tctggaattt tcaaaaaaat acttcagaga aaatgcctag tatctacaaa 240
totgtgacca tcaatacttc caaggagatg atgtgcttca gtgacttccc tgtccotgat 300
cablttccca actacatgca caactccaaa ctcatggact acttcgggat gtatgocaca 360
caclttggcc hcttgaatca caatcgtttt aspacfgaag tgcacaaglt gaggagacac 420
ccagattttt ctatcaaatg acaabgggat gtgtgtgtgt agactgaga gaaacagag 480
actttgtgt tttatgggggt cttagtttgc agtggacccc acacagatcc ctacttccca 540
cttcagtcct tccaggttat ggagaaattt gaagcctggt atttccatag tcgggaaatcc 600
aaaagtccc aggacttttc agggaaaaa atcatagtga tccgacttgg aaattctgga 660
gtggatattg cggtaggagct cagtcgtgta gcaaaacagg ttatattcct tagtaotaga 720
cgtggatcat ggaatttaca ccgtgtttgg gataatgggt atcccatgga tagttcattt 780
ttcactcgtt tcaatagttt tctccagaaa atactaacta caccacaaat aaataaccag 840
ctagagaaaa taatgnaactc asgatttaat catgagcact gtggcctgca gectcagcac 900
agagctttaa gtccagccc aactgttagt gatgacccgc caaatccact aatttttggg 960
aaagtcacag taaagcccag cgtgaaggag ttccacgaaa cagatgcaat ttttgaacc 1020
agcactgtag aggagaaat tgaatgttgc atctttgcta caggatacag ttttttttt 1080
tctttccttg atgtctgat caaggttact aacaatgaag tatctctgta taagcttatg 1140
ttccctcctg acctggagaa gccaaccttg cctgtcctcg gtcttatcca accactggcc 1200
atcctcttcc ctatgacaga gctccaactc cgttgggtca cccagatgct caaaggtctg 1260
atcaaatlac octcagcga gaacatgat gcagatattg cccagagaaa aaggctatg 1320
gaaaacggt atgtaaagac acccggccac acaatccag tggatccact tgagtacatg 1380
gatagatttg ccatgcccagc aggggtgaaa cccaacctgc tctctcctt tctctcagat 1440
ccaagcttg ccatgaggt ttctcttggc cctcgacccc calaacgta caactccat 1500
ggcccgcgga aatgggatg gcccccggga gctaacctga cccagagaga gaggatcatc 1560
aagcccctga ggaactgcct tactagttag gacagccac catctcaca gctctcttgg 1620
ataaagatg ccccagtag cctgycattt ctggctctgt gcttggcata ctttcgatat 1680
actccttacc gtaaatgaaa ataaatgaaa gaacactgag ggggaaaaag atggaatagt 1740
ttcta 1765

<210> 22
<211> 1587
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7476203CB1

<400> 22
atgctctccc tgetcagtag gctggcttta ctggccatct ccttctctgt cctgaaactt 60
ggcaactctt gctggagagc gactgtgttt octctggcc cactcccctt ccccaacttt 120
ggaaacctgt ggcagctatg ctttccgag cctccccttt cacttaaaaa ctttccagag 180
aagataggsa atacttttat gcaactggga agcagtggtg ttccactgpc attgcccattg 240
ttcaacgtga ctttccatcc ctgaaaccaa ggcgtgttat gtaagccact tattacttcc 300
cotaaacctt tccctaccag aaatccaggc atcatctgca gcagcggcca cactgycctg 360
caaaagagac gcttctgccc ggtgatgatt cgaaggttag gccatggcaa gctgycctgt 420
gaggtgcagc tgcagaaga ggcagcagag ctggcagaag ccttccgcca ggagcaggtt 480
aagagaccct tgcaccctca ggtatccatt gtcaggcca cagtccaggt cactcgggpc 540
cttgggttgg gccaccactt cctcttagag gatccactt tccaggaact gactccagpc 600
atcgactttg gectgacatt tctcagcact gtttggcpc agctgtatga cgttlttccc 660
tggccctct gccaccctcc aggaccctcc caggagatat ttaggtaaca agaggtctgt 720
ctgagcttaa tccaccagga gatcaccagc cacaactca gggcaccgga ggcctccagc 780
gacttcaatc gctgctacct gcccagatc tcaagccca tggatgacc tgtctccaca 840
ttcaaccagg acaacctggt ccaggtggtg atcgacctgt tcttgggag caccyaccu 900
acagccacca cctgtgctg ggcactcctc cacatgatcc agcagggagc tgtccaggag 960

WO 02/12467

PCT/US01/24382

```

acgggtgcagc tggagctgga caggtgtgtg ggtgtgtccc cggttgtctg ctatgaagac 1020
cgcaagcgac tgccttaaac catgtctgtcc tcaatgacgt gcaagcgcctc agcagcgtca 1080
tggcatgggg tgccttgggc cagtgtgtga cctcaaccg tgtgtgcagc tctccctgga 1140
gcaagggcac cctcaactta ccaaacctgg cctctgtgct ctatgacccct gagtgtgtgg 1200
agaccctctg acagttcaac cctggccact tctggcaaa ggatgaaac tttgtgcca 1260
atgagccttt cctgccatc tctgcaggtc ctatggctca cccagcagac cagctggctc 1320
aatggagct cttctgcat tttgccacc tctccaggac ctttgggtc caactgcca 1380
aagggagccc ggggctcaag ctgagttaca tctttggcgg cacttggcaa ccccagccc 1440
aggagatctg cgcagtccc cgcctgagca gcccagccc tggctctagg gaggatggc 1500
tgtagccact ggggtctcgg agcctgtcc cccatgaagt ccttctcag tctctttgg 1560
tctctgcaaa gttagaaaa aggagga 1587

```

```

<210> 23
<211> 1038
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7478583CB1

```

```

<400> 23
atgaagcag csgtllggt tggctcaaaa gatgaagag ttgaagaac tgaacctaa 60
gaattacagg acaatgaagt taaggttaaa gtatcttggg ctggcatttg tggtaacagt 120
ttacatgaat acttagaagg cctataat tttcaaacag aaagccaga tccattcta 180
ggtcaaaaag cgcagttac attgggtcat gaatttgcag gtgtagtga gaaactggt 240
tcccagttta caaaalita taaaygcyat cyagittgag ttaacctac agtttcaaaa 300
cgtgaaaang aagaaaatat tgacctttat gatggttatt caattatagg cttaggttct 360
gatgggtgat tgcagagtt tacaatgag ccggaagaaa atgtttata actaccagat 420
atgtttctg ataaagaagg tgccttctc gaaaccaag cgttgcagt tcaagcaat 480
aaagaggtg aagtctctat tgggtatac gtatgctatt ttgtgcagg acaatttga 540
ttattaacag tcttagcagc caaagcagct ggtgaaata aaatattgt ttctgatta 600
tcagaaagaa gactaagtas agctaagca ctatgycgca cctatgctat aaactctgt 660
aagcagatc cagttagatg tattaatgag tacaagaaa atggttga tgatcttll 720
gaatggctg gtttagcacc aaccttaaa tctctatag atgttcaaa gcaagaggt 780
acagttggt tcttctctat ttttggctat cctatagat ggaatccat gcaattaact 840
aatcacggag taaaacttac ttctacaatt gcaatacac ctactacat ccaaccaaca 900
attgacttaa tcaagyaag taatttaac gttaaaaly tagtlaclga lgaatlyag 960
ttagaanaa tcttagaatc aggtattgaa caactgtaa atgataaac tcaagcaaaa 1020
atattaatta aattataa 1038

```

```

<210> 24
<211> 1584
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7478585CB1

```

```

<400> 24
atgaaatcat taatgatgg cyotaatggt ggcgtcggta aacatctcgt acgtaaaacta 60
aatgcaagag atgttgatt taogccgggt gttgaaaaa agnacaagt tgaagotta 120
aaagcagatg gtatcagatc aacttaact gatgttgc aaacatctat tgatgatta 180
atagaaatat ttaaaccgta tgaccaaatc ctttttctg tgggtctctg tggaaagca 240
ggtgacgato aaacaatct atgagatta gacggtcag tgaagcaat taaagcaat 300
gaaccgttg gctgtaacc ctltgttatg gtatcaactc agcattcac tgaagagct 360
tttagatgct caggcagact gaaccatc accattgcta aacattatg ggtatgata 420
ttaagccatg caaattaaa atataaccat gtaactcag gogcaatac taacaatac 480
gaaacycaac aatcaatat gagtgtca tttgaaatg tacaatacc gtcatacag 540
agagaagatg tagcagaagt gottgtttct gyttaactg atgaaactt caaagttag 600
cttgcacaa gaccacaaag tatccctcaa gacgatgat ttgatttga aacatagna 660

```

WO 02/12467

PCT/US01/24382

```

actogagaac cacatgcagg tgaggttcaa gtagaatcca tttatgtatc tgtagatcct 720
tccatgcagag gcagaatgaa tgalacaaaan agttatgttc aacctttcca agtgaatgaa 780
ccattacaag gtcababtyt tggaaaagtc acacaaatoga acygtgaacg tctatctgtt 840
ggagattatg tccaggcat attacacatg aaaaagataa alacagtgaa tggagacgat 900
gtgacccctg tggcafcaaa agatgtacaa ttaactttat abttgagtgf tttggcatg 960
ccggaaaga cggcatatcc aggatgctt caaatgggtc aaccacaatc tggcagagc 1020
gltgtcgtgt cagctgcacc aggtgcagta ggalctgtcg taggacaatc tgcataagatt 1080
aaagcgcaa aagttgtcgg tattgtcgtt ggttaagcaga aaacaacata tttacaagat 1140
gaattaggat ttgatgoggc cattgactat aaacaagatg atttgcaca gcaactcgaa 1200
cggcctgtac cagatggtat tgatgtgtat tttgaaaatg tagcggcgtg aattctgtal 1260
gaagtgttta aacacttaaa tggatttcoa cggcttcogg tabtgggtgc aatttcaaga 1320
tataataatg aaaaagcaga tattggaaca cgtatccaag gaacgttgat taanaataca 1380
gcaatgagc aaggttttgt agtagcaaca ttegtgakt aktttanaga agcagcagat 1440
caactgcac aatgggtgtc tgaaggtaaa abbaatttg aagtgaagat agatgaagt 1500
ttgacatt tacctctcgc attcagaag ttattacag gtgataatt cggtaacaa 1560
gtgtcaaaa toaagaaga atag 1584

```

```

<210> 25
<211> 1500
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7479904CB1

```

```

<400> 25
atggatgaga aatccaaaca gctgctgcta gotttgggta tgotcttctt atttgcctg 60
atcgtctctc aatcagtggt ccccgcaaca gaatgcccgc tctctcgcct ccaagcgttc 120
agctcccagg tgcgggaccc gtaaccgtcg gaggatgaga gctccgcag gttctgtccc 180
cgtacaatt tcaaccggg cgaactctcg gcaaggttag acttgcacat caaggggat 240
gaccigatcg tgttctcgca cabcacgaag accgggggca ccaacttccg ccgcaactg 300
tgtcgttaaca tccagctgca gcaaccgtgc gagtcccgcg tgggtcagaa gaaatgcaat 360
tgcacaccgg cgggtaagcg gyaaacctgg ctctctccca ggttctccac gggctggagc 420
tgcgggttgc accgcaactg gaccagctgc accagctgtg tgcctcctg ggtgagagc 480
aagcgcagc ccagcgtgag accgtccaag tggagattt ttcaagattc agatcaaga 540
agtaaggata aacgggttcc tccaacacat aaccgagcgc ccaactctcc gtcactcaaa 600
aagaccggga acacatctaa gagtgggag aacttcaact acatcaaat cctccgagac 660
ccagttccc ggtactttag tgaatggagg catgtccaga gaaagggcaac atggaagca 720
tccctgcatg tctgcatgag aagcctcaac acctccagag agctgcccag ctgctcaact 780
ggcagatgact ggtctagctg ccccctcaaa gatttatgag actgtcccta caactagcc 840
aacaaccgcc aggttgcact gctctccgac ctgacctcgg taggctgcta caactctct 900
gtcagcctg aaaaagcaag aaacaaggtc ctctggaaa gtgccagtc aactctgaag 960
ccatcggcgt tottccgctc cactgagttt cagcaggaac ccaatctct gtttgaana 1020
acctcaaca tgaactttat ttcccaattt accoagata ataccactag ggcctctagt 1080
gtagagatca atgaggaat: tcaaaagcgt attgagggca tgaattttct ggratggag 1140
tggtaacgct atgccaaga cctttttttg cagagttatc agtttatgag ccaagaagag 1200
atcagagag ccaggcaaaa gctcagaaa caacgcaaat tctgaaagg aagctctct 1260
cagaccatt tccagacca ggtcagggc cagagccaga atccgaata gaatcagagt 1320
cagaaccoca atccgaatgc caatcagaac ctgaactcaga atctgatca gaactctact 1380
cagatttga gccagaagya gaaccgggaa agcccgaagc agaacctcag caagagcagc 1440
aatgataca ccagcaatg caccacagac tacataggca gtgtagaaa atggcgitta 1500

```

```

<210> 26
<211> 669
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7480367CB1

```

WO 02/12467

PCT/US01/24382

```

<400> 25
atggcagaga agcccaagct ccactactcc aatgcacggg gcagtatgga gtccattcgg 60
tggctccctgg ctgcagctgg agtagcgttg gaagagaant tctagaaac tgcagaagat 120
ttggcagagt taagaatgga tggagatttg ctgttccagg aagtccaact ggttgagatt 180
gcaggatga agctgggtga gccagagccc attcttaact acattgccag caatacaac 240
ctttatgga aagacatgca ggagagagcc ctgattgata tgaaccocga aggtatagta 300
gattgacttg aaatgatcct tcttctgtcc atatgtcaac cagagyaag agatgccaag 360
actgcccctgg tcaaaagaga aataaaaaat cgtactctcc ctgcccctga aaaagtatta 420
aagagccaca gacaagacta cctgtttggc aacaagctga gctgggctga catcaacctg 480
gtgaaacttt tctactcgt ggaagagctt gactcagtc ttatctccag ctccctctgt 540
ctgaaggccc tgaaaaccag aatcagcaac ctgccaccg tgaagaagtt tctgcagctc 600
ggcagccaga gaaagccctcc catggatgag aaactcttag aagaagcaag gaagatttcc 660
aggtttaa

```

```

<210> 27
<211> 3551
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 8069390CB1

```

```

<400> 27
ggcggcggagc agctccggcg gcgagacggg ggcggcggc cggggctctg gggggaccgg 60
tttgaagac tttccggcc tgcagattgg ccttaagaga agcagggagc cacatactgc 120
tgcggccca gaactggcag agagaaggtt gccatggctg ctgttgacag tttctacctc 180
ttgtacagg aaatgcocag gctctgcaat tgcataatgg aagctctagc tttggttga 240
gcctgggata cggccagaaa aagcatcact gtcactcttg acctttacag cctgacagg 300
ctgcaallta tcccgcgct gggagcaga gcagacttga tcaagcagta tggagatagg 360
gcgttgtca gcggtgcaac agatgggatt ggaagacct acgctgaaga gttagaagag 420
cgaagtcca atataactc gattagtcgg aacgaggaga agttgcaggt tgttgctaaa 480
gcatatgcc acactcaaa agtgaacct gatattatg ttgcgacttc cagcagcgtt 540
cgtgagatc accttccat tgcagaagcc ctgaaggaca aagacgttgg catcttgta 600
aataacgtgg gtgtgttcta tccctaccg cagtatttca ctacgtctc cgaaggcaag 660
ctctgggaca tcaataaagt gaacattgcc gcgctagtt tgatggcca tgttgggta 720
ccgggaatgg tggagagaaa gaaagtgcc atcgtcaga tctctctgg ctccctctgc 780
aaaccactc ctccagctgg tgaatttct gcttcaagg ctbatttga caacttcag 840
agagcctgc aatataasta tgcctcaaa ggaactcttg taagagctc aatcccttc 900
tatgtagcca ccagatgac agcaccacgc aactttctg acaggtctc gtggttggg 960
ccttcgcaa aagctctatg acatcctgt gtttctact ttgggatcc caaaggacc 1020
acagatatt ggtcccattc tattcagttt ctttttgcac agtatatgcc tgaatgctc 1080
tgggtgtggg gagcaaatat tctcaacct tcaactcga aggaagcct atcctgcaca 1140
gcctgagctt ggtggccacc ttggaagtt ttgcaacct ctgggaacct cgtattctg 1200
acatttga aaacacattt aatttatctc ctgtgttcca ttgctgatta ttcagcacc 1260
tgctgattc toatttga aaacacata ataccgtcag agtgcctga aaaacctaa 1320
gggtgtgtgg atggcacagg atcaataatg cctgagcctg attgacgaca tctacattc 1380
agtgctttt cctcaagctg ttgaaagtt acgctttct gtgttctag agcaacgca 1440
gtcaatttt gaataatca atgatttctc agtctctata atttcagat ttgttttta 1500
aggaatgg accatttca ttagagagtt gtgtgtgtt ttaaatgtc atcaagaag 1560
actactgaaa agtattattt tgtaactaag attgctgtga ctattaggaa aaactcgtt 1620
gtattgata gctotagctg ttgactatc tgtaatgaaa atgctgcaat caactggta 1680
ttcaataga gaacgtgtg tgtgctgtg tgtgtgctt ttgagcaact tbatattag 1740
ttcaatatt ttaaaaaga tttttgtca ggtgactta acatggactc ttataggta 1800
ttaaaacat ctgattattt ccttttctc ctaaataago ctaccaaat toatgctgtt 1860
gggttgcac gaatgatctt acctctaca tbatattgt gcttttcaa atcgtctatg 1920
gaatgaactt attcagat ttgtatgtg aaggaaac cgtggtctc tttgattta 1980
taaggcaatt ctgttgata aatagtatt tctcagctc tgaccattt tataactga 2040
attagcctt tttagcttg ttabatctgg ttttctacg ttttctatg taatattat 2100
ccattccgt agcattatt atagaaatg taagtattta tggaaatga aaataggac 2160
aaattacgt tgtgacatat ctgtcaaat aagttagaag ctlattctg gtttgltaa 2220
tgatttatt tattgtagt aataccttta ctggtgtgaa gataattatg cacaacctt 2280

```

WO 02/12467

PCT/US01/24382

```

cacaatacgc gttaacattg aaaccclgga aatgtcctta ggttgggtca tataaagcca 2340
aacatttttg agaaccaatg acctagtgct ttgaaactgt taagtcaacta tatgaatag 2400
acaattatgt cacatttaaa attcaagact cggcaattgt atactgatgc agaacctatg 2460
agattgtgta anngtctgga ctctctgggc attttttctg tgcactgata atctatcata 2520
agcagcccta agcaacagtt latgaacaca attttgcccc tgcactgccc taccggattt 2580
ccagatgtga cttgcaactca gaagatcaagt ggtcaacttc agaatgtott ccacgcttag 2640
atcaatgttt cagaacttag atgtgaaaat ctacaacctg gtagatgtgt tgaagcccca 2700
ggtttttagt gagtttgctt ggaatcctct tgactctatg ccacattgac gtgaactttg 2760
atgalaata agcagcagca actctatgtg aaaaataggt caggtagtta tatgtaagt 2820
tactgtgtcc agtaaatgct tagattgata aattaggtat ggaatccatc agtgttaact 2880
gatgagata gttgaacaca ccttctcagt gatgagttaa actctctccc ttggcaggac 2940
atgggcaaac atgctgattg gtgcaaatgt gttgcagcag tgtccatagc tgcagtgaaa 3000
gatgagagac aagacctctc ctatggtttc tagotttcat taatgtattt ttttcccc 3060
gagctaatit gaaagtgtat tggcaacctg tggatggggt ctcatlaaga atgtgggaaa 3120
tggggcccca gtgcggtggc tcacacctgt aatccacgca gtttgaaagg ccacggcagg 3180
tggatccctt gatcccagga ggtcagagcc agcctgggga acacatctgt tctctacaaa 3240
aaatacaaaa attagccagg caggtgggtg catgctgta gtcccagcta cttgggagcc 3300
tgagccagga gaattttttg agcccagat ccaagagttg aagttagcca agatcgtggc 3360
actgcactcc agccttgaga cagagcgaga cctgtctcca aaaaaaaaaa aagcaactgtg 3420
gaaatagaa ccttctgaag tlaactctgt aattgaaagt tlaacataaa aagtatgtt 3480
tgtttctctt tgaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaatggcg 3540
tcgcaagctt a

```

```

<210> 28
<211> 2178
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7473869CB1

```

```

<400> 28
tcgggtgggg agtagagtag gtaagcgtgt ttctttatgt ggggagggta tgtgtaaccg 60
agaggttttc actgccaatt atggaagtag cctctccttg atattgtlga gttgggatlg 120
tgcaacaaca ttgtactggt gtgggttgyta atttcgctca gatggcataa gtgagaactg 180
aagacttcag tgggatcag aggcacgggt cctgttagtt atgotcttgy gatcaagtgc 240
gctgtctgag ctgaaagga atgtgaktca atctcttcaa cctttasggt gatgagatgt 300
ttgagccagc tgragaatgg acacactatg aagctgtggc tagaaagga ctggtatgt 360
cccatctctc ggccagatgt actggggatg tccagacaga tgctgtcatg ggttagtgagg 420
gcccactctg cacacagacc agtggctgct tgcagttcaa tgcctgtgat ccagagtgtg 480
ttcaaaagtg actctctcgc tctcttgacc tctctctcca ggcagaaag gctgcaagct 540
gctgctatg tgatgcctga gcaacaagcc aaggaaactga actaagtctt tctgttaagt 600
ctgagtttg tcaatggcag gtttacttgt ggcacgctct ctctgacctt gggggttccg 660
ttctctactc ggggacctgt gattgaaaag atctccocgc ggcgcgcgcc gctacctggt 720
gcccaggtgt gcttgcagga gtctgggggc cagctggcct gyalglactg cagcaacggcg 780
gaaagttaca agtggctgag cttaagcag agctgtctgt acctgaaagt gtaagcgccg 840
ggcgccgcgc ccggggatcc ccagctgcca gtgatggtct ggttcccggy aggcacctc 900
atcgtggggc ctgcttctta tcaagagggc tctgaacttg ccoccccgca gaaagtgggt 960
ctggtgtttc tgcagccag gctcggcacc ttgggttccc tgaagcagga cgaagccacc 1020
ggcgccggga actgggggct gctggaccag atggcgctc tgcgctgggt gcagaggaac 1080
atccagacct tcgggggaga cccaggaat gtgacctgt tcggccagtc ggcgggggcc 1140
atgagcaact caggactgat gatgcaccc ctagcctcgg gtctctcca tcgggccaatt 1200
tcccaggtgt gcccgcgctt altcagactt tctcaacta gtaacccaact gaaagtggcc 1260
aagaagtttg cccactlggc ttgatgcaac caaacacaga cacagatcct ggtaaactgc 1320
ctggggccac tatcaaggac caagtgatg cgtgtgtcca acagatgag attctctcaa 1380
ctgaacttcc agagagacc ggaagatatt atctggtcca tgaacctgt ggtgagatgt 1440
gtggtgatcc cagatgaccc ttigtgtctc ctgaccagg gaaagtgttc atctgtgcc 1500
taactcttag gtgtcaacaa cctggaattc aattggctct tgccttatat catgaagttc 1560
cagtaaac ccagggcagat gagaagaa accaactacta agatgtctgt gatgaccocg 1620
accctgttgg tgagggaccc agctggcagg ggtgctcagt tcggacaggy ttgaccoccc 1680
tgtttttttt aaactagtag ctgctctttg caaaggggct cccagccagg gtaaggatct 1740

```

WO 02/12467

PCT/US01/24382

```

ttcttggagg gtctggggtt tgcctgggga tcagatgact gcttacaggt aaggtgctca 1800
gggtcacagag ggcagttatg cagcaaaatc aggggttaca atcagcaagag acagaaactt 1860
tcccaaggag ctcccttttc ccctctccca ggcacaaaac tctggggggs ctgagatggt 1920
atccasgtoa ctgggtggcc ccctctctgg ccagctggca cccagccttc agttaaagtt 1980
ggcctccctca ctccagaatat caccagggag cagtaacca lttccgttga ggaataactg 2040
gacaaatgtoa atgagcaatga ctggaagatg ctacgaaacc gtatgatgga catagttoaa 2100
gatgccaatt tctgtatgca cacaactgag actgctcact acccccgaga tgcgggcttt 2160
cctgtctacc tgtatgaa 2178

```

```

<210> 29
<211> 2081
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Inocyte ID No: 7478588CE1

```

```

<400> 29
gtctgaggct ggtcagaagc tggttacaat tccccccgac ccagtaactg ctggcagggg 60
tttagagccag ataaagttgt gctcacacac tgtagacaag gctacacgac catccacagt 120
gttgccatcc acagttgtgc cacaactcct gcccaacaga gtagctgctt ggaacatgag 180
gttgattctg tctggagacc tcaccctctg cctgatggcc cagacgacct tgggtgcttt 240
gcacacaagc aggcctcaag tggtaacaa atatggaaac ctgcaaggaa aacagatgca 300
tgtgggggag caacccaatc aagtcttttt aggagtcctc ttctccagac ctcccttagg 360
tatctcaagg ttgcaacctc cagaaccccc ggagccctgg aaaggaaatca gagnetgtao 420
caactaaccc cctgggtgac tgcaggagtc ctggggccag ctggcctcga tgtactgacg 480
caacggggaa cgttacaagt ggtctgcttt cagcagggac tgtctgtacc tgaacgtgta 540
cggccggggc cggagccagc gggatcccca gctgcaggtg atggtctggt tccggggagg 600
cgctctaatc gttggcctgt ctctcttgta cgaaggctct gacttggctg cccggagaaa 660
agtgtgtctg gtgtttctgc agcacaagct cgtctctctc ggtctctgta gcaacgacga 720
cagccacgag ccggggaaact gggggctgct ggaacagatg gogctctgct gctgggtgca 780
ggagaaacac gcaagccttg ggggagaccc aggaaatgtg aacctgttcg gccagtcggc 840
gggggcaatg agcaatcag gactgatgat gtcaccccta gccctgggtc tcttccatcg 900
ggcaatttcc cagagtgaga ccgcttattt cagaactttc atcaatgata acccaatgaa 960
agtggccaag aaggtttccc acctggctgg atgcaacca aacagcaaac agatcctggt 1020
aaaactgctg agggcaactat cagggaccac ggtgatgctg gttctcaaca agabgagatt 1080
cttccactg ancttccaga gagaaccgga agagatata tggtaactga cccctgtggt 1140
gaaaggtgtg tgatccagc atgaaccttt ggtctctctg acccagggga aggtttcttc 1200
gttgcctaac ctctctaggtt tcacaacctt ggaattccat ttgctcttgc ctatatacat 1260
gaaagttccg ctaaacccgc agggatgag aaaggaaac atcaactaga tgcctgggag 1320
taccgcaccc ctgttgaata tcaccagga gcaagtaaca ctgtgtgtg aggtactcct 1380
ggacaatgtc aatgagcatg actggaagat gtaacgaaac cgtatgatgg acatagtcca 1440
agatgccact tctgtgatg ccacaactga gactgctaac taccaccgag atgcccgcct 1500
ccctgtctac ctgatgaaat ttgagcaaca cgtctgtgga ataactgta aaccccgcac 1560
tgaatgggca gcccabgggg atgagatgta ctctctcttt gggggccctt tgcacaacgg 1620
cctttccatg gtaagggaga aggaacttag cctccagatg atgaatact gggcaacttt 1680
tgcocgcaca ggaaccccca atgatgggaa tctgcccctg tggcaacctt caacaagga 1740
tgaanaatgc ctgcagctgg attttaccac aagatggggc atgaagctca agggagaaga 1800
gatggctttt tggatgagtc tgtaccagtc tcaagacctt gagaagcaga ggcacttcta 1860
aggttgctta tgcaggaagg agccaagag ggttttgcoc caacactcca ggcctctggg 1920
agactagcca tggacatacc tggggacaag agttctaccc accccaattt agaactgca 1980
gagctccctg ctgcctccag gccaaagcta gagcttttgc ctgttgtgtg ggaactgca 2040
tgccttttcc agcctgaact cccatgatgc cctctactt c 2081

```

```

<210> 30
<211> 2642
<212> DNA
<213> Homo sapiens
<220>
<221> misc_feature

```

WO 02/12467

PCT/US01/24382

<223> Incyte ID No: 55046125CB1

```

<400> 30
ggagcggccc gggctgggtc cgcgcctcag tgcgcggggc cctcctaggg gtgtgtctca 60
cggaltccac acccggccgc tctgggaaa gccccgaag gcgtctcttt ccttggcggg 120
accgcgctgc gccaccgctt tgcgctgct gtcccggggc cgcgcgaggc ggtatgcagt 180
cctcaggcga cgttggggacc tggctccct ctgcggggcc ctgtctactc ggggctctgc 240
cgccttgggc cactcgtgta agcactgtct cgttgcgata ttctccaaga ttttggccc 300
catggccaag gtcgggaaac tggatgagaa atccaacaag ctgctgctag ctttgggtat 360
gctcttccca tttgcccgtg tctctccca atactgtgtc cccggccagc aatggccagc 420
ctcgcgcctg tctgtgccc gctacacttt ccccgccggc gacctcttgc gcaaggtaga 480
cttggacttc aagggggatg acctgatctg gtctctggcc atccagaaga ccgggggaaa 540
cactttgggc cgcacttgg tgcctaacat cccagctggg cagcgcctgc agtgcctgct 600
gggtcagaag aaatgcactt gccaccggcc gggtaagcgg gaacctggcc tctctccag 720
gttctccacg ggcctgggctt ggcgggttgc cgcgcactgg accagctca ccagctgtgt 780
gccctccggt gtggacggca agcgcgacgc caggtgaga ccgtccagga acttccaata 840
catacaccac ctccgagacc cagtgtcccg gtacttgagt gagtggaggc atgtccagag 900
agggcaaca tggaaagcat ccttgcctgt ctgcgatgga aggcctccaa cctccgaga 960
gtgcccagc tgetacactg gcgatgactg gtctggctgc cccctcaag agtllatgga 1020
ctgtctctac aatctagcca acacccggca ggttgcctgt ctctccgacc tgaacctggg 1080
aggctgtcac aacctctctg tcatgctgta aaagcaaaag aaccaagctc tcttgggaaa 1140
tgcacaatca aatctgaagc acatggcttt ctctggctcc actgagtttc agcggagaac 1200
ccaatctctg tttgagaana ccttcaacat gaactttatt tccgacttta cccagtataa 1260
taccactagg gcccttagtg tagagatcaa tgaaggaaat caaaaagcta ttgagggaat 1320
gaatttctgt gatattggat tgtacagcta tgccaaagac ctttttttgc agaggtatca 1380
gtttatgagg cagaagaagc atcagaagcc cagcgaagag cgtcaggaac aacgcaaat 1440
tctgaaggga aggcctcctc agacccttt ccaagaccag ggtcaggccc agagccagaa 1500
tccgaatcag aatcagagtc agaacccaaa tccgaaatgc aatcagaacc tgaactagaa 1560
tctgatgcag aatctgactc agagttttag ccagaaagag aaccgggaaa gcccgagcca 1620
gaactcaggc aaggagcaga atgataaac cagcaatgpc accaagact acatagggag 1680
tgtagagaaa tggcgttaaa tggctcaaaa agcctgtcac ataactctc caaagcgcga 1740
ctgaaagatg ggcataagct aaaaatgaaa agtctccaaa cacactctgc ttcctcatt 1800
ggggaagttt taaaaaaaag tttgatgtt gcccttccag ttgcctttca atcagttgt 1860
atactgtgtg taggtaaaac aaatctcaat atggnaataa atgtctctt tgggggtgga 1920
ctaaataga aatccgaaag ccaaaccaaga ctcaaccaga atgtctgttt agatatttta 1980
agaagttctt aaattagtta tggagcaaaa gtgaaacat aaaaatgtac catttaact 2040
atggttaaga aatggacttt aaatlalca tgalacactg ttaaaacca atctlggaat 2100
caaatatttt ttccaggggt ggaataatgt ataacataa agaaactaaa atgaaacta 2160
aaacttttta tttctctctg atttbaaaa ggaatctatt taatagaaat aacaactgat 2220
ggtgaactct accgagctgt agaaaataaa aaatctctc ccaaacctgg gtatgtttat 2280
gtccaatat tggcttttca agaacaggac tcaatctctg atatttaaga gatgtttaa 2340
atlttaaac tttctacact tctactgttt aaaggtttta ccaaggggtg atctcacatt 2400
aaacaaaaca ctttttttcc aaaaataaat accaatgtaa agatctaat tccagggcct 2460
ttcagggcac tgaatttcca ccaactctgg aatcattttg gcctgctctc tcaatcatt 2520
taaggctctc ctgaattgtg ctcaatccaa attaaccat gatatagatc tttctcatt 2580
atcaaatggg tgtgtgtctg aagttatgt ggtataaat cctggatata agtcaggact 2640
tt 2642

```

<210> 31

<211> 2080

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3538709CB1

<400> 31

```

tgttttggcg gccgcgggat gccctgcgc tgaaccgcaag gggcaggtgc ccgcccggct 60
agaagcaacc ggcctgaccc cgcgcaacca tgltaaacagc gccagcaygc ggaagctgpc 120
ttctccgctc gggaccctct cgcgccgacc cgggcccggc ggccctcgat gaggcaaac 180

```

WO 02/12467

PCT/US01/24382

```

catgctgacc ggggtgacc acggtatctt ctggtccctg ctgggcaacc ccccaaccg 240
cvtggggcca ctggagagcg tcyagtccag cgatggctac acctttagt aggtcaagcc 300
cggccgctgt ctgcygggtga agcatgcaag acccgcccca gccgctgccc cacctcaacc 360
atcatccgca tctctggatg cagcccaggg ggaacctctc ggtttggtcc gctgtcagcg 420
cggatcaacc gtagtaccga atggcggttt gctggtagaa aacctggacc gagccctcg 480
agccgaacct ctacaacggc agaatggctc tggggagccg ccggccgccc tggaggtgga 540
gctggcaagt ccggccggca gccatggccg ctgtgcccc ggcagcgcag gcagcggcag 600
ggcgaatggc agtggttggc ggccggggcg agccaggcgc cccaagagga caatccatct 660
tgaactgtag aagcgcataa ctagtgcata agccgcccag gccagctggy tgcctctttt 720
catccatggt gtccgctgtt cctctgccc ctggaaggag cagctggact tctttgtgcg 780
cctaggctat gaggtygtgg ctctgacct ggcggccacc gggccagct ctgccccca 840
ggtggccgca gctacaacct tctatggcct ggtgaggacc atggcagca tottaagcg 900
ctatgccaag aagcgaatg tgcctatgy ccattccctac ggtgtctctt tctgcaall 960
cttgcacat ggtaccacg acctagtga caggtgata atgataatg gggggggccc 1020
taacgctgtg gagcccagct tctgctcaat cttaacatg ccaacctggc tctgcaactg 1080
cttgtgcacc tgcctggcct gggacttctc caagcccgcc ttccgcccgc aaggagcca 1140
ggagaagcag ctgttaagg aggcacaagc tttaacgtg tcatctctg tactccggcg 1200
catgatgagc ggcagctact ggcggaggg ccagcaggtc taccagccg agctcaaccg 1260
gccgctcctg ctgtcccagc gcatgcagca taagtttgt ccggtggagg aagaccagcg 1320
catggccagc atctctgccc tggcaatctc gaagctcact gacagggcca gccacatggt 1380
gatgctgaaa tgcctgaga cggccaacac gctgtccac gaattccctg tctgggacc 1440
cagccctctg cccaaggtc tacccggacc actgcccggc cctccagaa accagagta 1500
gccctgggca cggcgggcca tgccttggg agccaccgag ccagcagagc agcccgagc 1560
ctgcccagc tctgcaagc agaccacctg ggcggccctg tgcctccgtt gggcggggcc 1620
agttcaaggc gaccccccca ggtgcctgg cggggcgctg gcatccgagc gagcccagc 1680
gacattccgc tctccgttc cgtcccgccg gcccactcgg cgttttgggg ccgacgcccg 1740
gacctcaagc gaagatgacc ttgtacagaa gctctccct accctccccc caacgcccag 1800
gcccaagcag gcccccaacc ccgctgtctt ccgtgtcagc cgtgcttct cctgggacc 1860
acgaccacca cagggacctc cagggcccca tccggtatc ccagacctt cctacccccc 1920
atctctggc gctgggagct atttttgcc aagggggggg gatggggggg ctggccacca 1980
cgaactgca catctcaact tctaactcaa taacagaaag tgaacaatgy aaaaaaaaa 2040
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2080

```

```

<210> 32
<211> 2219
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Inocyte ID No: 71563101CB1

```

```

<400> 32
ggcctccaat ctgcttccat gggggttggc tttctgagtg ggagaaatga ctctaactg 60
gagacatttg ctgagacctt tgtgctgggt cacttccctg cccaggtacc ttgagatga 120
tcttctcctg agcctaggtt ctctccggac atcagtaaac aagctcagtc ttcatacaa 180
gcccagaatg cctccatggt acttcaatgc tgaagatacc cagctccctg gctacaaccg 240
tctctgaaa atccacaagc aacatcttcc tctctggttg acggcatatc tccgaacc 300
cctgtgcttc aaccagggc acctgagtg gctcttctg gctgagaaa accagatacc 360
ggaattcttt tccgggcttg ttactgtcag tgttggcaac tgcacccaa aggtgaaagc 420
agtggcaaaa aagcagctcg gccgctgtg gcatcaagc accgtctctt tccacctcc 480
aatgcatgaa tatgcaagaa agcttgcgc acttctctc gagcctotta aggtcattt 540
cttggtagac agtggctcag aagccaatga gctggccatg ctgatggcca gggcgcact 600
aaacaacata gacatcattt ctctcaagag agcctaccat ggatgcagtc ctacacact 660
tggcttgaca aacgtaggga tctcaagat ggaactccct ggtgggacag gttgccaacc 720
aaacaatggt ccagatgttt ttctggccc ttggggagga agccactgct gagattctcc 780
agtccaaca atcaggaagt gcaactgtgc accagactgc tgcacaagta aagatcagta 840
tattgaaa ttcaagata cgttggacc atctgtggcc aagtcactg ctgattttt 900
gcacacacct atcaaggtg tgaatgagc tctccagac ccaaggggt tctcaagaa 960
agcctttgag ctggtccag caagggagc cgtgtgact gccagtagc tgcagacag 1020
atltggaag ttgggctctc acttctgggg ctcccaacc cagatgtcc tgcctgact 1080
tgcacactg gctaaaagga ttgggaatgg ctctccactg gcagcagta taaccactcc 1140

```

WO 02/12467

PCT/US01/24382

```

agagattgcc aaatcttggc cgaaatgcct gcagcacttc aacacctttg gagggaaacc 1200
catggcctgt gccattggat ctgctgtgct tgaagtgatt aaagaagaaa atctacaagg 1260
aaacagtcga gaagttggga cctcacatgtt actaaagttt gctaaagctgc gggatgaatt 1320
tgaatttgtt ggagactgcc gagcaaaagg cctcaatgata ggaatagaaa tggtagaaga 1380
taagataagc tgtcggcctc ttcccctgta agaagtaaat cagatccatg agpactgcaa 1440
gcacatggga ctctctgttg gcagagcgag cattttttot cagacatctc gcattggccc 1500
ctcaatgtgc atcactaaac cagaagttga ttttgcagta gaagtatttc gttctgcctt 1560
aacccaacac atggaagaaa gagctaagta acattgtcag aaataaata aaccacaagt 1620
ctcaagaatt tgccacgtat gttcaagggg gaatttgaag aatctcagaa ccactgggat 1680
ccagaaaaag cctgcagctc tcccagggag ctgtaaaagt catggttgac tgcctaccaa 1740
ccatatttgt tagcagagcc cctcttatct tgagaactcc attcttcagg gaaaggatct 1800
ccctagctca gagaataaat cctcaatagt ttatgttagg tatggtaatt tgattccctc 1860
ttgcagtgat tggttatgct atgatatgtt gatgtatttt tttccagtga atcttgaaga 1920
aaaatctttt ggtggaggtc ccttcaggga aagttttctt caecctcact ctccagttca 1980
agaaagatgt tctctctgtt gcgctgagaa caccatattt tcatgacagc attcctggca 2040
ccatgtaccg cggctgttag tcatgaggac aacccttttt ggtgaggttg gaagatggat 2100
ggagccaagc tgccttagta tgtcaagaaa gcactcaatt aagcattcct ggagccccc 2160
tacctcaggg cctcttgata tttgaggtaa taaaattcat tgttctgtat aaaaaaaaa 2219

```

<210> 33

<211> 681

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7472027CB1

<400> 33

```

atgagatcat gtgaaagacc cgaactacag ttgattggly taccgaaag tgcagggag 60
aatggaaaca agttggaaaa cacatttcag gatattatcc aggaactctt ccccaacctc 120
cgaagacggg ccaacattca aattcagatg gctggtggat ctatctggat tgaagggatt 180
ccttttcccc gcaataattt tcaagacctg aqactgttgc angatgaat tgtgttggc 240
gatgaagatg tcaattacat ttcttaccoc aagtcaggaa gcttttggat agtggaaatc 300
atcagttotg tccactccaa gggagatcct agttgggtcc aatcttggat tccctgggat 360
cgttcaccat ggatagaagt taaacgtaag aaagcaggtt tagagagtca gaagggccca 420
cactctcaca cctcccacct tcccattcag ctcttcccca agtcattott gaattccaag 480
gccaagtgtc tttatcctca tttctctcatg ctgttggttc tcatctcagg acataagagc 540
cagttggagta ttgatataaa gatctctgaa aatgcagag caactccaaa actgggaat 600
ggcaaaagat tggaaaggtt tggagggctc agaagaagac aggaagatga aagaagttg 660
gaatttttta gagactgtta a

```

<210> 34

<211> 399

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7480358CB1

<400> 34

```

agcccagacc agagtgtggc cctcggagtc ataagtatga aggtcatgta catgttgaag 60
ggccagagcc cggtgacggg caccatccac tttgagcaga aggaaaatga accatttatg 120
gtgtcagaat gcattacagg attgactgaa cgcagacaca gattccatgt tcatcagttt 180
ggagataata caccagpctg taccagggca gttccttact ttaatccttt aaccaaaacc 240
cacagtgggc caagatcaaa gagagggagg ttggagactt gggaaatgg gcgctggcca 300
aagatgtgtc gccacatgct tgttcaagat tctctggctc cactctcagg acactattcc 360
atcaatgccc acaaatggtt gtccatgaca aacagatga

```

<210> 35

<211> 2302

WO 02/12467

PCT/US01/24382

<212> DNA
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <223> Incyte ID No: 1618256CB1

<400> 35
 taactaatcg gaaggcgtg tgaattcac gttctcggg tagctgtaac gcgcggtata 60
 gttcttaactc aagtcgtgccc ctgccagaaa ggggaagatg caagcgggga gaggcctaga 120
 accgccatca aagctgcagc agctcctggg gctggcgcca caaagggagg agggaggagc 180
 ctggpccctg agaaggttct tgggaaagt tggctgagc caapagtgg gcggtggctg 240
 ggaagggggg gaggccggg ccagatttcc ctgtgtgct ccttggggc ctgagctgga 300
 cagaggcaaa aatctgctaa ctcaaggggc agactcaac aagcctgga gcaagcctgg 360
 gaaatgacc ccgatctcc aaccagtccc ttcccagct gcacggctgt ctccagctgt 420
 ctctgccctt cttctcggcc ctggctccat ctctctgta cctcaacctt cctctgcca 480
 catgggccc ctctctctct caggagcct ggggcttgg ggaacctgga gctctggggt 540
 gctctggga gtcttcaatga caattggctt tgcactccag ctctggggag ggccttcca 600
 gaggaggcta cctgggctac agctccgaca gccctcggcc ccatccctac gaccagccct 660
 tccgtccctg ccaccccggc agcgaactgt gttctctgaa acacataat ccgggagcag 720
 ctctgtgctg agctgtctc accgtcttgg ggaaccagac gggctgagct tgcacctcc 780
 tgcctgctac cagtttggct accccaagct ctccaggcc tctagggtaa aaggtaccgc 840
 ccacccgggt gaggccccc agctcccctt ccaactctc tgtaccoca tgaggttcaa 900
 cctgnaagag gtactctcag tcatgcttc tgacagcttc tttttttcca ttgtccgaga 960
 ccacagcct ctggctcct ctgctctct ctactataa tccacctcat cagccttccg 1020
 caagtcaaca tcttggctct ccttctggc caactctga ggccttaca ggcctggggc 1080
 cclyggggac caclacgctc gcaacttact atggtttggc ttgggctgcc cctttcccc 1140
 agagaagagg gccaaagag ggaatttca tccccccaga gaccccaacc ccccaagct 1200
 gcaagtcttg cctctgggt ctggccctcg agcccgaacc ctcaatccca atgcccctat 1260
 cctctctgtt tccactgta ctgatactg cagccagaba tcaagccctg cctcttcca 1320
 ttggggctct taactctca tccagtgagg tctggctgga ctggactctg tctttgact 1380
 ggtctatgtg gctgagtact tccatgagtc atgtgtctg ctggcagatg cctgtgctg 1440
 gggctatgat gactgtgtgg gcttctgca caatcccag gctggcaata agcaggccct 1500
 cagcaactgc agcaacagtg gactgactgc ggaagaccgg cagctyaacty caggycccgc 1560
 agcotggaac aacctggact gggctctcta tgtccacttc aacgcagtc tctgggcaag 1620
 gatagagaaa taaggccagg gccggctgca gaccagctg gccagactcc gggctgccc 1680
 agaggcccta gcgaaacatt gctgtgtagg ggggtgaggt tctgacccca aatacatca 1740
 tpatccctg tccgcccct tccagtttgg tccagctaa gttttggct abatacttg 1800
 gagtgagtg agccccag acaagagga atgtgagcgc ctgactacc ctgagctca 1860
 gtacagggac aagctggatg tcaagcagtt ccccctaac gctcactgc cctcaagac 1920
 ttcaaggcca ctctcccctt aaacatcaga ctacagattt aggtggaaga gcaagccatg 1980
 ttgaagggca catgtgatga gtggggggca gcaagatgcc atttctgcat ctcccagaag 2040
 gyatgagct tgtcccgat gcaagcccc tctctgctgg gctcccagca gtgcttccct 2100
 cctccacct ccaactcttt tgtctttcc ccccaacttt ttttttttg aaacygagtc 2160
 ttgctctgt ccccagcctg gagtgcagtg gcaatgactc ggcactgac aacctctgct 2220
 tcccaggttc aagcattct cctgctcag ctccagagt agctaggatt acagatagct 2280
 gcaaccatc ccggttaatt tt 2302

<210> 36
 <211> 3341
 <212> DNA
 <213> Homo sapiens
 <220>
 <221> misc_feature
 <223> Incyte ID No: 3387823CB1

<400> 36
 gaaacgaaga gaaacctctg tccctggcag aatctgcatg tacatttttt gtctgtcttt 60
 gctctctctc tctctgctg gcccaattgca gagagatttg gaagtttcca accttgggt 120
 gtaactctag ctcaactcac ctcccttggc gaaagacagt gtygcagcgc cctlyctgtc 180
 tgcactaac cctgactgtg ttggcatgag cccctctgca accagcaaaa accggcgggc 240

WO 02/12467

PCT/US01/24382

```

tgcacatggc aatgccagcc ctgggaagct ggagcactcc aaggctgcc tctccgtgca 300
cgttccaggc atgaaaccgat acttccagcc ttctaccag cccaatgagt gtggcaaaagc 360
cctctgtgtg agccagagat tgatggaact ggatgagctc tatgagtttc cagagtatttc 420
cggagacccc accagtacc tggctttgag aaacctcact ctggcactgt ggtatcctaa 480
ctgcaaaaga gctcttactc ctcaagaatg tattctcaac abctctgtcc ggggtctcgt 540
gcgatttcca tgccttcagg aagtggagag aatcctgtat ttatgacca gaaaaggtct 600
catcaacact ggagttctca gcctgggagc cgcaccagat attctnctta aggactaoca 660
caataaatac gtcacatcta tcgggggtgg tccagcagga tttagcagct ctaggccaact 720
gcataacttt ggaattaagg tgactgtcct ggaagccaaa gacagaattg gaggccagat 780
ctgggatgat aaatctttta aagccttacc agtgggaaga ggagctcaga ttgtcaatgg 840
gtgtattaac aaccagtag cattaatgtg tgaacaactt ggcctcagca tgcataaatt 900
tggagaaga tgtagcttaa ttccaggaagg tggagaata actgacocca ctattgcaaa 960
gscatggat ttcaattha atgctctctt gpatgtgtc tctgagtgga gaaagataa 1020
gactcagctc caagatgtcc ctlttaggaga aaagatagaa gaaatttaca aggcatttt 1080
taaggatctt ggtatccaat tcaagttagt ggagggcag gtctctcagt tccatctcag 1140
taacctggag taecctgtgt gcagcaactc tcaaccagta tctgtctcgt cgtgggacca 1200
caatgaatic ttggccagat ttgctgttga ccaactctct ctaactcccg ggtactcgtt 1260
gataattgaa aaactggcag aagggcttga cattcaactc aaatctccag tgcagtgtat 1320
tgatattctt ggagatgaag tgcaggttac cactacagat ggcacagggg atctctgaca 1380
aaaggtatta gtcactgtac cactggcttt actacagaaa ggtgccatto agtttaatc 1440
accgttgtca gagaagaaga tgaaggctat caaccagctc ggcagcagca tcaattgaaa 1500
gatgtctctg caatttccgt atagattttg gacagtaaa gtcacagggg ctgacttttt 1560
tggtaoagtt cctccaggtg ccaagcaagc agggcttttt gcccgtcttt atgacatgga 1620
tcccagaag aagccagcag tgcctatgtc tgtgattgcc ggggagctgt tccatctcgt 1680
gagccaccctg gacgcaaac aggtgtctca ccaagtctat gccacgtcc gggagctgtt 1740
ccaggagcag gaggctccag atcccacaaa gtattttgtc actcgttggg gccacagccc 1800
atgatccag atggcctaca gttttgtgaa gacaggtgga agtggggagg cctacagatc 1860
caltgtgaa gacaltcaag gaaccttctt ttctcgtggt gaggcaacaa accagcattt 1920
cccacaact gttaccgggg catatttgag tggcgttca gaaagcaagca agattgagc 1980
attttaaaga tccggtggac ccagctttct tctgtacccc agatggggaa atttgaatca 2040
catgttaaac ctcagtttta taagaggggg aaaaaacct cctacacag taacaclgaa 2100
atgtttctaa ggcgatatga taatgcaaac ctatttcaac acctbaaaag cactgacctc 2160
aaaaaacctt atagcaactt agatttaatt gcaatttcca taggttcaac tactgtgtaa 2220
agctcgtgatt tcaaatanaa gcagaatgta agtttcaagt gaggccatgg atttgaattg 2280
tcaatggctg gaagttccct tttagattca cattttatct ggctgataaa ttctcaca 2340
ttgagaacc aagtcaatca agcaggaatc atttaaaac cagataaagc catgttttta 2400
ttctgtgaca atttatcagt atctttacca atagacctta atttttatct aggtccaata 2460
ttgagctttt acttaaaatt tagatagaac ttttttttgg atacagcaca aactccagtt 2520
gacagtaaaa tgaagctctt agtlatittg tattgtacat atttccctct actgggtgtt 2580
caaaagaact ttaattcnaa gtaacctttg tgatanaatg ttttagattt gtcacccatc 2640
tggcaaaaca gaaaagtttc cagataggtt tigtatcatt gagaatgag caagatagt 2700
gtgggctcca ractatagac acagaatata gcttttctt aagcccaat ttgggtgata 2760
ggcaacttta aatatcctta attttggcaa ccaactgcaa aaaaaacttg tcaqaataat 2820
taaccaagc cctctccac ttcttttatt taagaacact gattcaattg ctaggaaatc 2880
ttttgcagat tttctttac agtattccat aggcaggtcc actggaaaac tgcagaaaaa 2940
tgtgagctct cctggtaaat agtatcaatt ttataagcta tabtttaag ccaataagaa 3000
atggcagata ttactttta tctttttttt aaaaaacccc atgacagaaa ccagctcaat 3060
aatatctcat tctaaataa aaactgggtt gcaaggtctt cagatgctt atltgcccag 3120
aaactcagat atcaaggtta gaataatgct tttgttttgg aacataatct actggtttgg 3180
tttaagaag ggaactaata tgaacttaaa gagaactcna atattgata ttttaaaat 3240
ttaaaagtag gtoagtttat aacagtaaa taccataacc ccaagaatgt gcagtgaacc 3300
tcaggcgggg atccttagtt ctaacggcag cccctagggc g 3341

```

```

<210> 37
<211> 777
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 55142051CE1

```

WO 02/12467

PCT/US01/24382

```

<400> 37
atgtccctctg ccaattgcaatt ggccttccctg ccaactggctgg taacattgctt ggtgctgctac 60
ggccactact tccgafctgtt ggtgcgcacg gctcttcttc gaagcctccg agactgcctg 120
tcaggcctgc gpatcggagg gctggccttc agctaacgtgc taccocctgc cctgcccgtg 180
gacctgtgtc acatcctcac caccctggac cactggagca gccctgcca gtaactgagc 240
cacatggggc ctgtcaaaag tcagatcctg atcgctgggg tggaggagaa ggcctctgct 300
tgtgtgctgg aattgggaaac ctactgttga tactotaccc tgcctattgc ccgagccctg 360
ccccctgggg gtcgctctct tactgtggag cgggacccac gcaacgcagc agtggctgaa 420
aaactcatcc gccctggccgg ctttggatgag cacatggctg agctcatcgt gggcagctca 480
gaggaactga tcccgtgccc acgcaaccag tatcagctga gtccggcaga cctggctcct 540
ctggcaacacc ggcacacgatg ttaacctgag gacctgcaag tgcctggagg ccaatgacctg 600
ctgcacagcag gtcacaccctg gctggctgac catgtgctct tccctggctg acccggcttc 660
ttgcagtatg ctaagagctg tggcctgtac cgtgtccgcc tccaccacac tggccttcca 720
gacttccctg caatcaagga tggaaatgct cagctaacct atgctggacc aggtgta 777

```

```

<210> 38
<211> 3600
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> Incycle ID No: 7395274CB1

```

```

<400> 38
gagagcagag caggacactg gcgcccgagg tcaaggcagct gcgtgcgctt ctcctccagg 60
cagcaagggg aaccocagag cgcgcgcgcc cggacacatg cgtctccagg gccctgcag 120
cgcgcggccg aggacccccc ctggcccggc gcctcctctg gtgcctctc gggctgctg 180
cgcctggacc ccagcggggg cgcctgtgtg ctatgcggcc tctgtgcctc gctgggctgg 240
agctggctgc gggggggccc gggcggagcc atcccggcc ggcaccagcc ctgacctgtg 300
gtggcaactc tgggtcaact gctgtgctct ccttctctcc gggcggggag ctggctgagc 360
agcaggaacc gggcccagag gattgatccc tctgtatag gcccgaggtt gctcctggct 420
caactagccc gcgtgtacgg cagcatcttc agctcttcta tccgcaacta cctggctggc 480
gtcctcagcg acttccacag cgtgcgcagc gcctctgtgc agcagccaga ggtctctcagc 540
gaccgccggc ggggtccgct catctccatc gtcaccaagg agaaaggggt tgtgtttgca 600
cattatgctc cgtctggag acaaccaagg aagttctctc attcaactct tcttcaattt 660
gggttgggaa aaacttagctt ggagcccagc attattgagg agttcaata tgtgaaagca 720
gaaatgcaaa agcaccggag agacccttc tgcctcttct ccaicctcag caatgctctc 780
tctaaactca ttgtctctct gtgctttggc cagcgtttgt attaacctaa taagttagttc 840
aagaaaatgc ttgttttatc gtcacagggc ctagaactct gctgaaacag tcaagtctcc 900
ctggtcaaca tatgcccctg gctttattac ctctccttgg gaccatttaa ggaattaa 960
caaatggaaa aggataatac cagtttctct aaaaaactca tcaagacca tcaagatctc 1020
ctggatagag agaaccctca ggaactccta gacatgtacc ttctccactt ggaagaggag 1080
aggaaaaata atagtaaacg cagtttggat gaaggtactc tatttatctc cattggggat 1140
ctctttattg ctgggaactga taccacnaact aactctttgc tctggtgctc gctgtatctg 1200
tcgtgcaacc ccgatgtaca agaaaaggtt calgaagaaa ttgaaagagt cattgggccc 1260
aacggagctc ctctccctac agcaagggcc cagatgctct acacagagc caaccactg 1320
gaagtgcaga ggcataactgt ggtgggtccc ctgtcaactc ctcaatgac ctccagagac 1380
acagtgctcc aaggtatcac cactctcaaa ggcacattga tcttaccaca cctgtgtcta 1440
gtacatagag acccagccat ttgggagaaa cccgagagatt totaccctaa tctattctgt 1500
gatgaccaa gacaactaat taaaaagaaa acccttattc cttttgggat agggagagcg 1560
gtgtgtatgg gagaacaact ggcacaagat gaattattcc taatgtttgt gagcctaagt 1620
cagagtttgg caittgtctt acctgagatc totaagaagc cctcctgac tggagatctt 1680
ggtctaacctc tagcccccac tccatttaat ataacatttt caaggagatg aagagatctc 1740
caaggaagag atggtaaaaa gatatataaa tacatattctc totaagcaga ttcttctcac 1800
tgcgaagac agtgaactca gcaactcagt ggaaccaagc tggctcaga ggtcggagag 1860
aggttagac accctggagc gtttcaactc gtagaattcc ttagcagatc acttccagca 1920
ttttagtaat gcaagttctgt gatttggggg atagaaaaca aagtaactat gaaacgggat 1980
atctggattt tacttgaagt gcttccaac gatgggcaac tottctcatt tottagtgc 2040
tcagacatcc catatgtaaa atgagagtaa taaaacttgg ctctctctca cctctcagca 2100
ctaagatgg tcaaatgctc taactctttt ctgatatctc taaaatgctg taaagtctg 2160
gagaagaact tcaggagaag aagatctatc agctggcttt taaagacnta tgaaccaatg 2220

```

WO 02/12467

PCT/US01/24382

```

aaagtgggtg tcagcctgga atgctttgtc agagatgggt gtggatttag gttatactgg 2280
gggagaactt ttctcagcac agattctatg ccagcttctt tgggcttggc ctgtcactat 2340
ctttttgttt atgattttag tttttacttc ttgttagatg gggatgaagt ggactctgtc 2400
ggtctacttg aggaaaaag aaattatac tttaaaaaat ccttgtggg attattatct 2460
aaatttatat gtctaaactc taataaactc acaggaacng tgagccttgc tacttcttta 2520
gtactttctt ggcagaahtc cttctactg agtllattgc aaagatgcag ctctaaccttt 2580
ttacttaagg cctgaatggt gacatgggg attttgatac tgggactcat caggaaagga 2640
ttctgtcttc aaactatact gaacattcct gtcactagct ccttgccacc agggcccaatg 2700
catctgatcc ttgaatatac tctcaagaa ttcactctct ttttattaag agaactaaat 2760
tgtttctaaa tglagatggt cctctggaa aagcagtttt cagcaggggt ggtaacccct 2820
tcagaggag tttggaaatg tgtgggtatg attottggtt atcataatga tgggggtgct 2880
actggccttc tgotgcaatg ggaccaggat gctaaatgtc aaggtagtcc tatacagtga 2940
agaatgtccc tgotcaagat gccaggattc ccccagtga gaacatgctc taaggatga 3000
ccaccctttt cttttattct cccacagtgc tccatgtaca gaagtaagca tagcagtcct 3060
atgagcaacc acattctctg acccttctct atgtctgctc taacttaat cctttacttg 3120
tatgtttctg taattcttac ataaattcta ttaagagggt ggcatactgt agtggatgaa 3180
gotgaggctt atagtaggta aggcacaaag ttaaaagta acatcactgg gtttcaaac 3240
tactgtcttc tglactaaa gaacacttcc agaaccactt ctgtactctg ccaccacttg 3300
atcccataac aggtaccccc ttggcctcat gotggagttg tgtgtctctg tottcaatcc 3360
aggctgagct ccttggagtg aggatgttgt gctgtttgccc tccctcacag tgccttggto 3420
ttagtggatg cccagttctc ttgtgaatga cttttangaa gtgtacttaa gagaanaatc 3480
ctacttact tgaataatca caagtcatgt ttttgtgct taaagtgat aeatcagtg 3540
atattatttg ttaatgtcon ttaagccag tttttaaaa aaaaaaana aaaaaaana 3600

```

【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
14 February 2002 (14.02.2002)

PCT

(10) International Publication Number
WO 02/012467 A3(51) International Patent Classification: C12N 15/53,
15/54, 9/02, 9/04, 9/06, 9/10, 9/14, 9/18, C07K 16/40,
C12Q 1/68, G01N 33/68, A61K 67/027, A61K 38/44,
38/45

(21) International Application Number: PCT/US01/24382

(22) International Filing Date: 3 August 2001 (03.08.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/223,055 4 August 2000 (04.08.2000) US
60/224,728 11 August 2000 (11.08.2000) US
60/226,440 18 August 2000 (18.08.2000) US
60/228,067 24 August 2000 (24.08.2000) US
60/230,063 31 August 2000 (31.08.2000) US
60/232,244 13 September 2000 (13.09.2000) US
60/234,269 20 September 2000 (20.09.2000) US(71) Applicant (for all designated States except US): INCYTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(72) Inventors: and

(75) Inventors/Applicants (for US only): BAUGHN, Mariah,
R. [US/US]; 14244 Santiago Road, San Leandro, CA
94577 (US). BRUNS, Christopher, M. [US/US]; 575 S.
Rengstorff Avenue #126, Mountain View, CA 94040 (US).
DAS, Debopriya [IN/US]; 1179 Bonita Avenue, Apt. 3,
Mountain View, CA 94040 (US). DELEGEANE, Angelo,
M. [US/US]; 594 Angus Drive, Milpitas, CA 95035 (US).
DING, LI [CN/US]; 3353 Alamo Street, #146, Palo Alto,
CA 94306 (US). ELLIOTT, Vicki, S. [US/US]; 3770 Politan
Place Way, San Jose, CA 95121 (US). GANDHI, Ameena,
R. [US/US]; 837 Robie Avenue, #1, Menlo Park, CA 94025
(US). GRIFFIN, Jennifer, A. [US/US]; 33691 Mello Way
#11, Fremont, CA 94555 (US). HAFALIA, April, J.,
A. [US/US]; 2227 Calle de Primavera, Santa Clara, CA
95054 (US). KHAN, Farrah, A. [IN/US]; 3617 Central
Road #102, Glenview, Illinois 60025 (US). LAL, Preeti
[IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US).
LEE, Sally [US/US]; 825 East Evelyn, #425, Sunnyvale,
CA 94086 (US). LU, Dyung, Aina, M. [US/US]; 233 Coy
Drive, San Jose, CA 95123 (US). LU, Yan [CN/US]; 3885
Corina Way, Palo Alto, CA 94303 (US). PATTERSON,Chandra [US/US]; 490 Sherwood Way #1, Menlo Park,
CA 94025 (US). RAMKUMAR, Jayalaxmi [IN/US];
34359 Maybird Circle, Fremont, CA 94555 (US). RING,
Huijun, Z. [US/US]; 625 Orange Avenue, Los Altos, CA
94022 (US). SANJANWALA, Madhu, S. [US/US]; 210
Sylvia Court, Los Altos, CA 94024 (US). TANG, Yi, Tom
[US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US).
THANGAVELU, Kavitha [IN/US]; 1950 Montecito
Avenue 23, Mountain View, CA 94043 (US). THORN-
TON, Michael [US/US]; 9 Medway Road, Woodside,
CA 94062 (US). TRIBOULEY, Catherine, M. [FR/US];
1121 Tennessee Street, #5, San Francisco, CA 94107 (US).
WALIA, Narinder, K. [US/US]; 890 Davis Street #205,
San Leandro, CA 94577 (US). WARREN, Bridget, A.
[US/US]; 10130 Parwood Drive #2, Cupertino, CA 95014
(US). YANG, Junming [CN/US]; 1739 Bank Lane, San
Jose, CA 95129 (US). YAO, Monique, G. [US/US]; 111
Frederick Court, Mountain View, CA 94043 (US). YUE,
Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087
(US).(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AH, AG, AI, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GI, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK,
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,
MZ, NO, NZ, PL, PT, RO, RU, SD, SI, SG, SK, SL,
SJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.(84) Designated States (regional): ARIPO patent (GI, GM,
KL, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD,
TG).Published:
with international search report(88) Date of publication of the international search report:
27 December 2002

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

(54) Title: DRUG METABOLIZING ENZYMES

(57) Abstract: The invention provides human drug metabolizing enzymes (DME) and polynucleotides which identify and encode DME. 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 DME.



WO 02/012467 A3

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International Application No. PCT/US 01/24382
A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C12N15/53 C12N15/54 C12N9/02 C12N9/04 C12N9/06 C12N9/10 C12N9/14 C12N9/18 C07K16/40 C12Q1/68 G01N33/68 A01K67/027 A61K38/44 A61K38/45		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K C12Q G01N A01K 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, PAJ, BIOSIS, MEDLINE, SEQUENCE SEARCH		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 06830 A (UNIV SYDNEY ;WEISS ANTHONY STEVEN (AU)) 19 February 1998 (1998-02-19) the whole document	1-10, 12-14, 16,18, 19, 22-27, 29-44
X	TRACKMAN P C ET AL: "CLONING OF RAT AORTA LYSYL OXIDASE COMPLEMENTARY DNA COMPLETE CODONS AND PREDICTED AMINO ACID SEQUENCE" BIOCHEMISTRY, vol. 29, no. 20, 1990, pages 4863-4870, XP00219862 ISSN: 0006-2960 the whole document	1-7,9, 10, 12-14, 25,29, 31, 35-37, 43,44
	---	---
	-/--	
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "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 24 May 2002		Date of mailing of the international search report 25. 09. 2002
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2880 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 851 epo nl, Fax: (+31-70) 340-3016		Authorized officer Oderwald, H

Form PCT/ISA/210 (second sheet) (July 1997)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 01/24382

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE SWALL [Online] EBI Hinxton, UK; 24 November 1998 (1998-11-24) LANGENAU D M ET AL.: "Progesterin-upregulated mRNAs in the yellow perch ovary" Database accession no. AF103901 XP002199863 abstract	1-7,12
A	--- US 5 744 355 A (WOOD THOMAS C ET AL) 28 April 1998 (1998-04-28) the whole document	
A	--- US 5 756 299 A (GOLI SURYA K ET AL) 26 May 1998 (1998-05-26) the whole document	
A	--- US 6 060 253 A (KRAUSZ KRISTOPHER W ET AL) 9 May 2000 (2000-05-09) the whole document	
E	--- WO 01 92495 A (MEYERS RACHEL ;MILLENNIUM PHARM INC (US)) 6 December 2001 (2001-12-06) the whole document	1-16,18, 19,22-44
P,X	--- ITO HIROMU ET AL: "Molecular cloning and biological activity of a novel lysyl oxidase-related gene expressed in cartilage." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 276, no. 26, 29 June 2001 (2001-06-29), pages 24023-24029, XP002199860 ISSN: 0021-9258 the whole document	1-7, 9-15,22, 25,26, 29,31, 35-37, 43,44

Form PCT/ISA/210 (continuation of second sheet) (July 1999)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 01/24382**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

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

1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
see FURTHER INFORMATION sheet PCT/ISA/210
2. Claims Nos.: (20, 21 complete), (23, 24 partially)
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

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

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

see additional sheet

1. As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
Claims 1-19, 22-44 partially, 45, 64 completely.

Remark on Protest

- The additional search fees were accompanied by the applicant's protest.
- No protest accompanied the payment of additional search fees.

International Application No. PCT/US 01/24382

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 18 and 24 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
Although claims 32 and 34 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.2

Claims Nos.: (20, 21 complete), (23, 24 partially)

Present claims 20, 21, 23 and 24 relate to a product/compound defined by reference to a desirable characteristic or property, namely an agonist or an antagonist of a polypeptide.

The claims cover all products/compounds having this characteristic or property, whereas the application provides support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT for only a very limited number of such products/compounds. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Independent of the above reasoning, the claims also lack clarity (Article 6 PCT). An attempt is made to define the product/compound by reference to a result to be achieved. Again, this lack of clarity in the present case is such as to render a meaningful search over the whole of the claimed scope impossible.

Consequently, the search has been carried out for those parts of the claims which appear to be clear, supported and disclosed, namely those parts relating to the products/compounds antibodies, antisense and ribozymes as antagonist (p. 33, 34, 69 and 73 of the description). No search has been carried out for the product/compound agonist.

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

International Application No. PCT/US 01/24382

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

1. Claims: (1-19, 22-44 partially), (45, 64 complete)

An isolated polypeptide comprising an amino acid as set forth in SEQ ID NO: 1. A polynucleotide encoding said polypeptide (SEQ ID NO: 20). A cell, a transgenic organism, an antibody, compositions comprising said polypeptide or polynucleotide. A method for detecting a target polynucleotide or a polypeptide, methods for treating a disease, methods for screening agonists, antagonists, binding compounds, modulators, a method of assessing toxicity, methods of diagnosing using said polypeptide, polynucleotide and antibody.

2. Claims: (1-19, 22-44 partially), (46, 65 complete)

same as invention 1 but comprising SEQ ID NO: 2 and 21.

3. Claims: (1-19, 22-44 partially), (47, 66 complete)

same as invention 1 but comprising SEQ ID NO: 3 and 22.

4. Claims: (1-19, 22-44 partially), (48, 67 complete)

same as invention 1 but comprising SEQ ID NO: 4 and 23.

5. Claims: (1-19, 22-44 partially), (49, 68 complete)

same as invention 1 but comprising SEQ ID NO: 5 and 24.

6. Claims: (1-19, 22-44 partially), (50, 69 complete)

same as invention 1 but comprising SEQ ID NO: 6 and 25.

7. Claims: (1-19, 22-44 partially), (51, 70 complete)

same as invention 1 but comprising SEQ ID NO: 7 and 26.

8. Claims: (1-19, 22-44 partially), (52, 71 complete)

same as invention 1 but comprising SEQ ID NO: 8 and 27.

9. Claims: (1-19, 22-44 partially), (53, 72 complete)

same as invention 1 but comprising SEQ ID NO: 9 and 28.

International Application No. PCT/US 01/24382

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

10. Claims: (1-19, 22-44 partially), (54, 73 complete)
same as invention 1 but comprising SEQ ID NO: 10 and 29.
11. Claims: (1-19, 22-44 partially), (55, 74 complete)
same as invention 1 but comprising SEQ ID NO: 11 and 30.
12. Claims: (1-19, 22-44 partially), (56, 75 complete)
same as invention 1 but comprising SEQ ID NO: 12 and 31.
13. Claims: (1-19, 22-44 partially), (57, 76 complete)
same as invention 1 but comprising SEQ ID NO: 13 and 32.
14. Claims: (1-19, 22-44 partially), (58, 77 complete)
same as invention 1 but comprising SEQ ID NO: 14 and 33.
15. Claims: (1-19, 22-44 partially), (59, 78 complete)
same as invention 1 but comprising SEQ ID NO: 15 and 34.
16. Claims: (1-19, 22-44 partially), (60, 79 complete)
same as invention 1 but comprising SEQ ID NO: 16 and 35.
17. Claims: (1-19, 22-44 partially), (61, 80 complete)
same as invention 1 but comprising SEQ ID NO: 17 and 36.
18. Claims: (1-19, 22-44 partially), (62, 81 complete)
same as invention 1 but comprising SEQ ID NO: 18 and 37.
19. Claims: (1-19, 22-44 partially), (63, 82 complete)
same as invention 1 but comprising SEQ ID NO: 19 and 38.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/US 01/24382

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9806830	A	19-02-1998	AU 733946 B2 31-05-2001
			AU 3689497 A 06-03-1998
			WO 9806830 A1 19-02-1998
			EP 0960192 A1 01-12-1999
			JP 2000517172 T 26-12-2000
			NZ 333577 A 27-10-2000
US 5744355	A	28-04-1998	US 5714594 A 03-02-1998
US 5756299	A	26-05-1998	AU 5594098 A 03-07-1998
			WO 9826076 A1 18-06-1998
US 6060253	A	09-05-2000	NONE
WO 0192495	A	06-12-2001	AU 6516201 A 11-12-2001
			WO 0192495 A2 06-12-2001
			US 2002068322 A1 06-06-2002

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 K 38/53	A 6 1 P 1/00	4 B 0 6 5
A 6 1 K 39/395	A 6 1 P 1/16	4 C 0 8 4
A 6 1 K 45/00	A 6 1 P 3/00	4 C 0 8 5
A 6 1 K 49/00	A 6 1 P 5/00	4 H 0 4 5
A 6 1 K 51/00	A 6 1 P 27/02	
A 6 1 P 1/00	A 6 1 P 35/00	
A 6 1 P 1/16	A 6 1 P 37/00	
A 6 1 P 3/00	A 6 1 P 43/00	1 1 1
A 6 1 P 5/00	C 0 7 K 16/40	
A 6 1 P 27/02	C 1 2 N 1/15	
A 6 1 P 35/00	C 1 2 N 1/19	
A 6 1 P 37/00	C 1 2 N 1/21	
A 6 1 P 43/00	C 1 2 N 9/02	
C 0 7 K 16/40	C 1 2 N 9/10	
C 1 2 N 1/15	C 1 2 P 21/08	
C 1 2 N 1/19	C 1 2 Q 1/48	
C 1 2 N 1/21	C 1 2 Q 1/68	A
C 1 2 N 5/10	G 0 1 N 33/15	Z
C 1 2 N 9/02	G 0 1 N 33/50	Z
C 1 2 N 9/10	G 0 1 N 33/53	D
C 1 2 P 21/08	G 0 1 N 33/53	M
C 1 2 Q 1/48	G 0 1 N 33/566	
C 1 2 Q 1/68	C 1 2 N 5/00	A
G 0 1 N 33/15	A 6 1 K 37/48	
G 0 1 N 33/50	A 6 1 K 49/02	A
G 0 1 N 33/53	A 6 1 K 37/50	
G 0 1 N 33/566	A 6 1 K 37/52	
	A 6 1 K 37/54	
	A 6 1 K 37/60	

- (31)優先権主張番号 60/228,067
(32)優先日 平成12年8月24日(2000.8.24)
(33)優先権主張国 米国(US)
- (31)優先権主張番号 60/230,063
(32)優先日 平成12年8月31日(2000.8.31)
(33)優先権主張国 米国(US)
- (31)優先権主張番号 60/232,244
(32)優先日 平成12年9月13日(2000.9.13)
(33)優先権主張国 米国(US)
- (31)優先権主張番号 60/234,269
(32)優先日 平成12年9月20日(2000.9.20)
(33)優先権主張国 米国(US)

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CO,CR,CU,CZ,DE,DK,DM,DZ,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,S

D, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW

- (72)発明者 ブラン、クリストファー・エム
アメリカ合衆国カリフォルニア州9 4 0 4 0・マウンテンビュー・# 1 2 6・サウスレングスト
フアベニュー 5 7 5
- (72)発明者 ダス、デボプリヤ
アメリカ合衆国カリフォルニア州9 4 0 4 0・マウンテンビュー・アパートメント 3・ボニータ
アベニュー 1 1 7 9
- (72)発明者 デレジーン、アンジェロ・エム
アメリカ合衆国カリフォルニア州9 5 0 3 5・ミルピタス・アングスドライブ 5 9 4
- (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9 4 3 0 6・パロアルト・# 1 4 6・アルマストリート 3 3 5
3
- (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 ガンディー、アミーナ・アール
アメリカ合衆国カリフォルニア州9 4 0 2 5・メンロパーク・# 1・ローブルアベニュー 8 3 7
- (72)発明者 グリフィン、ジェニファー・エイ
アメリカ合衆国カリフォルニア州9 4 5 5 5・フレモント・# 1 1・メローウェイ 3 3 6 9 1
- (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・コーレデプリマベータ 2 2 2 7
- (72)発明者 カーン、ファラ・エイ
アメリカ合衆国イリノイ州6 0 0 2 5・グレンビュー・# 1 0 2・セントラルロード 3 6 1 7
- (72)発明者 ラル、ブリーティ
アメリカ合衆国カリフォルニア州9 5 0 5 6・サンタクララ・ピーオーボックス 5 1 4 2
- (72)発明者 リー、サリー
アメリカ合衆国カリフォルニア州9 4 0 8 6・サニーバイル・# 4 2 5・イーストイープリン 8
2 5
- (72)発明者 リュ、デュング・アイナ・エム
アメリカ合衆国カリフォルニア州9 5 1 2 3・サンノゼ・コイドライブ 2 3 3
- (72)発明者 リュ、ヤン
アメリカ合衆国カリフォルニア州9 4 3 0 3・パロアルト・コリーナウェイ 3 8 8 5
- (72)発明者 パターソン、チャンドラ
アメリカ合衆国カリフォルニア州9 4 0 2 5・メンロパーク・# 1・シャーウッドウェイ 4 9 0
- (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9 4 5 5 5・フレモント・メイバードサークル 3 4 3 5 9
- (72)発明者 リング、ヒュイジュン、ジー
アメリカ合衆国カリフォルニア州9 4 0 2 2・ロスアルトス・オレンジアベニュー 6 2 5
- (72)発明者 サンジャンワラ、マデュー・エス
アメリカ合衆国カリフォルニア州9 4 0 2 4・ロスアルトス・シルビアコート 2 1 0
- (72)発明者 タング、ワイ・トム
アメリカ合衆国カリフォルニア州9 5 1 1 8・サンノゼ・ランウィックコート 4 2 3 0
- (72)発明者 サンガベル、カピサ
アメリカ合衆国カリフォルニア州9 4 0 4 3・マウンテンビュー・2 3・モンテシトアベニュー
1 9 5 0
- (72)発明者 ソートン、マイケル
アメリカ合衆国カリフォルニア州9 4 0 6 2・ウッドサイド・メッドウェイロード 9
- (72)発明者 トリボレー、キャサリン・エム
アメリカ合衆国カリフォルニア州9 4 1 0 7・サンフランシスコ・# 5・テネシーストリート 1
1 2 1

- (72)発明者 ワライヤ、ナリンダー・ケイ
 アメリカ合衆国カリフォルニア州 9 4 5 7 7 ・サンレアンドロ・# 2 0 5 ・デービスストリート
 8 9 0
- (72)発明者 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 9 5 0 1 4 ・クーペルティノー・# 2 ・パーウッドドライブ 1
 0 1 3 0
- (72)発明者 ヤング、ジュンミング
 アメリカ合衆国カリフォルニア州 9 5 1 2 9 ・サンノゼ・バンクレーン 1 7 3 9
- (72)発明者 ヤオ、モニック・ジー
 アメリカ合衆国カリフォルニア州 9 4 0 4 3 ・マウンテンビュー・フレデリックコート 1 1 1
- (72)発明者 ユエ、ヘンリー
 アメリカ合衆国カリフォルニア州 9 4 0 8 7 ・サニーベイル・ルイスアベニュー 8 2 6

F ターム(参考) 2G045 AA34 AA35 AA40 BA11 BB50 DA12 DA13 DA14 DA36 FB02
 FB03
 4B024 AA01 AA11 BA08 BA10 CA04 DA02 DA05 DA06 DA11 DA12
 EA04 GA11 HA12
 4B050 CC03 DD01 LL01 LL03
 4B063 QA18 QA20 QQ22 QQ26 QQ44 QR08 QR42 QR56 QS02 QS25
 QS34 QX02
 4B064 AG27 CA10 CA20 CC24 DA01 DA13
 4B065 AA01X AA57X AA72X AA90X AB01 BA02 CA28 CA29 CA44 CA46
 4C084 AA01 AA02 AA07 AA16 DC01 DC22 DC23 DC24 DC25 MA52
 MA55 MA57 MA59 MA60 MA63 MA65 MA66 NA06 NA14 ZA332
 ZA662 ZA752 ZB072 ZB262 ZC022 ZC212
 4C085 AA13 AA14 AA16 BB11 BB41 BB43 CC22 CC23 EE01 GG01
 HH20 KA03 KA04
 4H045 AA11 CA40 DA76 DA89 EA20 EA50 FA72 FA74

專利名稱(譯)	藥物代謝酶		
公開(公告)号	JP2004518407A	公開(公告)日	2004-06-24
申請号	JP2002517758	申請日	2001-08-03
[標]申請(專利權)人(譯)	洞察Genomics公司		
申請(專利權)人(譯)	洞察基因组公司		
[標]發明人	ボーグンマライアアール ブランクリストファーエム ダスデボプリヤ デレジーンアンジェロエム デイングリー エリオットビッキーエス ガンディーアミーナアール グリフィンジェニファーエイ ハファリアエープリルジェイエイ カーンファラエイ ラルプリーティ リーサリー リュデュングアイナエム リュヤン パターソンチャンドラ ランクマールジャヤラクシミ リングヒュイジュンジー サンジャンワラマデューエス タングワイトム サンガベルカピサ ソーントンマイケル トリボレーキャサリーンエム ワライヤナリンダーケイ ワレンブリジットエイ ヤングジュンミング ヤオモニークジー ユエヘンリー		
發明人	ボーグン、マライア・アール ブラン、クリストファー・エム ダス、デボプリヤ デレジーン、アンジェロ・エム デイング、リー エリオット、ビッキー・エス ガンディー、アミーナ・アール グリフィン、ジェニファー・エイ ハファリア、エープリル・ジェイ・エイ カーン、ファラ・エイ ラル、プリーティ リー、サリー リュ、デュング・アイナ・エム リュ、ヤン パターソン、チャンドラ ランクマール、ジャヤラクシミ リング、ヒュイジュン、ジー		

サンジャンワラ、マデュー・エス
タング、ワイ・トム
サンガベル、カピサ
ソートン、マイケル
トリボレー、キャサリーン・エム
フライヤ、ナリンダー・ケイ
ワレン、ブリジット、エイ
ヤング、ジュンミング
ヤオ、モニーク・ジー
ユエ、ヘンリー

IPC分类号	G01N33/50 A61K38/00 A61K38/43 A61K38/44 A61K38/45 A61K38/46 A61K38/53 A61K39/395 A61K45/00 A61K49/00 A61K51/00 A61P1/00 A61P1/16 A61P3/00 A61P5/00 A61P27/02 A61P35/00 A61P37/00 A61P43/00 C07K16/40 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N9/00 C12N9/02 C12N9/10 C12N15/09 C12N15/53 C12N15/54 C12P21/08 C12Q1/48 C12Q1/68 G01N33/15 G01N33/53 G01N33/566
CPC分类号	A61K38/00 A61P1/00 A61P1/16 A61P3/00 A61P5/00 A61P27/02 A61P35/00 A61P37/00 A61P43/00 C12N9/00
FI分类号	C12N15/00.ZNA.A A61K39/395.D A61K39/395.N A61K45/00 A61K49/00.A A61P1/00 A61P1/16 A61P3/00 A61P5/00 A61P27/02 A61P35/00 A61P37/00 A61P43/00.111 C07K16/40 C12N1/15 C12N1/19 C12N1/21 C12N9/02 C12N9/10 C12P21/08 C12Q1/48 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N5/00.A A61K37/48 A61K49/02.A A61K37/50 A61K37/52 A61K37/54 A61K37/60
F-TERM分类号	2G045/AA34 2G045/AA35 2G045/AA40 2G045/BA11 2G045/BB50 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA08 4B024/BA10 4B024/CA04 4B024/DA02 4B024/DA05 4B024/DA06 4B024/DA11 4B024/DA12 4B024/EA04 4B024/GA11 4B024/HA12 4B050/CC03 4B050/DD01 4B050/LL01 4B050/LL03 4B063/QA18 4B063/QA20 4B063/QQ22 4B063/QQ26 4B063/QQ44 4B063/QR08 4B063/QR42 4B063/QR56 4B063/QS02 4B063/QS25 4B063/QS34 4B063/QX02 4B064/AG27 4B064/CA10 4B064/CA20 4B064/CC24 4B064/DA01 4B064/DA13 4B065/AA01X 4B065/AA57X 4B065/AA72X 4B065/AA90X 4B065/AB01 4B065/BA02 4B065/CA28 4B065/CA29 4B065/CA44 4B065/CA46 4C084/AA01 4C084/AA02 4C084/AA07 4C084/AA16 4C084/DC01 4C084/DC22 4C084/DC23 4C084/DC24 4C084/DC25 4C084/MA52 4C084/MA55 4C084/MA57 4C084/MA59 4C084/MA60 4C084/MA63 4C084/MA65 4C084/MA66 4C084/NA06 4C084/NA14 4C084/ZA332 4C084/ZA662 4C084/ZA752 4C084/ZB072 4C084/ZB262 4C084/ZC022 4C084/ZC212 4C085/AA13 4C085/AA14 4C085/AA16 4C085/BB11 4C085/BB41 4C085/BB43 4C085/CC22 4C085/CC23 4C085/EE01 4C085/GG01 4C085/HH20 4C085/KA03 4C085/KA04 4H045/AA11 4H045/CA40 4H045/DA76 4H045/DA89 4H045/EA20 4H045/EA50 4H045/FA72 4H045/FA74
優先権	60/223055 2000-08-04 US 60/224728 2000-08-11 US 60/226440 2000-08-18 US 60/228067 2000-08-24 US 60/230063 2000-08-31 US 60/232244 2000-09-13 US 60/234269 2000-09-20 US
外部リンク	Espacenet

摘要(译)

本发明提供了鉴定和编码DME的人药物代谢酶 (DME) 和多核苷酸。 本发明还提供表达载体, 宿主细胞, 抗体, 激动剂和拮抗剂。 本发明还提供了用于诊断, 治疗或预防与DME异常表达有关的疾病的方法。

IMECE ステーション 外ID	IMECE ステーション NO.	IMECE ステーション ID	IMECE ステーション SEQ ID NO.	IMECE ステーションID
7442835	1	7442835001	20	7442835001
7472835	2	7472835001	21	7472835001
7476202	3	7476202001	22	7476202001
7478543	4	7478543001	23	7478543001
7479535	5	7479535001	24	7479535001
7479904	6	7479904001	25	7479904001
7480347	7	7480347001	26	7480347001
8063330	8	8063330001	27	8063330001
7473869	9	7473869001	28	7473869001
7479586	10	7479586001	29	7479586001
55046125	11	55046125001	30	55046125001
3538709	12	3538709001	31	3538709001
71563102	13	71563102001	32	71563102001
7472027	14	7472027001	33	7472027001
7480358	15	7480358001	34	7480358001
1618256	16	1618256001	35	1618256001
3387823	17	3387823001	36	3387823001
55142051	18	55142051001	37	55142051001
7395274	19	7395274001	38	7395274001