

(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2004-512818**(P2004-512818A)**(43) 公表日 **平成16年4月30日(2004.4.30)**

(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 O 1 K	67/027	A O 1 K	67/027		4 B O 2 4
A 6 1 K	38/00	A 6 1 K	45/00		4 B O 5 0
A 6 1 K	45/00	A 6 1 K	49/00	A	4 B O 6 3
A 6 1 K	49/00	A 6 1 P	1/00		4 B O 6 4
		審査請求 未請求 予備審査請求 有		(全 259 頁) 最終頁に続く	
(21) 出願番号	特願2001-577452 (P2001-577452)	(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 60	
(86) (22) 出願日	平成13年4月12日 (2001.4.12)	(74) 代理人	100089266		
(85) 翻訳文提出日	平成14年10月11日 (2002.10.11)		弁理士 大島 陽一		
(86) 国際出願番号	PCT/US2001/011869	(72) 発明者	ポリッキー、ジェニファー・エル		
(87) 国際公開番号	W02001/079468		アメリカ合衆国カリフォルニア州95118・サンノゼ・ジャービスコート 1511	最終頁に続く	
(87) 国際公開日	平成13年10月25日 (2001.10.25)				
(31) 優先権主張番号	60/197,590				
(32) 優先日	平成12年4月13日 (2000.4.13)				
(33) 優先権主張国	米国 (US)				
(31) 優先権主張番号	60/198,403				
(32) 優先日	平成12年4月19日 (2000.4.19)				
(33) 優先権主張国	米国 (US)				
(31) 優先権主張番号	60/200,185				
(32) 優先日	平成12年4月28日 (2000.4.28)				
(33) 優先権主張国	米国 (US)				

(54) 【発明の名称】 薬剤代謝酵素

(57) 【要約】

本発明は、ヒト薬剤代謝酵素 (DME) と、DMEを同定及びコードするポリヌクレオチドとを提供する。本発明はまた、発現ベクター及び宿主細胞、抗体、アゴニスト、アンタゴニストを提供する。更に、本発明は、DMEの異常な発現に関連する疾患の診断・治療・予防方法を提供する。

【特許請求の範囲】

【請求項 1】

単離されたポリペプチドであって、

(a) SEQ ID NO: 1 乃至 SEQ ID NO: 10 (SEQ ID NO: 1-10) からなる群から選択されたアミノ酸配列を有するポリペプチドと、

(b) SEQ ID NO: 1-10 からなる群から選択されたアミノ酸配列と少なくとも 90% の配列同一性を有する天然のアミノ酸配列と、

(c) SEQ ID NO: 1-10 からなる群から選択されたアミノ酸配列を含むポリペプチドの生物学的に活性な断片と、

(d) SEQ ID NO: 1-10 からなる群から選択されたアミノ酸配列を含む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単離されたポリペプチド。 10

【請求項 2】

SEQ ID NO: 1-10 からなる群から選択された請求項 1 の単離されたポリペプチド。

【請求項 3】

請求項 1 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 4】

請求項 2 のポリペプチドをコードする単離されたポリヌクレオチド。

【請求項 5】

SEQ ID NO: 11-20 からなる群から選択された請求項 4 の単離されたポリヌクレオチド。 20

【請求項 6】

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

【請求項 7】

請求項 6 の組換えポリヌクレオチドで形質転換された細胞。

【請求項 8】

請求項 6 の組換えポリヌクレオチドを含む遺伝子組換え生物。

【請求項 9】

請求項 1 のポリペプチドを生産する方法であって、 30

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

(b) そのように発現したポリペプチドを回収するステップとを含むことを特徴とする請求項 1 のポリペプチドの生産方法。

【請求項 10】

請求項 1 のポリペプチドに特異的に結合する単離された抗体。

【請求項 11】

単離されたポリヌクレオチドであって、

(a) SEQ ID NO: 11-20 からなる群から選択されたポリヌクレオチド配列を有するポリヌクレオチドと、 40

(b) SEQ ID NO: 11-20 からなる群から選択されたポリヌクレオチド配列と少なくとも 90% の配列同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチド配列と、

(c) 前記 (a) のポリヌクレオチドに相補的なポリヌクレオチドと、

(d) 前記 (b) のポリヌクレオチドに相補的なポリヌクレオチドと、

(e) 前記 (a) 乃至 (d) の RNA 等価物とで構成される群から選択された単離されたポリヌクレオチド。

【請求項 12】

請求項 11 のポリヌクレオチドの少なくとも 60 個の連続するヌクレオチドを含む単離さ 50

れたポリヌクレオチド。

【請求項 13】

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

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

(b) 前記ハイブリダイゼーション複合体が存在するか否かを検出し、存在する場合には任意選択でその収量を測定するステップとを含むことを特徴とする標的ポリヌクレオチドの検出方法。 10

【請求項 14】

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

【請求項 15】

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

(a) ポリメラーゼ連鎖反応増幅を用いて、前記標的ポリヌクレオチドまたはその断片を増幅するステップと、 20

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

【請求項 16】

請求項 1 のポリペプチド及び医薬的に容認できる賦形剤を含む組成物。

【請求項 17】

前記ポリペプチドが、SEQ ID NO: 1-10 からなる群から選択されたアミノ酸配列を含むことを特徴とする請求項 16 の組成物。

【請求項 18】

機能的な DME の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 16 の組成物を投与することを含むことを特徴とする治療方法。 30

【請求項 19】

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

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

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

【請求項 20】

請求項 19 のスクリーニング方法によって同定されたアゴニスト化合物及び医薬的に容認できる賦形剤を含む組成物。 40

【請求項 21】

機能的な DME の発現の低下に関連する疾患や病態の治療方法であって、そのような治療が必要な患者に請求項 20 の組成物を投与することを含むことを特徴とする治療方法。

【請求項 22】

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

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

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

【請求項 23】

請求項 22 のスクリーニング方法によって同定されたアンタゴニスト化合物及び医薬的に容認できる賦形剤を含む組成物。

【請求項 24】

機能的な DME 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 23 の組成物を投与することを含むことを特徴とする治療方法。

【請求項 25】

請求項 1 のポリペプチドに特異的に結合する化合物をスクリーニングする方法であって、
(a) 請求項 1 のポリペプチドを好適な条件下で少なくとも 1 つの試験化合物と結合させるステップと、

(b) 請求項 1 のポリペプチドと前記試験化合物との結合を検出して、請求項 1 のポリペプチドと特異的に結合する化合物を同定するステップとを含むことを特徴とするスクリーニング方法。 10

【請求項 26】

請求項 1 のポリペプチドの活性を変化させる化合物をスクリーニングする方法であって、
(a) 請求項 1 のポリペプチドを、その活性が許容される条件下で少なくとも 1 つの試験化合物と結合させるステップと、

(b) 前記試験化合物の存在下での請求項 1 のポリペプチドの活性を評価するステップと、

(c) 前記試験化合物の存在下での請求項 1 のポリペプチドの活性と、前記試験化合物の不在下での請求項 1 のポリペプチドの活性とを比較するステップとを含み、
前記試験化合物の存在下での請求項 1 のポリペプチドの活性の変化が、請求項 1 のポリペプチドの活性を変化させる化合物の存在を示唆すること特徴とするスクリーニング方法。 20

【請求項 27】

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

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

(b) 前記標的ポリヌクレオチドの発現の変化を検出するステップと、

(c) 様々な量の前記化合物の存在下での前記標的ポリヌクレオチドの発現と、前記化合物の不在下での前記標的ポリヌクレオチドの発現とを比較するステップとを含むことを特徴とするスクリーニング方法。 30

【請求項 28】

試験化合物の毒性を評価する方法であって、

(a) 核酸を含む生体サンプルを前記試験化合物で処理するステップと、

(b) 処理した前記生体サンプルの核酸と、請求項 12 のポリヌクレオチドの少なくとも 20 の連続するヌクレオチドを含むプローブをハイブリダイズさせるステップであって、このハイブリダイゼーションが、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 11 のポリヌクレオチドのポリヌクレオチド配列またはその断片を含むポリヌクレオチドである、前記ステップと、 40

(c) ハイブリダイゼーション複合体の収量を定量するステップと、

(d) 前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処理の生体サンプルにおけるハイブリダイゼーション複合体の収量と比較するステップとを含み、

前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量の差が試験化合物の毒性を示唆すること特徴とする試験化合物の毒性評価方法。

【請求項 29】

生体サンプルにおいて、DME の発現に関連する病態または疾患の診断検査を行うための方法であって、

(a) 前記生体サンプルと請求項 10 の抗体とを、前記抗体と前記ポリペプチドが結合し 50

て抗体－ポリペプチド複合体が形成される条件下で結合させるステップと、

(b) 前記複合体を検出するステップとを含み、

前記複合体の存在が前記生体サンプル内に前記ポリペプチドが存在することと相関することとを特徴とする方法。

【請求項 30】

抗体であって、

(a) キメラ抗体と、

(b) 単鎖抗体と、

(c) Fab 断片と、

(d) F(ab')₂ 断片と、

(e) ヒト化抗体とであることを特徴とする請求項 10 に記載の抗体。

10

【請求項 31】

請求項 10 の抗体及び容認できる賦形剤を有する組成物。

【請求項 32】

DME の発現に関連する病態または疾患の診断方法であって、そのような治療が必要な患者に請求項 31 の組成物を有効量投与することを含むことを特徴とする診断方法。

【請求項 33】

前記抗体がラベルされていることを特徴とする請求項 31 の組成物。

【請求項 34】

DME の発現に関連する病態または疾患の診断方法であって、そのような治療が必要な患者に請求項 33 の組成物を有効量投与することを含むことを特徴とする診断方法。

20

【請求項 35】

請求項 10 の抗体の特異性を有するポリクローナル抗体を作製する方法であって、

(a) 抗体反応が起こる条件下で、SEQ ID NO: 1-10 のポリペプチドまたはその免疫原性断片を含む一群より選択されたアミノ酸配列を有するポリペプチドで動物を免疫化するステップと、

(b) 前記動物から抗体を単離するステップと、

(c) 前記単離された抗体を前記ポリペプチドでスクリーニングして、SEQ ID NO: 1-10 を含む一群より選択されたアミノ酸配列を有するポリペプチドに特異的に結合するポリクローナル抗体を同定するステップとを含むことを特徴とするポリクローナル抗体作製方法。

30

【請求項 36】

請求項 35 の方法によって生成される抗体。

【請求項 37】

請求項 36 の前記抗体及び好適な担体を有する組成物。

【請求項 38】

請求項 10 の抗体の特異性を有するポリクローナル抗体を作製する方法であって、

(a) 抗体反応が起こる条件下で、SEQ ID NO: 1-10 のポリペプチドまたはその免疫原性断片を含む一群より選択されたアミノ酸配列を有するポリペプチドで動物を免疫化するステップと、

(b) 前記動物から細胞生成抗体を単離するステップと、

(c) 細胞生成抗体と不死化した抗体とを融合させ、融合細胞生成モノクローナル抗体を形成するステップと、

(d) ハイブリドーマ細胞を培養する過程と、

(e) SEQ ID NO: 1-10 よりなる一群より選択されたアミノ酸配列を有するポリペプチドに特異的に結合するモノクローナル抗体培養より単離するステップとを含むことを特徴とするモノクローナル抗体作製方法。

40

【請求項 39】

請求項 38 の方法によって生成される単クローン抗体。

【請求項 40】

50

請求項 39 の前記抗体及び好適な担体を有する組成物。

【請求項 41】

前記抗体が F a b 発現ライブラリをスクリーニングすることで生成されることを特徴とする請求項 10 に記載の抗体。

【請求項 42】

前記抗体が組み換え型免疫グロブリンライブラリをスクリーニングすることで生成されることを特徴とする請求項 10 に記載の抗体。

【請求項 43】

サンプルにおいて、SEQ ID NO: 1-10 よりなる一群より選択されたアミノ酸配列を有するポリペプチドを検出する方法であって、

10

(a) 請求項 10 の抗体と前記ポリペプチドとの特異的な結合が許容される条件下で、前記抗体を前記サンプルと共にインキュベートするステップと、

(b) 特異的な結合を検出するステップとを含み、

前記特異的な結合が前記サンプルにおける SEQ ID NO: 1-10 よりなる一群より選択されたアミノ酸配列を有するポリペプチドの存在を示唆することを特徴とする方法。

【請求項 44】

サンプルにおいて、SEQ ID NO: 1-10 よりなる一群より選択され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) 請求項 10 の抗体と前記ポリペプチドとの特異的な結合が許容される条件下で、前記抗体を前記サンプルと共にインキュベートするステップと、

20

(b) 前記サンプルより前記抗体を分離させ、SEQ ID NO: 1-10 よりなる一群より選択されたアミノ酸配列を有する精製ポリペプチドを得るステップ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】

30

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】

40

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 11 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 11 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 11 に記載のポリペプチド。

【請求項 58】

50

SEQ ID NO: 14 のアミノ酸配列を含む請求項 11 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 11 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 11 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 11 に記載のポリペプチド。

【請求項 62】

SEQ ID NO: 18 のアミノ酸配列を含む請求項 11 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 11 に記載のポリペプチド。

【請求項 64】

SEQ ID NO: 20 のアミノ酸配列を含む請求項 11 に記載のポリペプチド。

【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（薬物動体学）は、その効果、毒性、およびその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除の3つのプロセスが薬物動体学を司っている。これらのプロセスは、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学的特性のほとんどを変えるため、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。

【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、一部の患者のグループ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）。

【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、およびフェーズ II に分類され、それらは連続している。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、レボドパ）。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

10

20

30

40

50

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

10

【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。

【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 (ExPASy ENZYME EC 1. 14. 14. 1; Prosite PDOC00081 Cytochrome P450 cysteine heme-iron ligand signature; PRINTS EP450 I E-Class P450 Group I signature; Graham-Lorence, S. および Peterson, J. A. (1996) FASEB J. 10: 206-214. を参照)。

30

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

50

において同定された (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 の大きなファミリーの中に含まれる (PRINTS EP450I E-Class P450 Group I signature)。

【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 α ヘリックスと約22%の β 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 (Prosite 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、DNAと付加物を形成する反応性中間体に化学物質を代謝することで起こる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、発癌を引き起こすス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

30

【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 (APCED) 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)。

40

【0012】

チトクローム P450 における突然変異は、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である副腎皮質過形成、プソイドビタミンD-欠損くる病、脳腱黄色腫症、進行性の神経障害によって特徴付けられる脂質貯蔵病、早期アテローム性動脈硬化症、白内障、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, pp. 1968-1970 ; Takeyama, K. 他 (1997) Science 277 : 1827-1830 ; Kitanaka, S. 他 (1998) N. Engl. J. Med. 338 : 6

50

53-661 ; OMIM *213700 Cerebrotendinous xanthomatosis ; および OMIM #122700 Coumarin resistance)。チトクロームP450タンパク質アロマターゼの極端に高い発現レベルが、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は、チトクロームP450触媒サイクルへの電子の小供与体として広く見られる。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。

【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）の活性型ヘモグロビン（ferrohemoglobin）へ還元する。酸化剤即ち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）。

【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 Endocrinology and Metabolism 11 : 315-319を参照）。

【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は最も豊富に存在するビタミンDの前駆体であって、更に酵素25-ヒドロキシビタミンD 1 α -ヒドロキシラーゼ（1 α -ヒドロキシラーゼ）によって、活性型である1 α , 25-ジヒドロキシビタミンD（1 α , 25（OH）₂D）に腎臓において代謝される。1 α , 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 α -ヒドロキシラーゼ活性は、酵素産物（1 α , 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、成長因子、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 α -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 α , 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 α , 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D 24-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)。

【0017】

ビタミンD 25-ヒドロキシラーゼ、 1α -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D 25-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($1\alpha, 25(\text{OH})_2\text{D}$)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、 1α -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あひる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D 25-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D 25-ヒドロキシラーゼの欠損がしてもくる病が起きないことから、 $25(\text{OH})\text{D}$ 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、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、基質の通常範囲外の求核性の窒素、硫黄、およびリンのヘテロ原子を酸化する。チトクロームP450と同様に、FMOはミクロソーム酵素であってNADPHおよび O_2 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モチーフおよびFAD結合領域を含む（Steher, M. 他（1998）Trends Biochem. Sci. 23 : 56-57 ; PRINTS FMOXYGENASE 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ソへの酸化、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。

【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の肺発現型は、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。

【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、魚臭症候群を引き起こす（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関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。β-アミノプロピオノ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応答して放出されるホルモンのレベルの上昇に応答して増大する。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-β、血小板由来成長因子、アンギオテンシン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）。

【0026】

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。この反応は、グリシンおよびプリンの新規の合成、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-ジヒドロ葉酸 + NADPH $\rightarrow$ 5, 6, 7, 8-テトラヒドロ葉酸 + NADP⁺ である。この酵素は、trimethoprim およびメトトレキサ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 或いは複数の DHFR 遺伝子の複製の結果として、メトトレキサ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 and Co., Inc. New York. pp. 511-5619)。

10

【0027】

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。

20

【0028】

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 Aldose-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)。

30

【0029】

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。

40

【0030】

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 イソ酵素はエタノールおよ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

50

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。

【0031】

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-β-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-β-デヒドロゲナーゼ、エストラジオール-17-β-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-β-ヒドロキシステロイドデヒドロゲナーゼ、20-β-ヒドロキシステロイドデヒドロゲナーゼ、リピ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-α-ヒドロキシステロイドデヒドロゲナーゼ、cis-1, 2-dihydroxy-3, 4cyclohexadiene-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、ビフェニル-2, 3-ジヒドロ-2, 3-ジオールデヒドロゲナーゼ、N-acetylmannosamine 1dehydrogenase、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)。

【0032】

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送を触媒する。この転送は、通常は嗅覚性ヘテロ原子 (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、並びにビリルビン、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。

【0033】

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P450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)。

【0034】

薬剤代謝に関係するUGTは、UGT1およびUGT2の2つの遺伝子ファミリーによってコードされる。UGT1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。UGT2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、UGT2AおよびUGT2Bの2つのファミリーに分類される。2Aサブファミリーは嗅上皮で発現し、2Bサブファミリーは肝臓ミクロソームで発現する。UGT遺伝子における突然変異は、高ビリルビン血症 (OMIM#143500 Hyperbiliru

b i n e m i a 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 (O M I M # 2 1 8 8 0 0 C r i g l e r - N a j j a r s y n d r o m e)、ギルバール病と呼ばれる軽度の高ビリルビン血症 (O M I M * 1 9 1 7 4 0 U G T 1) に関係する。

【0035】

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 (S T) は、補助因子 3'-ホスホアデノシン-5'-ホスホ硫酸 (P A P S) から基質へ S O₃ を転移してこの反応を触媒する。S T 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

【0036】

S T 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種類を越える S T 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 S T は、それらの基質選択性に基づいて、アリー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

【0037】

S T 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、S T 発現プロファイルの差異並びにチトクローム P 4 5 0 等の他の D M E 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 S T の高発現が、U D P グルクロニ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。

【0038】

S T 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 l e v a d o p a 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 S T 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。この最後の酵素は、細胞生化学における S T 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、成熟した硫酸ケラタン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 (N a k a z a w a, K. 他 (1984) J. Biol. Chem. 259 : 13751-7 ; O M I M * 2 1 7 8 0 0 M a c u l a r d y s t r o p h y, c o r n e a l)。

30

【0039】

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 (G l c N A c) オリゴ糖鎖にガラクトース (G a l) 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 (K o l b i n g e r, F. 他 (1998) J. Biol. Chem. 273 : 433-440 ; A m a d o, 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。β 1, 3-ガラクトシルトランスフェラーゼは、G a l (β 1-3) G l c N A c 結合を有する I 型炭化水素鎖を形成する。既知

50

のヒトおよびマウス $\beta 1$, 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 β -N-アセチルグルコサミン $\beta 1$, 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 β -N-アセチルグルコサミン $\beta 1$, 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、brainiacタンパク質が $\beta 1$, 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

【0040】

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 ($\beta 1$ -4) GlcNAc結合を有するII型炭化水素鎖の形成を触媒する。 $\beta 1$, 3-ガラクトシルトランスフェラーゼの場合と同様に、可溶型の酵素が、膜結合型の切断によって形成される。 $\beta 1$, 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 (Yadav, S. および Brew, K. (1990) J. Biol. Chem. 265 : 14163-14169 ; Yadav, 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)。 $\beta 1$, 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上に炭化水素鎖を合成するのに加えて、いくつかの特定の役割を有する。哺乳動物において、 α -ラクトアルブミンを有するヘテロ二量体の一部として $\beta 1$, 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 $\beta 1$, 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 $\beta 1$, 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/基底板相互作用、正常な細胞遊走、および転移性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)。

20

30

【0041】

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、D101、E104、およびR131は、GSHの結合に重要である (Lee, H-C

40

50

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)。

【0042】

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し得る。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、dihalomethanes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)。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αGST, 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)。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。

10

【0043】

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する。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関係し、薬剤に応答して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関係する (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。

20

30

【0044】

γ-グルタミルトランスぺプチダーゼ

γ-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、γ-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システインが集まった領域を細胞に提供する。γ-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γ-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)。

40

【0045】

アシルトランスフェラーゼ

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-C

50

o A 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。

【0046】

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)。

10

【0047】

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徹底的に研究された。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 の鋳型のプロモータ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) Current Opinion in Cell Biology 12:326-333 および Berger, S. L. (1999) Current Opinion in Cell Biology 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 α/β ヒドロラーゼフォールド (α/β 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>)。

20

30

40

【0048】

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、同じ酵素によってある程度 O-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 つのステップでアセチル基を転移するサイトゾル酵素である。第 1 のステップでは、アセチル基がアセチル-CoA から活性部位シ

50

ステイン残基に転移され、第2のステップ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。

【0049】

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NAT1およびNAT2が存在し、マウスには第3型酵素であるNAT3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（NAT1は広範に発現されるが、NAT2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、NAT1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NAT2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、ダプソン、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

【0050】

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NET2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、NET1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）。NAT1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NAT1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-acetyltransferase1）。

20

【0051】

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-L-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

【0052】

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

【0053】

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媒して2-oxoadipate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

50

トランスフェラーゼの欠損が *pleiotropic* 効果に関係すると思われる (Buchhl i, R. 他 (1995) J. Biol. Chem. 270:29330-29335)。

【0054】

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、レボドパ、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型よりも位置特異的である。この酵素の可溶型の翻訳は、完全長mRNA (1.5 kb) 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 (1.3 kb) 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。

10

【0055】

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-propionophenone) 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、レボド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、fenoldopam、アポモルフィン、およびα-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される。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)。

20

30

【0056】

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O₂ およびH₂O₂ への不均化反応を触媒する。この不均化反応の速度は、拡散を制限し、後に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、或る細菌種の周辺質でも同定された。銅-亜鉛スーパーオキシドジスムターゼは元気のいい酵素であって、蛋白分解に高い耐性を有し、尿素およびSDSによって変性される。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)。

40

【0057】

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解凍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アップレギュレートにより、酵母細胞が更なる過酸化ストレス

50

に適應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)。

【0058】

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倍アップレギュレートされる。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2桁多く発現し、極めて高い割合で発現酵素を分泌する。この結果、結核菌によって *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)。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)。

10

【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。Tyre 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行き止まり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)。

20

30

【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、イソソームにおいてスフィンゴイリ分子が蓄積され、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)。

【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

40

50

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)。

【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)。

【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バリエーションとして発現される。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バリエーションが存在する (Conti, M. および S. -L. C. Jin (1999) Prog. Nucleic Acid Res. Mol. Biol. 63:1-38)。多数の PDE ファミリー、イソ酵素、およびスプライスバリエーションの存在は、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 (Houslay, 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バリエーション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)。

【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)。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】

PDE 3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PDE 3の競合的基質 (competitive substrate) として作用する。PDE 3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PDE 3ファミリーは、cilostamide、enoximone、およびlixazaloneなどの特定のインヒビターに対する感受性を有する。PDE 3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)。

10

【0067】

PDE 4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、PDE 4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。PDEインヒビターは、喘息、慢性塞栓性肺疾患、およびアトピー性湿疹の治療薬として臨床試験が行われている。PDE 4の既知の4つ全てのイソ酵素は、マウスの行動記憶を改善することが分かっている化合物であるインヒビターロリプラムに対して感受性が高い (Barad, M. 他 (1998) Proc. Natl. Acad. Sci. USA 95: 15020-15025)。PDE 4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)。

20

【0068】

PDE 5は、基質としてのcGMPに対して高い選択性を有し (Turko, I. V. 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PDE 5のリン酸化にとって重要であると思われる。PDE 5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はPDE 5およびPDE 1に対して効果がある。PDE 5に対する特異性を得るためにザプリナストを改良してsildenafilを生成した (VIGRA; 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)。PDE 5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)。

30

【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PDE 6は、光伝達カスケードの重要な要素である。PDE 6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、PDE 6はまた、PDE 6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)。PDE 6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 (Su

40

50

ber, 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。

【0070】

PDE の PDE7 ファミリーは、複数のスプライスバリエント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、および脾臓で見られるが、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)。

【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、ビンポセチン、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)。

【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、ビンポセチン、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、および精巣で発現される (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. Nucleic Acid Res. Mol. Biol. 63:1-38)。保存された推定上の亜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 c G M P 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) X n F X₃ D E のタンデムリピートを含む (M c A l l i s t e r - L u c a s, L. M. 他 (1993) J. B i o l. C h e m. 268:22863-22873)。NKXnDモチーフは変異誘発によって見られ、c G M P 結合に重要である (T u r k o, I. V. 他 (1996) J. B i o l. C h e m. 271:22240-22244)。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。

10

【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 c A M P のレベルを上昇させる薬剤によって阻害される (V e r g h e s e, M. W. 他 (1995) M o l. P h a r m a c o l. 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K_m c A M P P E D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。

【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 (P e r r y, M. J. および G. A. H i g g s (1998) C u r r. O p i n. C h e m. B i o l. 2:472-481; T o r p h y, T. J. (1998) A m. J. R e s p i r. C r i t. C a r e M e d. 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。ロリプラムはまた、i n v i t r oでH I V - 1の複製を促すことが認められたリポ多糖 (L P S) 誘導性T N F - aを阻害することが分かった。従って、ロリプラムはH I V - 1の複製を阻害すると考えられる (A n g e l, J. B. 他 (1995) A I D S 9:1137-1144)。更に、ロリプラムが、T N F - a、T N F - 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、遅発性ジスキネジア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 (S o m m e r, N. 他 (1995) N a t. M e d. 1:244-248; S a s a k i, H. 他 (1995) E u r. J. P h a r m a c o l. 282:71-76)。

20

30

【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 (B a n n e r, K. H. および C. P. P a g e (1995) E u r. R e s p i r. 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、T N F - aの生成を阻止し、H I V - 1の複製を阻害し得る (A n g e l 他, 前出)。

40

【0078】

PDEは、様々な細胞型の細胞増殖に影響を与え (C o n t i 他 (1995) E n d o c r i n e R e v. 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D U 1 4 5およびL N C a Pの成長は、c A M P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 (B a n g, Y. J. 他 (1994) P r o c. N a t l. A c a d. S c i. U S A 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 (M a t o u s o v i c,

50

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) ことが提案された。癌治療は、PDEを腫瘍の特定の細胞区画に送達し、細胞死を導くことであると記載されている (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 diazoxon および 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。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読み終わりステップを触媒する。読み終わり (chain-terminating) ステップは、脂肪酸アシル鎖を脂肪酸シン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読み終わり酵素 (chain-terminating enzyme) 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イ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読み終わり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) が、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考えられる。

【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、トロンビン、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लग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、ピロリジジナルカロイド、ステロイド、p-ニトロフェニル酢酸、マラチオン、butanillicaine、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）。

10

【0083】

neurologins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、neurologins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. 271:2676-2682）。

20

【0084】

スクアレノエポキシダーゼ

スクアレノエポキシダーゼ（スクアレノモノオキシゲナーゼ、SE）は、ミクロソーム膜結合FAD依存性オキシドレドクターゼであって、真核細胞のステロール生合成経路における初めの酸素負荷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)-オキシドスクアレノに変換し、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関係す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. pp. 554-560およびSakakibara, J. 他 (1995) 270:17-20）。

30

【0085】

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→ゲラニルピロリン酸→ファルネシルピロリン酸→スクアレノ→スクアレノエポキシド→ラノステロール→コレステロール

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-

40

50

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）。

【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 α/β ヒドロラーゼフォールド（ α/β 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、Streptomyces aureofaciens由来ブロモペルオキシダーゼA2（bromoperoxidase A2）、シュードモナス-プチダ由来のhydroxymuconic semialdehyde hydrolases、および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への加水分解、threo-9, 10-dihydroxyoctadec-12 (Z)-enoic acid（leukotoxin diol）、およびcis-12, 13-epoxyoctadec-9 (Z)-enoic acid（isoleukotoxin）からその対応するdiol threo-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）。

【0088】

チロシン触媒作用に関与する酵素

コハク酸とピルビン酸か、或いはフマル酸とアセト酢酸へのアミノ酸チロシンの分解には多数の酵素が必要であり、多数の中間化合物が生成される。加えて、多くの生体異物化合物は、チロシン分解経路の一部である1或いは複数の反応が用いられて代謝され得る。この経路は初め細菌で研究されたが、チロシン分解は様々な生物において起こることが知られ、多数の同様の生物学的反応が関係すると思われる。

【0089】

コハク酸とピルビン酸へのチロシンの分解に関係する酵素には、4-ヒドロキシフェニルピルビン酸オキシダーゼ、4-ヒドロキシフェニル酢酸3-ヒドロキシラーゼ（4-hy

droxyphenylacetate 3-hydroxylase)、3, 4-ジヒドロキシフェニル酢酸 (3, 4-dihydroxyphenylacetate 2, 3-dioxygenase)、5-カルボキシメチル-2-ヒドロキシムコン酸セミアルデヒドデヒドロゲナーゼ (5-carboxymethyl-2-hydroxymuconic semialdehyde dehydrogenase)、トランス, シス-5-カルボキシメチル-2-ヒドロキシムコン酸イソメラーゼ (trans, cis-5-carboxymethyl-2-hydroxymuconate isomerase)、ホモプロトカテチュ酸イソメラーゼ/デカルボキシラーゼ (homoprotocatechuate isomerase/decarboxylase)、cis-2-oxohept-3-ene-1, 7-dioate hydratase、2, 4-dihydroxyhept-trans-2-ene-1, 7-dioate aldolase、およびコハク酸セミアルデヒドデヒドロゲナーゼ (succinic semialdehyde dehydrogenase) が含まれる。

10

【0090】

チロシンのフマル酸塩とアセト酢酸塩への分解に関係する酵素には、4-ヒドロキシフェニルピルビン酸ジオキシゲナーゼ、ホモゲンチシン酸-1, 2-ジオキシゲナーゼ、マレイルアセト酢酸イソメラーゼ、およびフマリルアセトアセターゼが含まれる。コハク酸/ピルビン酸経路からの中間体が受け入れられた場合は、4-ヒドロキシフェニル酢酸1-ヒドロキシラーゼ (4-hydroxyphenylacetate 1-hydroxylase) が関係し得る。

20

【0091】

異なった生物においてチロシン代謝に関係する更なる酵素には、4-クロロフェニル酢酸-3, 4-ジオキシゲナーゼ (4-chlorophenylacetate-3, 4-dioxygenase)、芳香族アミノトランスフェラーゼ、5-oxopent-3-ene-1, 2, 5-tricarboxylate decarboxylase、2-oxo-hept-3-ene-1, 7-dioate hydratase、および5-カルボキシメチル-2-ヒドロキシムコン酸イソメラーゼ (5-carboxymethyl-2-hydroxymuconate isomerase) が含まれる (Ellis, L. B. M. 他 (1999) Nucleic Acids Res. 27:373-376; Wackett, L. P. および Ellis, L. B. M. (1996) J. Microbiol. Meth. 25:91-93; および Schmidt, M. (1996) Amer. Soc. Microbiol. News 62:102)。

30

【0092】

ヒトにおいて、チロシン分解経路の酵素における後天性或いは先天性の遺伝的欠陥が、遺伝性チロシン血症 I 型を引き起こし得る。この疾患の1つの型である遺伝性チロシン血症 I 型 (HT1) は、チロシンをフマル酸とアセテート酢酸に代謝する生物における経路の最後の酵素である酵素フルマリルアセトアセターゼヒドロラーゼの欠損によって引き起こされ得る。HT1 は幼児期に始まる進行性の肝傷害によって特徴づけられ、肝癌のリスクが高い (Endo, F. 他 (1997) J. Biol. Chem. 272:24426-24432)。

40

【0093】

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。

【0094】

(発明の要約)

50

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」、「DME-8」、「DME-9」、および「DME-10」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。本発明の一実施態様では、(a) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-10 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アミノ酸配列を含む天然ポリペプチドと、(c) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-10 と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一群から選択された単離されたポリペプチドを提供する。別法では、SEQ ID NO: 1-10 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

【0095】

更に本発明は、(a) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-10 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を含むポリペプチドと、(c) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドの免疫原性断片とで構成される一群から選択されたポリペプチドをコードする単離されたポリヌクレオチドを提供する。別法では、このポリヌクレオチドは、SEQ ID NO: 1-10 からなる一群から選択されたポリペプチドをコードする。別法では、このポリヌクレオチドは、SEQ ID NO: 11-20 からなる一群から選択される。 20

【0096】

更に、本発明は、(a) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-10 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を含むポリペプチドと、(c) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドの免疫原性断片とで構成される一群から選択されたポリペプチドをコードするポリヌクレオチドと機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドを提供する。別法では、本発明は、この組換えポリヌクレオチドで形質転換された細胞を提供する。更なる別法では、本発明は、この組換えポリヌクレオチドを含む遺伝子組換え生物を提供する。 30

【0097】

また、本発明は、(a) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-10 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を含むポリペプチドと、(c) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-10 とからなる一群から選択されたアミノ酸配列を含むポリペプチドの免疫原性断片とで構成される一群から選択されたポリペプチドの生産方法を提供する。この方法は、(a) このポリペプチドの発現に好適な条件下で、このポリペプチドをコードするポリヌクレオチドと機能的に結合されたプロモーター配列を含む組換えポリヌクレオチドで形質転換された細胞を培養するステップと、(b) このように発現したポリペプチドを回収するステップとを含む。 40

【0098】

更に、本発明は、(a) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-10 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アミノ酸配列を含む天然のポリペ 50

プチドと、(c) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-10 とからなる一群から選択されたアミノ酸配列を含むポリペプチドの免疫原性断片とで構成される一群から選択されたポリペプチドに特異的に結合する単離された抗体を提供する。

【0099】

更に、本発明は、(a) SEQ ID NO: 11-20 からなる一群から選択されたポリヌクレオチド配列を含むポリヌクレオチド配列と、(b) SEQ ID NO: 11-20 からなる一群から選択されたポリヌクレオチド配列と90%以上の配列同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチド配列と、(c) 前記(a)に相補的なポリヌクレオチド配列と、(d) 前記(b)に相補的なポリヌクレオチド配列と、(e) 前記(a)乃至(d)のRNA等化物とで構成される一群から選択された単離されたポリヌクレオチドを提供する。別法では、このポリヌクレオチドは、少なくとも60個の連続するヌクレオチドを含む。

10

【0100】

更に本発明は、(a) SEQ ID NO: 11-20 からなる一群から選択されたポリヌクレオチド配列を含むポリヌクレオチドと、(b) SEQ ID NO: 11-20 からなる一群から選択されたポリヌクレオチド配列と90%以上の配列同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチドと、(c) 前記(a)に相補的なポリヌクレオチド配列と、(d) 前記(b)に相補的なポリヌクレオチド配列と、(e) 前記(a)乃至(d)のRNA等化物とで構成される一群から選択されたポリヌクレオチド配列を有する標的ポリヌクレオチドをサンプルにおいて検出する方法を提供する。この方法は、(a) 前記サンプル内の標的ポリヌクレオチドと相補的な配列を構成する少なくとも20個の連続するヌクレオチドを含むプローブと前記サンプルをハイブリダイズさせるステップであって、前記プローブと前記標的ポリヌクレオチドまたはその断片とでハイブリダイゼーション複合体が形成される条件下で、前記プローブが前記標的ポリヌクレオチドに特異的にハイブリダイズする、該ステップと、(b) 前記ハイブリダイゼーション複合体の存在するか否かを検出し、存在する場合には随意選択でその収量を測定するステップとを含む。別法では、前記プローブは、少なくとも60個の連続するヌクレオチドを含む。

20

【0101】

更に本発明は、(a) SEQ ID NO: 11-20 からなる一群から選択されたポリヌクレオチド配列を含むポリヌクレオチド配列と、(b) SEQ ID NO: 11-20 からなる一群から選択されたポリヌクレオチド配列と90%以上の配列同一性を有するポリヌクレオチド配列を含む天然のポリヌクレオチド配列と、(c) 前記(a)に相補的なポリヌクレオチド配列と、(d) 前記(b)に相補的なポリヌクレオチド配列と、(e) 前記(a)乃至(d)のRNA等化物とで構成される一群から選択されたポリヌクレオチド配列を有する標的ポリヌクレオチドをサンプルにおいて検出する方法を提供する。この方法は、(a) ポリメラーゼ連鎖反応増幅を用いて、前記標的ポリヌクレオチドまたはその断片を増幅するステップと、(b) 増幅された前記標的ポリヌクレオチドまたはその断片が存在するか否かを検出し、存在する場合には随意選択でその収量を測定するステップとを含む。

30

40

【0102】

更に本発明は、(a) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-10 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アミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1-10 からなる一群から選択されたアミノ酸配列を含むポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-10 とからなる一群から選択されたアミノ酸配列を含むポリペプチドの免疫原性断片とで構成される一群から選択された効果的な量のポリペプチド及び好適な医薬用賦形剤を含む組成物を提供する。一実施例では、SEQ ID NO: 1-10 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

50

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。

【0103】

更に本発明は、(a) SEQ ID NO: 1-10からなる一群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-10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アミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1-10からなる一群から選択されたアミノ酸配列を含むポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-10とからなる一群から選択されたアミノ酸配列を含むポリペプチドの免疫原性断片とで構成される一群から選択されたポリペプチドのアゴニストとして効果的な化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを含むサンプルを化合物に曝露するステップと、(b) このサンプルのアゴニスト活性を検出するステップとを含む。別法では、本発明は、この方法によって同定されたアゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的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

【0104】

更に、本発明は、(a) SEQ ID NO: 1-10からなる一群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-10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アミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1-10からなる一群から選択されたアミノ酸配列を含むポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-10とからなる一群から選択されたアミノ酸配列を含むポリペプチドの免疫原性断片とで構成される一群から選択されたポリペプチドのアнтаゴニストとして効果的な化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを含むサンプルを化合物に曝露するステップと、(b) このサンプルのアンタゴニスト活性を検出するステップとを含む。別法では、本発明は、この方法によって同定されたアンタゴニスト化合物と好適な医薬用賦形剤とを含む組成物を提供する。更なる別法では、本発明は、この組成物の患者への投与を含む、機能的DMEの過剰な発現に関連した疾患やその症状の治療方法を提供する。

20

【0105】

更に本発明は、(a) SEQ ID NO: 1-10からなる一群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-10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アミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1-10からなる一群から選択されたアミノ酸配列を含むポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-10とからなる一群から選択されたアミノ酸配列を含むポリペプチドの免疫原性断片とで構成される一群から選択されたポリペプチドに特異的に結合する化合物をスクリーニングする方法を提供する。この方法は、(a) このポリペプチドを好適な条件下で少なくとも1つの化合物と結合させるステップと、(b) このポリペプチドとこの試験化合物との結合を検出して、このポリペプチドと特異的に結合する化合物を同定するステップとを含む。

30

【0106】

更に本発明は、(a) SEQ ID NO: 1-10からなる一群から選択されたアミノ酸配列を含むポリペプチドと、(b) SEQ ID NO: 1-10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アミノ酸配列を含む天然のポリペプチドと、(c) SEQ ID NO: 1-10からなる一群から選択されたアミノ酸配列を含むポリペプチドの生物学的に活性な断片と、(d) SEQ ID NO: 1-10とからなる一群から選択されたアミノ酸配列を含むポリペプチドの免疫原性断片とで構成される一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。このスクリーニング方法は、(a) このポリペプチドを、その活性が許容される条件下で少なくとも1つの化合物と結合させるステップと、(b) この試験化合物の存在下でのこのポリペプチドの活性を評価するステップと、(c) この試験化合物の存在下で

40

50

のこのポリペプチドの活性と、この試験化合物の不在下でのこのポリペプチドの活性とを比較するステップとを含み、この試験化合物の存在下でのこのポリペプチドの活性の変化が、このポリペプチドの活性を調節する化合物の存在を示唆するという特徴を有する。

【0107】

更に本発明は、SEQ ID NO: 11-20 からなる一群から選択された配列を含む標的ポリヌクレオチドの発現を変化させるのに効果的な化合物をスクリーニングする方法であって、(a) この標的ポリヌクレオチドを含むサンプルを化合物に曝露するステップと、(b) この標的ポリヌクレオチドの発現の変化を検出するステップとを含む、該スクリーニング方法を提供する。

【0108】

本発明はさらに、試験化合物の毒性を評価する方法を提供する。この方法は、(a) 核酸を含む生体サンプルを前記試験化合物で処理するステップと、(b) 処理した前記生体サンプルの核酸をプローブとハイブリダイズするステップと、(c) ハイブリダイゼーション複合体の収量を測定するステップと、(d) 前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量を、未処理の生体サンプルにおけるハイブリダイゼーション複合体の収量とを比較するステップとを含み、前記処理した生体サンプルにおけるハイブリダイゼーション複合体の収量の差異が試験化合物の毒性を示唆する。この方法における前記プローブは、(1) SEQ ID NO: 11-20 からなる一群から選択されたポリヌクレオチド配列と、(2) SEQ ID NO: 11-20 からなる一群から選択されたポリヌクレオチド配列と少なくとも90%の配列同一性を有する天然のポリヌクレオチド配列と、(3) 前記(1)に相補的なポリヌクレオチド配列と、(4) 前記(2)に相補的なポリヌクレオチド配列と、(5) 前記(1)乃至(4)のRNA等価物とで構成される一群から選択されたポリヌクレオチド配列を含むポリヌクレオチドの連続する少なくとも20個のヌクレオチドを含む。また、前記ハイブリダイゼーションは、前記プローブと前記生体サンプルの標的ポリヌクレオチドとの間で特異的なハイブリダイゼーション複合体が形成される条件下で行われる。また、前記標的ポリヌクレオチドが、(1) SEQ ID NO: 11-20 からなる一群から選択されたポリヌクレオチド配列と、(2) SEQ ID NO: 11-20 からなる一群から選択されたポリヌクレオチド配列と少なくとも90%の配列同一性を有する天然のポリヌクレオチド配列と、(3) 前記(1)に相補的なポリヌクレオチド配列と、(4) 前記(2)に相補的なポリヌクレオチド配列と、(5) 前記(1)乃至(5)のRNA等価物とを含む。代替的に前記標的ポリヌクレオチドは前記ポリヌクレオチド配列の断片である。

【0109】

(本発明の記載について)

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

【0110】

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

【0111】

本明細書で用いた全ての科学技術用語は、別の方法で定義されていない限り、本発明の属する技術分野の一般的な技術者が普通に解釈する意味と同じである。本明細書で記述したものと類似、或いは同等の全ての装置及び材料、方法は本発明の実施及びテストに使用できるが、好適な装置及び材料、方法をここに記す。本明細書に記載の全ての文献は、本発明に関連して使用する可能性のある文献に記載された細胞系、プロトコル、試薬、ベクタ

10

20

30

40

50

ーを記述し開示するために引用した。従来の発明を引用したからと言って、本発明の新規性が損なわれると解釈されるものではない。

【0112】

(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。

【0113】

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

【0114】

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

20

【0115】

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アレル変異配列との不適當或いは予期しないハイブリダイゼーション、並びに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形性を含む。コードされたタンパク質も変異され得り、サイレント変化を生じ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、負に荷電したアミノ酸にはアスパラギン酸及びグルタミン酸が含まれ、正に荷電したアミノ酸にはリシン及びアルギニンが含まれ得る。類似の親水性の値をもち極性非荷電側鎖を有するアミノ酸には、アスパラギン、グルタミン、セリン、トレオニンが含まれ得る。類似の親水性の値をもち非荷電側鎖を有するアミノ酸には、ロイシン、イソロイシン、バリン、グリシン、アラニン、フェニルアラニン及びチロシンが含まれ得る。

30

【0116】

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

40

【0117】

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

【0118】

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。

50

【0119】

用語「抗体」は、抗原決定基と結合可能な Fab 及び F(ab')₂、及びそれらの断片、Fv断片などの無傷の分子を指す。DMEポリペプチドと結合する抗体は、抗体を免疫する小ペプチドを含む無傷の分子またはその断片を用いて作製可能である。動物（例えば、マウス、ラット、若しくはウサギ）を免疫化するのに使用されるポリペプチド或いはオリゴペプチドは、RNAの翻訳から、或いは化学的に合成可能であり、必要に応じて担体タンパク質と結合させることも可能である。ペプチドと化学的に結合した一般に用いられる担体は、ウシ血清アルブミン、チログロブリン、及びキーホールリンペットヘモニアン（KLH）を含む。次ぎに、この結合したペプチドを用いて動物を免疫化する。

【0120】

10

用語「抗原決定基」は、特定の抗体と接触する分子の領域（即ちエピトープ）を指す。タンパク質或いはタンパク質の断片が、宿主動物を免疫化するのに用いられるとき、このタンパク質の種々の領域は、抗原決定基（タンパク質上の特定の領域或いは三次元構造体）に特異的に結合する抗体の産生を誘発し得る。抗原決定基は、抗体と結合するために無傷の抗原（即ち、免疫応答を引き出すために用いられる免疫原）と競合し得る。

【0121】

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

20

【0122】

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

30

【0123】

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

【0124】

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

40

【0125】

「コンセンサス配列」は、不要な塩基を分離するためにDNA配列の解析を繰り返し行い、XL-PCRキット（PE Biosystems, Foster City CA）を用いて5'及び／または3'の方向に伸長され、再度シーケンシングされた核酸配列、

50

または G E L V I E W 断片構築システム (G C G , M a d i s o n , W I) または P h r a p (U n i v e r s i t y o f W a s h i n g t o n , S e a t t l e W A) 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 c D N A や E S T、またはゲノム D N A 断片から構築された核酸配列を指す。伸長及び重複の両方によって構築されるコンセンサス配列もある。

【0126】

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

元の残基

保存的な置換

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

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

【0127】

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

【0128】

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

【0129】

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

【0130】

用語「断片」は、D M E または D M E 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 (p a r e n t s e q u e n c e) と同一であるがその配列より長さが

10

20

30

40

50

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

10

【0131】

SEQ ID NO: 11-20の断片は、例えば、この断片を得たゲノム内の他の配列とは異なる、SEQ ID NO: 11-20を明確に同定する固有のポリヌクレオチド配列の領域を含む。SEQ ID NO: 11-20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: 11-20に関連ポリヌクレオチド配列から区別する類似の方法に有用である。ある断片と一致するSEQ ID NO: 11-20の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に測定できる。

【0132】

SEQ ID NO: 1-10のある断片は、SEQ ID NO: 11-20のある断片によってコードされる。SEQ ID NO: 1-10のある断片は、SEQ ID NO: 1-10を特異的に同定する固有のアミノ酸配列の領域を含む。例えば、SEQ ID NO: 1-10のある断片は、SEQ ID NO: 1-10を特異的に認識する抗体の開発における免疫原性ペプチドとして有用である。ある断片と一致するSEQ ID NO: 1-10の正確な断片の長さや領域は、その断片の目的に基づいて当分野で一般的な技術によって日常的に決定できる。

20

【0133】

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

30

【0134】

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

【0135】

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

【0136】

ポリヌクレオチド配列間の同一性のパーセントは、MEGALIGN version 3.12e配列アラインメントプログラムに組込まれるCLUSTAL Vアルゴリズムのデフォルトパラメータを用いて決定可能である。このプログラムはLASERGENEソフトウェアパッケージの一部であり、分子生物学分析プログラム一式(DNA STAR, Madison WI)である。このCLUSTAL Vは、Higgins, D. G. 及び P. M. Sharp (1989) CABIOS 5:151-153、Higgins, D. G. 他 (1992) CABIOS 8:189-191に記載されている。ポリヌクレオチド配列の対のアラインメントの場合、デフォルトパラメータは、Ktuple=2、gap penalty=5、window=4、「diagonals saved」=4と設定する。「重み付けされた」残基重み付け表が、デフォルトとして選

40

50

択された。同一性のパーセントは、アラインメントされたポリヌクレオチド配列間の「類似性のパーセント」としてCLUSTAL Vによって報告される。

【0137】

別法では、National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S. F. 他 (1990) J. Mol. Biol. 215:403-410) が提供する、広く用いられている無料の配列比較アルゴリズム一式が、NCBI (Bethesda, MD) を含む幾つかのソース及びインターネット (<http://WWW.ncbi.nlm.nih.gov/BLAST/>) で入手可能である。このBLASTソフトウェア一式には、既知のポリヌクレオチド配列と様々なデータベースの別のポリヌクレオチド配列とのアラインメントに用いられる「blastn」を含む、様々な配列分析プログラムが含まれる。「BLAST 2 Sequences」と呼ばれるツールが入手可能であり、2つのヌクレオチド配列の対を直接比較するために用いられる。「BLAST 2 Sequences」は、<http://WWW.ncbi.nlm.nih.gov/gorf/bl2.html> にアクセスして、対話形式で利用できる。「BLAST 2 Sequences」ツールは、blastn 及び blastp (以下に記載) の両方に用いることができる。BLASTプログラムは、一般的には、デフォルトを設定するギャップ及び他のパラメータと共に用いられる。例えば、2つのヌクレオチド配列を比較する場合、ある者は、デフォルトパラメータに設定された「BLAST 2 Sequences」ツールVersion 2.0.12 (April-21-2000) でblastnを使用するであろう。そのようなデフォルトパラメータは、例えば、以下のようにする。

10

20

【0138】

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

30

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

【0139】

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

40

【0140】

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

【0141】

ポリペプチド配列間の同一性のパーセントは、MEGALIGN バージョン3.12e

50

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

【0142】

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

【0143】

Matrix: BLOSUM62

Open Gap: 11 及び Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

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

【0144】

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

【0145】

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

【0146】

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

【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、洗浄過程を行う際の温度によ

10

20

30

40

50

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

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

10

20

【0149】

用語「ハイブリダイゼーション複合体」は、相補的な塩基間の水素結合によって、形成された2つの核酸配列の複合体を指す。ハイブリダイゼーション複合体は溶液中（例えば、 C_0t または R_0t 分析）で形成されるか、或いは溶液中の1つの核酸配列と固体の支持物（例えば、紙、膜、フィルタ、チップ、ピン、或いはスライドガラス、または細胞及びその核酸を固定する任意の適当な基板）に固定されたもう一つの核酸配列とで形成され得る。

【0150】

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

30

【0151】

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

【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DME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DME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】

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

40

【0154】

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

【0155】

用語「調節」は、DMEの活性の変化を指す。例えば、調節によって、DME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。

【0156】

50

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

【0157】

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

10

【0158】

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

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

30

【0161】

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

【0162】

プローブ及びプライマーの準備及び使用方法については、例えば、Sambrook, J. 他による、1989年、名称「Molecular Cloning: A Laboratory Manual」、第2版の1-3巻（Cold Spring Harbor Press, Plainview NY）、またはAusubel, F. M. 他による、1987年、名称「Current Protocols in Molecular Biology」（Greene Publi. Assoc. & Wiley-Intersciences, New York NY）、並びにInnis他による、1990年、名称「PCR Protocols, A Guide to Methods and Applications」（Academic Press, San Diego CA.）を参照されたい。PCR用のプライマーの組は、例えば、Primer（Ver

40

50

sion 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA) などのそのような目的のためのコンピュータプログラムを用いて、ある既知の配列から引き出すことができる。

【0163】

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

10

20

30

【0164】

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

40

【0165】

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

【0166】

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

50

【0167】

「レポーター分子」は、核酸、アミノ酸または抗体の標識に用いられる化学的または生化学的な部分である。レポーター分子には、放射性核種、酵素、蛍光剤、化学発光剤、発色剤、基質、補助因子、インヒビター、磁気粒子及びその他の当分野で既知の成分が含まれる。

【0168】

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

【0169】

用語「サンプル」は、その最も広い意味で用いられている。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、cDNAと、組織と、組織プリント等を含み得る。

10

【0170】

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

20

【0171】

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

【0172】

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

30

【0173】

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

【0174】

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

【0175】

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

40

50

【0176】

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

10

【0177】

特定の核酸配列の「変異配列」とは、デフォルトパラメータ設定の「BLAST 2 Sequences」ツール Version 2.0.9 (May-07-1999) を用いる *blastn* によって、ある核酸配列のある長さに対する該特定の核酸配列の同一性が、少なくとも40%と決定された核酸配列のことである。このような核酸の対は、ある長さにおいて、例えば、少なくとも50%または60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或いはそれ以上の同一性を示し得る。ある変異配列は、例えば、「アレル」変異配列 (上述) または「スプライス」変異配列、「種」変異配列、「多型」変異配列と表すことができる。スプライス変異配列は基準分子と同一性が極めて高い可能性があるが、mRNA プロセッシング中のエキソンの択一的スプライシングによってポリヌクレオチドの数が多くなったり、少なくなったりする。対応するポリペプチドは、基準分子に存在する追加の機能ドメインを有したり、基準分子に存在するドメインが欠落したりし得る。種変異配列は、種によって異なるポリヌクレオチド配列である。得られるポリペプチドは、互いに高いアミノ酸同一性を有する。多型変異配列は、所定の種と種における特定の遺伝子のポリヌクレオチド配列が異なる。多型変異配列はまた、ポリヌクレオチド配列の1つのヌクレオチドが異なる「1ヌクレオチド多型」(SNP) も含み得る。SNPの存在は、例えば、或る集団、病態、病態の性向を示唆し得る。

20

30

【0178】

特定のポリペプチド配列の「変異体」とは、デフォルトパラメータ設定の「BLAST 2 Sequences」ツール Version 2.0.9 (May-07-1999) を用いる *blastp* によって、あるポリペプチド配列のある長さに対する該特定のポリペプチド配列の同一性が、少なくとも40%と決定されたポリペプチド配列のことである。このようなポリペプチドの対は、ある長さにおいて、例えば、少なくとも50%または60%、70%、80%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、或いはそれ以上の同一性を示し得る。

【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。

40

【0180】

表1は、本発明のポリヌクレオチド配列及びポリペプチド配列の識別番号を示す。各ポリヌクレオチドおよびそれに対応するポリペプチドは、1つのインサイトプロジェクト識別番号 (Incyte Project ID) に相関する。各ポリペプチド配列は、記載されているようにポリペプチド配列識別番号 (Polypeptide SEQ ID NO :) およびインサイトポリペプチド配列番号 (Incyte Polypeptide I

50

D) の両方によって示されている。各ポリヌクレオチド配列は、記載されているようにポリヌクレオチド配列識別番号 (Polynucleotide SEQ ID NO :) およびインサイトポリヌクレオチドコンセンサス配列番号 (Incysite Polynucleotide ID) の両方によって示されている。

【0181】

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

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インサイトポリペプチド配列番号 (Incysite Polypeptide ID) を示す。列3は、各ポリペプチドのアミノ酸残基数を示す。列4および列5はそれぞれ、GCG配列分析ソフトウェアパッケージのMOTIFSプログラム (Genetics Computer Group, Madison WI) によって決定された、潜在的なリン酸化部位および潜在的なグリコシル化部位を示す。列6は、シグネチャ (signature) 配列、ドメイン、およびモチーフを含むアミノ酸残基を示す。列7は、タンパク質の構造/機能の分析のための分析方法を示し、該当箇所にはさらに分析方法に利用した検索可能なデータベースを示す。

20

【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一覧にしたものであって、これらの特性は請求するポリペプ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 : 9は、ヒト推定N-アセチルトランスフェラーゼCamel1102 (GenBank ID g6651438) と32%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を参照)。BLASTの確率スコアは $6.50E-23$ であり、探しているポリペプチド配列アラインメントが偶然の一致により得られる確率を示す。SEQ ID NO : 9はまた、SPScanプログラムによって決定されたシグナルペプチドサイン配列、及び保存されたタンパク質ファミリードメインの隠れマルコフモデル (HMM) ベースのPFAMデータベースにおける統計的に有意なマッチを検索して決定されたアセチルトランスフェラーゼを有する。アセチルトランスフェラーゼに対するHMMの確率は、 $3.2e-12$ である (図3参照)。BLAST及びHMM解析を基にすると、SEQ ID NO 9のタンパク質はN-アセチルトランスフェラーゼである (表7を参照)。これらサイン配列のすべては、統計学的に高度に有意であり (しきい値については表7を参照)、例えば、アルド-ケト レダクターゼファミリーサイン配列に対するHMMの確率値は、 $5.6e-170$ である。BLAST、HMM、BLIMPS、MOTI、及びProfileScan解析を基に、SEQ ID NO 10のタンパク質がアルド/ケト レダクターゼであり、これに限定する物ではないが、グルコース及びカルボニル含有薬剤分子及びゼノバイオティックスを有するカルボニル含有糖及び芳香族化合物を減少させる。SEQ ID NO : 1-8は、同様の規則で解析され、アノテーションを付けた。SEQ ID NO : 1-10のアルゴリズム及びパラメーターは表7に記載されている。

30

40

【0184】

表4に示されているように、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、或いはこれらの2種類の配列のあらゆる

50

る組み合わせを用いて組み立てた。列1および列2はそれぞれ、本発明の各ポリヌクレオチドのポリヌクレオチド配列識別番号 (Polynucleotide SEQ ID NO :) およびそれに対応するインサイトポリヌクレオチドコンセンサス配列番号 (In cyte Polynucleotide ID) を示す。列3は、塩基対における各ポリヌクレオチド配列の長さを示す。列4は、例えば、SEQ ID NO : 11-20を同定するため、或いはSEQ ID NO : 11-20と関連するポリヌクレオチド配列とを区別するためのハイブリダイゼーションまたは増幅技術に有用なポリヌクレオチド配列の断片を示す。列5は、cDNA配列、ゲノムDNAから推定されるコード配列 (エキソン)、および/またはcDNAおよびゲノムDNAの両方からなる群に対応する識別番号を示す。これらの配列を用いて本発明の完全長ポリヌクレオチド配列を組み立てた。表4の列6および列7はそれぞれ、列5の配列に対応するcDNA配列およびゲノム配列の開始ヌクレオチド (5') 位置および終了ヌクレオチド (3') 位置を示す。 10

【0185】

表4の列5に示されている識別番号は、具体的には、例えばインサイトcDNAおよびそれらに対応するcDNAライブラリの識別番号を示す。例えば、3700065F6はインサイトcDNA配列の識別番号であり、SININOT05はそれが由来するcDNAライブラリの識別番号である。cDNAライブラリが示されていないインサイトcDNAは、プールされているcDNAライブラリ (例えば、71447910V1) に由来する。または、列5の識別番号は、ポリヌクレオチド配列の組み立てに用いたGenBankのcDNAすなわちEST (例えば、g5663459) の識別番号の場合もある。また 20
は、列5の識別番号は、Genscan分析によって推定されるゲノムDNAのコード領域の場合もある。例えば、GNN. g7025744_000055_002は、Genscan推定コード配列の識別番号であって、g7025744がGenscan分析によって得られたGenBankの配列の識別番号である。このGenscan推定コード配列は、配列を組み立てる前に編集する場合がある (実施例4を参照)。または、列5の識別番号は、"exon-stitching" アルゴリズムによってcDNAおよびGenscan推定エキソンの両方からなる群の場合もある。例えば、FL152824_00001はステッチ配列 ("stitched" sequence) であって、この152824はアルゴリズムに用いた配列のクラスター識別番号を表し、00001はアルゴリズムによって推定された数を表す (実施例5を参照)。または、列5の識別番号は、" 30
exon-stretching" アルゴリズムによってcDNAおよびGenscan推定エキソンの両方からなる群の場合もある (実施例5を参照)。場合によっては、列5に示されている配列の範囲と重複するインサイトcDNAの範囲が得られ、最終的なコンセンサス配列が決定されるが、それに相当するインサイトcDNAの識別番号は示されていない。

【0186】

表5は、インサイトcDNA配列を用いて組み立てられたこれらの完全長ポリヌクレオチド配列が由来する代表的なcDNAライブラリを示す。代表的なcDNAライブラリとは、上記ポリヌクレオチド配列の組み立ておよび決定に用いられたインサイトcDNA配列を最も多く含むインサイトcDNAライブラリのことである。表5に示されているcDN 40
Aライブラリを作製するために用いた組織およびベクターが表6に示されている。

【0187】

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。

【0188】

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 : 11-20からなる一群から選択された配列を含むポリヌクレオチド配列を提供する。配列表に示したSEQ ID NO : 1 50

1-20のポリヌクレオチド配列は、窒素系塩基のチミンがウラシルに置換され、糖鎖の背骨がデオキシリボースではなくリボースからなる等価RNA配列を含む。

【0189】

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と少なくとも70%のポリヌクレオチド配列同一性、或いは少なくとも85%のポリヌクレオチド配列同一性、更には少なくとも95%ものポリヌクレオチド配列同一性を有する。本発明の特定の実施形態は、SEQ ID NO: 11-20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: 11-20からなる一群から選択された配列を含むポリヌクレオチド配列の変異配列を提供する。上記したポリヌクレオチド変異配列は何れも、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。

10

【0190】

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、可能なコドン選択に基づいた組み合わせの選択によって作り出され得る可能なポリヌクレオチド配列の変異の全てが含まれ得る。これらの組み合わせは、天然の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

20

【0191】

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含む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。特定のコドンが宿主によって利用される頻度に基づいてコドンを選択して、ペプチドの発現が特定の真核細胞または原核宿主に発生する割合を高めることが可能である。コードされたアミノ酸配列を変えないで、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

【0192】

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。

【0193】

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: 11-20及びそれらの断片とハイブリダイズ可能なポリヌクレオチド配列が含まれる（例えば、Wahl, G. M. 及びS. L. Berger (1987) *Methods Enzymol.* 152: 399-407; and Kimmel, A. R. (1987) *Methods Enzymol.* 152: 507-511. を参照）。アニーリング及び洗浄条件を含むハイブリダイゼーションの条件は、「定義」に記載されている。

40

【0194】

当分野で周知のDNAのシーケンシング方法を用いて、本発明の何れの実施例も実行可能である。この方法には、例えばDNAポリメラーゼIのクレノウ断片、SEQUENASE (US Biochemical, Cleveland OH)、Taqポリメラーゼ (Applied Biosystems)、熱安定性T7ポリメラーゼ (Amers

50

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

10

【0195】

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

20

30

40

【0196】

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

【0197】

市販の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産物のヌ

50

クレオチド配列のサイズの分析、または確認が可能である。詳しくは、キャピラリーシーケンシングには、電気泳動による分離のための流動性ポリマー、及び4つの異なったヌクレオチドに特異的なレーザーで活性化される蛍光色素、放出された波長の検出に利用するCCDカメラを使用することが可能である。出力／光強度は、適切なソフトウェア（例えば、GENOTYPER及びSEQUENCE NAVIGATOR、PE Biosystems）を用いて電気信号に変換され、サンプルのローディングからコンピュータ分析までのプロセス及び電子データ表示がコンピュータ制御可能である。キャピラリー電気泳動法は、特定のサンプルに少量しか存在しない場合もあるDNAの小片のシーケンシングに特に適している。

【0198】

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。

【0199】

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。ランダムな断片によるDNAの混合や遺伝子断片と合成オリゴヌクレオチドのPCR再組み立てを用いて、ヌクレオチド配列の組換えが可能である。例えば、オリゴヌクレオチド媒介性定方向突然変異誘発を利用して、新しい制限部位を生成する突然変異の導入、グリコシル化パターンの変更、コドン選択の変更、スプライスバリエーションの作製等が可能である。

【0200】

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

【0201】

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成することが可能である。例えば、ペプチド合成は種々の固相技術を用いて実行可能である（例えば、Creighton, T. (1984) Proteins. Struc

10

20

30

40

50

t u r e s a n d M o l e c u l a r P r o p e r t i e s, W H F r e e m a n, New York NY, pp. 55-60; R o b e r g e, J. Y. ら (1995) S c i e n c e 269:202-204を参照)。また、合成の自動化は例えば A B I 431A ペプチドシンセサイザー (P E B i o s y s t e m s) を用いて達成し得る。更に D M E 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。

【0202】

このペプチドは、分離用高速液体クロマトグラフィー (例えば、C h i e z, R. M. 及び F. Z. R e g n i e r (1990) M e t h o d s E n z y m o l. 182:392-421を参照) を用いて実質的に精製可能である。合成されたペプチドの組成は、アミノ酸分析或いはシーケンシングにより確認することができる (例えば、C r e i g h t o n、前出、pp28-53を参照)。

【0203】

生物学的に活性な D M E を発現させるために、D M E 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 D M E をコードするポリヌクレオチド配列におけるエンハンサー、構成型及び発現誘導型のプロモーター、5' 及び 3' の非翻訳領域などの調節配列が含まれる。このようなエレメントは、その長さ及び特異性が様々である。特定の開始シグナルによって、D M E 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、A T G 開始コドン及びコザック配列などの近傍の配列が含まれる。D M E 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 A T G 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なければならない。外来性の翻訳要素及び開始コドンは、自然及び合成の様々なものから得ることが可能である。用いられる特定の宿主細胞系に好適なエンハンサーを含めることで発現の効率を高めることが可能である。(例えば、S c h a r f, D. 他 (1994) R e s u l t s P r o b l. C e l l D i f f e r. 201-18-162. を参照)。

【0204】

当業者に周知の方法を用いて、D M E 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、i n v i t r o 組換え DNA 技術、合成技術、及び i n v i v o 遺伝子組換え技術が含まれる。(例えば、S a m b r o o k, J. 他. (1989) M o l e c u l a r C l o n i n g. A L a b o r a t o r y M a n u a l, C o l d S p r i n g H a r b o r P r e s s, P l a i n v i e w N Y, 4 章及び 8 章, 及び 16-17 章; 及び A u s u b e l, F. M. 他. (1995) C u r r e n t P r o t o c o l s i n M o l e c u l a r B i o l o g y, J o h n W i l e y & S o n s, N e w Y o r k N Y, c h. 9 章及び 13 章 1-4 章を参照)。

【0205】

種々の発現ベクター／宿主系を利用して、D M E をコードする配列の保持及び発現が可能である。これらには、限定するものではないが、組換えバクテリオファージ、プラスミド、またはコスミド DNA 発現ベクターで形質転換された細菌などの微生物や、酵母菌発現ベクターで形質転換された酵母菌や、ウイルス発現ベクター (例えば、バキュロウイルス) に感染した昆虫細胞系や、ウイルス発現ベクター (例えば、カリフラワーモザイクウイルス、C a M V; タバコモザイクウイルス、T M V) または細菌発現ベクター (例えば、T i または p B R 3 2 2 プラスミド) で形質転換された植物細胞系や、動物細胞系などが

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

【0206】

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、これらのベクターを用いて、クローニングされた配列の in vitroでの転写、ジデオキシンスクリーニング、ヘルパーファージによる一本鎖の救出、入れ子状態の欠失を作り出すことが可能である (例えば、Van Heeke, G. 及び S. M. Schuster (1989) J. Biol. Chem. 264:5503-5509. を参照)。例えば、抗体の産生のためなどに多量の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、強力に発現を誘発するT5またはT7バクテリオファージプロモーターを含むベクターを使用できる。

【0207】

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

【0208】

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、例えば、CaMV由来の35S及び19Sプロモーターなどのウイルスプロモーターが単独で、或いはTMV (例えば、Coruzzi, G. ら. (1984) EMBO J. 3:1 50

671-1680 ; Broglie, R. ら (1984) Science 224 : 838-843 ; および Winter, J. ら (1991) Results Probl. Cell Differ. 17 : 85-105を参照) 由来のオメガリーダー配列と組み合わせて促進される。これらの作製物は、直接のDNA形質転換或いは病原体を介したトランスフェクションによって、植物細胞の中に導入可能である。(例えば、The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill NY, pp. 191-196を参照)。

【0209】

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

10

【0210】

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

20

【0211】

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の上の選択マーカー遺伝子を含む。ベクターの導入の後、細胞を選択培地に移す前に、強化培地で約1~2日の間増殖させる。選択マーカーの目的は選択的な媒介物に対する抵抗性を与えるとともに、その存在により導入された配列を確実に発現する細胞の増殖及び回収が可能となる。安定的に形質転換された細胞の耐性クローンは、その細胞型に好適な組織培養技術を用いて増殖可能である。

30

【0212】

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

40

50

を変える *trpB* 及び *hisD* が文献に記載されている（例えば、Hartman, S. C. 及び R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51 を参照）。アミノシアニン、緑色蛍光タンパク質 (GFP; Clontech)、 β グルクロニダーゼ及びその基質 GUS、ルシフェラーゼ及びその基質ルシフェリンなどの可視マーカーが用いられる。緑色蛍光タンパク質 (GFP) (Clontech, Palo Alto, CA) も使用できる。これらのマーカーを用いて、トランスフォーマントを特定するだけでなく、特定のベクター系に起因する一過性或いは安定したタンパク質発現を定量することが可能である（例えば、Rhodes, C. A. 他 (1995) *Methods Mol. Biol.* 55:121-131 を参照）。

10

【0213】

マーカー遺伝子の発現の存在／不在によって目的の遺伝子の存在が示されても、その遺伝子の存在及び発現の確認が必要な場合もある。例えば、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特定可能である。または、1つのプロモーターの制御下でマーカー遺伝子が 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。

【0214】

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。これらの方法には、DNA-DNA 或いは DNA-RNA ハイブリダイゼーションや、PCR 法、核酸或いはタンパク質の検出及び／または数量化のための膜系、溶液ベース、或いはチップベースの技術を含むタンパク質生物学的試験法または免疫学的アッセイが含まれるが、これらに限定されるものではない。

20

【0215】

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)、蛍光標示式細胞分取器 (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その他のアッセイは、当分野では十分に知られている。（例えば、Hampton, R. 他 (1990) *Serological Methods, a Laboratory Manual*, APS Press, St Paul, MN, Sect. IV; Coligan, J. E. 他 *Current Protocols in Immunology*, Greene Pub. Associates and Wiley-Interscience, New York, NY; 及び Pound, J. D. (1990) *Immunochemical Protocols*, Humana Press, Totowa NJ）。

30

40

【0216】

種々の標識技術及び結合技術が当業者には周知であり、様々な核酸アッセイおよびアミノ酸アッセイに用いられ得る。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。当分野では周知であり市販されているこのようなベクターを、T7, T3, または SP6 などの好適な RNA ポリメラーゼ及び標識されたヌクレオチドの追加によって、*in vitro* で RNA プローブの合成に用いることができる。これらの方

50

法は、例えば、Amersham Pharmacia Biotech及びPromega (Madison WI)、U. S. Biochemical Corp (Cleveland OH) が市販する種々のキットを用いて行うことができる。容易な検出のために用い得る好適なレポーター分子或いは標識には、基質、コファクター、インヒビター、磁気粒子、及び放射性核種、酵素、蛍光剤、化学発光剤、色素産生剤などが含まれる。

【0217】

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。形質転換された細胞から産生されたタンパク質が分泌されるか細胞内に留まるかは、使用されるその配列及び／またはそのベクターによる。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。

10

【0218】

更に、挿入した配列の発現調節能力または発現したタンパク質を所望の形にプロセッシングする能力によって宿主細胞株が選択される。このようなポリペプチドの修飾には、アセチル化、カルボキシル化、グリコシル化、リン酸化、脂質化 (lipidation)、及びアシル化が含まれるが、これらに限定されるものではない。タンパク質の「プレプロ」または「プロ」形を切断する翻訳後のプロセッシングを利用して、標的タンパク質、折りたたみ及び／または活性を特定することが可能である。翻訳後の活性のための特定の細胞装置及び特徴のある機構をもつ種類の宿主細胞 (例えば、CHO、HeLa、MDCK、MEK293、WI38) がAmerican Type Culture Collection (ATCC; Bethesda, MD) より入手可能であり、外来のタンパク質の正しい修飾及びプロセッシングを確実にするために選択される。

20

【0219】

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が、市販の親和性基質を用いて融合タンパク質の精製を促進し得る。このような部分には、グルタチオンSトランスフェラーゼ (GST)、マルトース結合タンパク質 (MBP)、チオレドキシン (Trx)、カルモジュリン結合ペプチド (CBP)、6-His、FLAG、c-mc、赤血球凝集素 (HA) が含まれるが、これらに限定されるものではない。GST及びMBP、Trx、CBP、6-Hisによって、固定されたグルタチオン、マルトース、フェニルアルシン酸化物 (phenylarsine oxide)、カルモジュリン、金属キレート樹脂のそれぞれで同族の融合タンパク質の精製が可能となる。FLAG、c-mc、及び赤血球凝集素 (HA) によって、これらのエピトープ標識を特異的に認識する市販のモノクローナル抗体及びポリクローナル抗体を用いた融合タンパク質の免疫親和性の精製ができる。また、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と精製の方法は、Ausubel, (1995、前出 ch 10) に記載されている。市販されている様々なキットを用いて、融合タンパク質の発現及び精製を促進できる。

30

40

【0220】

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またはT3、SP6プロモーターと機能的に結合したタンパク質をコードする配列の転写と翻訳をつなげる。翻訳は、例えば、³⁵Sメチオニンである放射能標識されたアミノ酸前駆体の存在の下で起こる。

【0221】

50

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。

【0222】

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（Colligan, 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、このような化合物に対するスクリーニングには、分泌タンパク質或いは細胞膜上のタンパク質の何れか一方として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、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

【0223】

あるアッセイは、単に試験化合物をポリペプチドに実験的に結合させ、結合を、フルオロフォア、放射性同位体、酵素抱合体またはその他の検出可能な標識により検出することができる。例えば、このアッセイは、少なくとも1つの試験化合物を、溶液中の或いは固体支持物に固定された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、細胞遊離剤、化学ライブラリまたは天然の生成混合物を用いて実施することができ、試験化合物は、溶液中で遊離させるか固体支持体に固定させる。

20

【0224】

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、或るいは部分的または逆アゴニスト等が含まれる。一実施例では、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の活性が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*または細胞遊離系と結合させてアッセイを実施する。これらアッセイの何れかにおいて、DMEの活性を調整する試験化合物は間接的に結合することが可能であり、試験化合物と直接接触する必要がない。少なくとも1つから複数の試験化合物をスクリーニングすることができる。

30

【0225】

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

40

50

c l e i c A c i d s R e s . 2 5 : 4 3 2 3 - 4 3 3 0) 。形質転換した E S 細胞を同定し、例えば C 5 7 B L / 6 マウス系等から採取したマウス細胞胚盤胞に微量注入する。胚盤胞を偽妊娠メスに外科的に導入し、得られるキメラ子孫の遺伝形質を決め、これを交配させてヘテロ接合性系またはホモ接合性系を作製する。このようにして作製した遺伝子組換え動物は、潜在的な治療薬や毒性薬剤で検査することができる。

【0226】

DME をコードするポリヌクレオチドを *i n v i t r o* でヒト胚盤胞由来の E S 細胞において操作することが可能である。ヒト E S 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 (T h o m s o n , J . A . 10
他 (1 9 9 8) S c i e n c e 2 8 2 : 1 1 4 5 - 1 1 4 7) 。

【0227】

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 E S 細胞に注入し、注入した配列を動物細胞ゲノムに組み込ませる。形質転換細胞を胞胚に注入し、胞胚を上記のように移植する。遺伝子組換え子孫または近交系について研究し、潜在的な医薬品を用いて処理し、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 を過剰に発現する哺乳動物近交系は、便利なタンパク質源となり得る (J a n n e , J . 他 (1 9 9 8) B i o t e c h n o l 20
. A n n u . R e v . 4 : 5 5 - 7 4) 。

【0228】

(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文脈における化学的及び構造的類似性が存在する。更に、DME の発現は、肝臓腫瘍細胞株、肝臓腫瘍組織、膵臓組織、脳下垂体腫瘍、脳腫瘍、小腸細胞、胎児の脳組織、及び不等皮質脳組織に密接に関連する。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の増大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、DM 30
E 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増大させることが望ましい。

【0229】

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、細胞増殖異常の異常、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 (A I D S) 及び副腎機能不全、成人呼吸窮迫症候群、アレルギー、強直性脊椎炎、アミロイド症、貧血、喘息、アテローム性動脈硬化症、自己免疫性溶血性貧血、自己免疫 40
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 (A P E C E D) 、気管支炎、胆嚢炎、接触皮膚炎、クローン病、アトピー性皮膚炎、皮膚筋炎、糖尿病、肺気腫、リンパ球毒素性一時性リンパ球減少症、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、膵炎、乾癬、ライター症候群、リウマチ様関節炎、強皮症、シェーグレン症候群、全身性アナフィラキシー、全身性エリテマトーデス、全身性硬化症、原発性血小板血症、血小板減少症、潰瘍性大腸炎、ウェルナー症候群、癌合併症、血液透析、体外循環、ウイルス感染症、細菌感染症、真菌感染症、寄生虫感染症、原虫感染症、蠕虫感染症、外傷が含まれ、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 50

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、脾臓、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-トゥース病及び神経線維腫症などの遺伝性神経病、10
 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-クリスチャン病、レトラ-シヴェ病、サルコイドーシス、エンプティセラ症候群、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適当抗利尿ホルモン（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、20
 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、30
 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、5α-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炎、色素性網膜炎、脈絡膜黒色腫、球後腫瘍、交叉腫瘍（chiasmatal tumor）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、40
 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、低カルシウム血症、低リン酸血症、思春期後小脳性運動失調症、チロシン血症が含まれ、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 50

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、 α -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、子癰前症、子癰、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。

【0230】

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、DME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

【0231】

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、実質的に精製されたDME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。

【0232】

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、DME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。

【0233】

更なる実施例では、DME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DME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、DME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DMEを発現する細胞または組織に薬剤を運ぶターゲッティング或いは運搬機構として間接的に用いられ得る。 20

【0234】

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DME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、DME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

【0235】

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

【0236】

DME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。 40
詳しくは、精製されたDME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DMEと特異的に結合するものを同定が可能である。DMEの抗体も、当分野で一般的な方法を用いて製造することが可能である。このような抗体には、ポリクローナル抗体、モノクローナル抗体、キメラ抗体、一本鎖、Fabフラグメント、及びFab発現ライブラリによって作られたフラグメントが含まれる。但し、これらに限定されるものではない。治療用には、中和抗体（即ち、二量体の形成を阻害するもの）が特に好ましい。

【0237】

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

【0238】

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或いはペプチド、またはそれらの断片は、天然のタンパク質のアミノ酸配列の一部と同一であることが望ましい。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

【0239】

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。これらの技術には、ハイブリドーマ技術、ヒトB細胞ハイブリドーマ技術、及びEBV-ハイブリドーマ技術が含まれるが、これらに限定されるものではない（例えば、Kohler, G. ら. (1975) *Nature* 256:495-497; Kozbor, D. ら. (1985) *J. Immunol. Methods* 81-8-42; Cote, R. J. ら. (1983) *Proc. Natl. Acad. Sci.* 80:2026-2030; Cole, S. P. ら. (1984) *Mol. Cell Biol.* 62:109-120を参照）。

20

【0240】

更に、「キメラ抗体」作製のために発達したヒト抗体遺伝子にマウス抗体遺伝子をスプライシングするなどの技術が、好適な抗原特異性及び生物学的活性を備える分子を得るために用いられる（例えば、Morrison, S. L. 他. (1984) *Proc. Natl. Acad. Sci.* 81-4851-4855; Neuberger, M. S. 他. (1984) *Nature* 312:604-608; Takeda, S. ら. (1985) *Nature* 314:452, 454を参照）。別法では、当分野で周知の方法を用いて、一本鎖抗体の産生のための記載された技術を適用して、DME特異性一本鎖抗体を生成する。関連する特異性を備えるが別のイディオタイプの組成の抗体は、ランダムな組み合わせの免疫グロブリンライブラリから鎖混合によって生成することもできる（例えば、Burton D. R. (1991) *Proc. Natl. Acad. Sci.* 88:11120-3を参照）。

30

【0241】

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

40

【0242】

DMEに対する特異的な結合部位を含む抗体も得ることができる。例えば、このような断片には、抗体分子のペプシン消化によって生成される $F(ab')_2$ 断片と、 $F(ab')$ 断片のジスルフィド架橋を減じることによって生成される Fab 断片が含まれるが、これらに限定されるものではない。別法では、Fab発現ライブラリを作製することによって、所望の特異性とモノクローナル Fab 断片の迅速且つ容易な同定が可能となる（例えば、Huse, W. D. ら. (1989) *Science* 254:1275-1281を参照）。

50

【0243】

種々のイムノアッセイを用いてスクリーニングし、所望の特異性を有する抗体を同定する。隔離された特異性を有するポリクローナル抗体またはモノクローナル抗体の何れかを用いる競合的な結合アッセイ、または免疫放射線アッセイ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調整の計測が含まれる。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、競合的な結合アッセイも利用することができる(Pound、前出)。

【0244】

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 $10^9 \sim 10^{12} \text{ L/mol}$ 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 $10^6 \sim 10^7 \text{ L/mol}$ 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. and Cryer, A. (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY)。

【0245】

ある下流での適用におけるこのような医薬品の品質及び適性を調べるために、ポリクローナル抗体医薬の抗体価及び結合活性を更に評価する。例えば、少なくとも $1 \sim 2 \text{ mg/ml}$ の特異的な抗体、好ましくは $5 \sim 10 \text{ mg/ml}$ 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。様々な適用例における抗体の特異性及び抗体価、結合活性、抗体の品質や使用法の指針は一般に入手可能である。(例えば、Catty、前出、及びColigan他、前出を参照)。

【0246】

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、治療目的で利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。

【0247】

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

10

20

30

40

50

t, W. and W. Walther (1994) Pharmacol. Ther. 63 (3): 323-347を参照)。その他の遺伝送達機構には、リポソーム系、人工的なウイルスエンベロップ、及び当分野で周知のその他の系が含まれる (Rossi, J. J. (1995) Br. Med. Bull. 51 (1): 217-225; Boado, R. J. 他 (1998) J. Pharm. Sci. 87 (11): 1308-1315; and Morris, M. C. 他 (1997) Nucleic Acids Res. 25 (14): 2730-2736. を参照)。

【0248】

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

【0249】

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な導入技術には、(i) 個々の細胞内へのDNAのマイクロインジェクション、(ii) 金粒子の打ち込み、(iii) リポソーム仲介性トランスフェクション、(iv) 受容体仲介性遺伝子導入、及び (v) DNAトランスポソン (Morgan, R. A. and W. F. Anderson (1993) Annu. Rev. Biochem. 62: 191-217; Ivics, Z. (1997) Cell 91: 501-510; Boulay, J-L. and H. Recipon (1998) Curr. Opin. Biotechnol. 9: 445-450) の使用が含まれる。

【0250】

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、PCDNA 3.1、EPITAG、PRCCMV2、PREP、PVAXベクター (Invitrogen, Carlsbad CA)、PCMV-SCRIPT、PCMV-TAG、PEGSH/PERV (Stratagene, La Jolla CA)、PTET- 50

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 β -アクチン遺伝子等)、(ii) 誘導性プロモーター (例えば、市販されているT-REXプラスミド (Invitrogen) に含まれている、テトラサイクリン調節性プロモーター (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. U. S. A. 89:5547-5551; Gossen, M. 他 (1995) Science 268:1766-1769; Rossi, F. M. V. and H. M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456))、エクジソン誘導性プロモーター (市販されているプラスミドPVGRXR及びPINDに含まれている: Invitrogen)、FK506/ラパマイシン誘導性プロモーター、またはRU486/ミフェプリストーン誘導性プロモーター (Rossi, F. M. V. and H. M. Blau, 前出)、または(iii) 正常な個体に由来する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。

【0251】

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

【0252】

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社から入手可能であり、公表データ (Riviere, I. 他. (1995) Proc. Natl. Acad. Sci. U. S. A. 92:6733-6737) に基づいている。上記データを引用することをもって本明細書の一部とする。このベクターは、VSV (Armentano, D. 他 (1987) J. Virol. 61:1647-1650; Bender, M. A. 他 (1987) J. Virol. 61:1639-1646; Adam, M. A. and A. D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. 他 (1998) J. Virol. 72:8463-8471; Zufferey, R. 他 (1998) J. Virol. 72:9873-9880) 等の乱交雑エンベロープタンパク質若しくは標的細胞上の受容体に対する親和性を有するエンベロープ遺伝子を発現する好適なベクター産生細胞系 (VPCCL) において増殖される。RIGGに付与された米国特許第5,910,434号 (「Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant」) において、レトロウイルスパッケージング細胞系を得るための方法が開示されており、引用することをもって本明細書の一部とする。レトロウイルスベクターの増殖、ある細胞集団 (例えば、CD4⁺T細胞) の形質導入、並びに形質導入した細胞を患者に戻す

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

【0253】

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は当分野では周知である。複製欠損型アデノウイルスベクターは、免疫調節タンパク質をコードする遺伝子を脾臓の無損傷の脾島の中に導入するために可変性であることが証明された (Csete, M. E. 他. (1995) Transplantation 27:263-268)。使用できる可能性のあるアデノウイルスベクターが、米国特許第5,707,618号 (「Adenovirus vectors for gene therapy」) に記載されており、引用することをもって本明細書の一部とする。アデノウイルスベクターについてはまた、Antinozzi, P. A. 他 (1999) Annu. Rev. Nutr. 19:511-544; and Verma, I. M. and N. Somia (1997) Nature 18:389:239-242を参照し、引用することをもって本明細書の一部とする。

【0254】

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分野では周知である。複製適格性の単純疱疹ウイルス (HSV) I型系のベクターは、霊長類の眼にレポーター遺伝子を送達するために用いられてきた (Liu, X. 他 (1999) Exp. Eye Res. 169:385-395)。HSV-1ウイルスベクターの作製は、DeLucaに付与された米国特許第5,804,413号 (Herpes simplex virus swains for gene transfer) に記載されており、引用することをもって本明細書の一部とする。米国特許第5,804,413号には、ヒト遺伝子治療を含む目的のために、好適なプロモーターのコントロールの下で、細胞に導入される少なくとも1つの内在性遺伝子を含むゲノムからなる組換えHSV d9 2についての記載がある。また上記特許には、ICP4、ICP27及びICP22のために除去される組換えHSV株の作製及び使用方法が開示されている。HSVベクターについては、Goins, W. F. 他 (1999) J. Virol. 73:519-532 and Xu, H. 他 (1994) Dev. Biol. 163:152-161を参照し、引用することをもって本明細書の一部とする。クローニングされたヘルペスウイルス配列の操作や、巨大ヘルペスウイルスのゲノムの異なった部分を含む多数のプラスミドをトランスフェクトした後の組換えウイルスの継代、ヘルペスウイルスの成長及び増殖、並びにヘルペスウイルスの細胞への感染は当分野で周知の技術である。

【0255】

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

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

10

【0256】

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

20

【0257】

酵素性RNA分子であるリボザイムは、RNAの特異的切断を触媒するために用いることができる。リボザイム作用の機構には、相補的な標的RNAへのリボザイム分子の配列特異性ハイブリダイゼーションが含まれ、ヌクレオチド鎖切断が続く。例えば、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。

30

【0258】

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

【0259】

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

40

【0260】

RNA分子を修飾することによって、細胞内の安定性を高め、半減期を長くすることがで

50

きる。可能な修飾には、分子の5'及び／または3'端部でのフランキング配列の追加、または分子のバックボーン内のホスホジエステル結合の代わりにホスホロチオネートまたは2' Oメチルを用いる修飾が含まれるが、これらに限定されるものではない。PNAの生成に固有のこの概念は、内在性のエンドヌクレアーゼによって容易には認識されないアデニン、シチジン、グアニン、チミン、及びウリジンのアセチル、メチル、チオール、及び同様の修飾形態だけでなく、イノシン、キューエオシン (queosine)、及びワイブトシン (wybutosine) などの従来のものでない塩基を含めることによって、これらの分子の全体に拡大することができる。

【0261】

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

【0262】

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

【0263】

ベクターを細胞または組織に導入する多数の方法が利用でき、in vivo、in vi

10

20

30

40

50

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

【0264】

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

【0265】

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

【0266】

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

【0267】

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

【0268】

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

【0269】

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輸送を促進し得る。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。このようにして作製された融合タンパク質は、マウスモデル系の脳を含む全ての組織の細胞に形質導入されることが確認されている（Schwarze, S. R. 他 (1999) Science 285: 1569-1572）。

【0270】

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

【0271】

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、

10

20

30

40

50

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

10

【0272】

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

【0273】

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

20

【0274】

（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。診断に有用な抗体は、治療のところで記載した方法と同じ方法で製剤される。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。これらの抗体は、修飾をして或いはしないで使用され、レポーター分子の共有結合性或いは非共有結合性の接着によって標識化され得る。当分野で周知の種々のレポーター分子が用いられるが、その内の幾つかは上記した。

30

【0275】

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。標準的な複合体形成の量は、測光法（photometric）などの種々の方法で定量され得る。被験者のDMEの発現の量、制御及び疾患、生検組織からのサンプルが基準値と比較される。基準値と被験者との間の偏差が、診断の指標となる。

40

【0276】

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。用いられるポリヌクレオチドには、オリゴヌクレオチド配列、相補的なRNA及びDNA分子、及びPNAが含まれる。このポリヌクレオチドを用いて、疾患と相関し得る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

50

節を監視する。

【0277】

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、或いはアレルや関連配列コードする自然界の配列のみを同定するかどうかによって決まるであろう。

【0278】

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: 11-20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。

【0279】

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。このようなベクターは市販されており、当業者には周知であり、好適なRNAポリメラーゼ及び好適な標識されたヌクレオチドを加えることによって、in vitroでRNAプローブを合成するために用いられる。ハイブリダイゼーションプローブは、例えば³²P或いは³⁵Sなどの放射性核種、或いはアビジン／ビオチン(biotin)結合系によってプローブに結合されたアルカリホスファターゼなどの酵素標識等の種々のレポーターの集団によって標識され得る。

【0280】

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、細胞増殖異常の異常、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、脾炎、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、脾臓、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)、脊髓形成異常症

10

20

30

40

50

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-トゥ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-クリスチャン病、レトラ-シヴェ病、サルコイドーシス、エンプティセラ症候群、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-Hearley-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 α -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 (chiasmata tumor) 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)、脂肪異常栄養症、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、低カルシウム血症、低リン酸血症、思春期後小脳性運動失調症、チロシン血症が含まれ、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、 α -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。DME をコードするポリヌクレオチド配列は、サザーン法やノーザン法、ドットプロット法、或いはその他の膜系の技術、PCR 法、ディップスティック (dips

t i c k)、ピン(p i n)、E L I S A式アッセイ、及び変異D M Eの発現を検出するために患者から採取した体液或いは組織を利用するマイクロアレイに使用することが可能である。このような質的或いは量的方法は、当分野では周知である。

【0281】

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

10

【0282】

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

20

【0283】

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

【0284】

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

30

【0285】

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

40

【0286】

或る実施態様において、D M Eをコードするポリヌクレオチド配列由来のオリゴヌクレオチドプライマーを用いて、一塩基多型(S N P)を検出し得る。S N Pは、ヒトの先天性または後天性遺伝病の原因となる場合が多いヌクレオチドの置換、挿入及び欠失である。限定するものではないが、S N Pの検出方法には、一本鎖立体構造多型(S S C P)及び蛍光S S C P(f S S C P)法が含まれる。S S C Pでは、D M E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(P C R)でD N Aを増幅する。このD N Aは、例えば病変或いは正常な組織、生検サンプル、体液等に由来し得る。このD N A内のS N Pは、一本鎖形状のP C R産物の2次及び3次構造に差異を生じさせる。この差異は非変性ゲル中でのゲル電気泳動法を用いて検出可能

50

である。f S C C Pでは、オリゴヌクレオチドプライマーを蛍光標識することによって、DNAシーケンシング装置などのハイスループット機器でアンプリマー (amplimer) の検出をすることが可能になる。更に、インシリコSNP (in silico SNP: is SNP) と呼ばれる配列データベース分析法は、共通のコンセンサス配列の構築に用いられる個々の重複するDNA断片の配列を比較することによって、多型を同定することができる。これらのコンピュータベースの方法は、DNA配列クロマトグラムの自動分析及び統計モデルを用いたシーケンシングエラーや研究室でのDNAの調整に起因する配列のばらつきを排除する。別法では、例えばハイスループットのMASSARRAYシステム (Sequenom, Inc., San Diego CA) を用いた質量分析によりSNPを検出し、特徴付ける。

10

【0287】

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

【0288】

更に別の実施例では、本明細書で記載した任意のポリヌクレオチド配列に由来するオリゴヌクレオチドまたはより長い断片を、マイクロアレイにおける標的として用いることができる。マイクロアレイを、上記したように多数の遺伝子の相対的な発現レベルを同時にモニタリングする転写イメージング技術に用いることができる。マイクロアレイはまた、遺伝子変異、突然変異及び多型の同定に用いることができる。この情報を用いて、遺伝子機能を決定し、疾患の遺伝的根拠を解明し、疾患を診断し、遺伝子発現に関連する疾病の進行/後退をモニタリングし、疾患の治療における治療薬の開発や活性のモニタリングを行うことができる。特に、患者にとって最適かつ有効な治療法を選択するために、この情報を用いて患者の薬理ゲノムプロファイルを作成することができる。例えば、患者の薬理ゲノムプロファイルに基づいて、患者に対して極めて効果的でありながら副作用を殆ど示さない治療薬を選択することができる。

20

30

【0289】

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、上記のようにタンパク質間相互作用、薬剤-標的相互作用及び遺伝子発現プロファイルをモニタリング及び測定することが可能である。

【0290】

特定の実施例は、或る組織または細胞型の転写イメージを生成する本発明のポリヌクレオチドの使用に関連する。転写イメージは、特定の組織または細胞型により遺伝子発現の包括的パターンを表す。包括的遺伝子発現パターンは、所定の条件下で所定の時間に発現した遺伝子の数及び相対存在量を定量することにより分析される (Seilliamer 他、米国特許第5, 840, 484号の "Comparative Gene Transcript Analysis" を参照。この特許に言及することを以って本明細書の一部とする)。従って、特定の組織または細胞型の転写物または逆転写物の全てに本発明のポリヌクレオチドまたはその相補配列をハイブリダイズすることにより、転写イメージが生成され得る。或る実施例では、本発明のポリヌクレオチドまたはその相補配列がマイクロアレイ上に複数のエレメントのサブセットを構成するハイスループット型でハイブリダイゼーションさせる。結果として得られる転写イメージは、遺伝子活性のプロファイルとなり得る。

40

【0291】

50

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

【0292】

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

【0293】

一実施例では、試験化合物の毒性は、核酸を含有する生体サンプルをその試験化合物で処理して評価する。処理した生体サンプル中で発現した核酸は、本発明のポリヌクレオチドに特異的な1若しくは複数のプローブでハイブリダイズさせ、それによって本発明のポリヌクレオチドに対応する転写レベルを定量することができる。処理した生体サンプル中の転写レベルを、非処理生体サンプル中のレベルと比較する。両サンプルの転写レベルの差が、処理されたサンプル中で試験化合物が引き起こす毒性反応を示唆する。

【0294】

別の実施例は、本発明のポリペプチド配列を用いて組織または細胞型のプロテオームを分析することに関連する。「プロテオーム」という用語は、或る特定の組織または細胞型におけるタンパク質発現の包括的パターンを指す。プロテオームを構成する各タンパク質は、個々に更なる分析をすることができる。プロテオーム発現パターン即ちプロファイルは、所定の条件下で所定の時間に発現したタンパク質の数及びそれらの相対的な存在量を定量することにより分析する。従って、ある細胞のプロテオームのプロファイルは、特定の組織または細胞型のポリペプチドを分離及び分析することにより作成し得る。或る実施例では、このような分離は2次元ゲル電気泳動によって行う。この2次元ゲル電気泳動法では、まず、1次元の等電点電気泳動によりサンプルからタンパク質を分離し、次に、2次元のドデシル硫酸ナトリウムスラブゲル電気泳動により分子量に従って分離する (前出のSteiner and Anderson)。これらのタンパク質は、通常クーマシーブルーまたはシルバーまたは蛍光染色などの染色剤を用いてゲルを染色して、分散した個別の位置にあるスポットとしてゲル中で可視化される。各タンパク質スポットの光学密度は、通常サンプル中のタンパク質レベルに比例する。異なるサンプル、例えば試験化合物ま

たは治療薬で処理済みまたは未処理のいずれかの生体サンプルから得られる等位置にあるタンパク質スポットの光学密度を比較し、処理に関連するタンパク質スポット密度の変化を調べる。スポット内のタンパク質は、例えば化学的または酵素的に切断した後、質量分析する標準的な方法を用いて部分的にシークエンシングする。スポット内のタンパク質の同一性は、好適には少なくとも5個の連続するアミノ酸残基であるその部分的な配列を、本発明のポリペプチド配列と比較することにより決定し得る。場合によっては、決定的なタンパク質同定のための更なる配列が得られる。

【0295】

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、マイクロアレイ上のエレメントとして抗体を用い、マイクロアレイをサンプルに曝露して各アレイエレメントへのタンパク質結合レベルを検出することによりタンパク質発現レベルを定量する (Lueking, A. ら. (1999) Anal. Biochem. 270:103-111、Mendoe, L. G. ら. (1999) Biotechniques 27:778-788)。検出は当分野で既知の様々な方法で行うことができ、例えば、チオール反応性またはアミノ反応性蛍光化合物を用いてサンプル中のタンパク質を反応させ、各アレイエレメントにおける蛍光結合の量を検出し得る。 10

【0296】

プロテオームレベルでの毒性シグネチャも中毒学的スクリーニングに有用であり、転写レベルでの毒性シグネチャと並行して分析するべきである。或る組織における或るタンパク質では、転写物の存在量とタンパク質の存在量との相関性が低いことがあるため (Anderson, N. L. and J. Seilhamer (1997) Electrophoresis 18:533-537)、プロテオーム毒性シグネチャは、転写イメージにはそれ程影響しないがプロテオームのプロファイルを変化させる化合物の分析において有用たり得る。更に、体液中での転写の分析は、mRNAが急速に分解するため困難である。したがって、このような場合にはプロテオームのプロファイル作成はより信頼でき、情報価値がある。 20

【0297】

別の実施例では、試験化合物の毒性は、タンパク質を含む生体サンプルをその試験化合物で処理して評価する。処理された生体サンプル中で発現したタンパク質を分離して、各タンパク質の量が定量できるようにする。各タンパク質の量を、未処理生体サンプル中の対応するタンパク質の量と比較する。両サンプル中のタンパク質の量の差は、処理されたサンプル中の試験化合物に対する毒性反応を示唆する。個々のタンパク質は、それらのアミノ酸残基をシークエンシングし、これらの部分配列を本発明のポリペプチドと比較することで同定する。 30

【0298】

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

【0299】

当分野で周知の方法でマイクロアレイを準備して使用し、分析する。(例えば、Brennan, T. M. 他 (1995) 米国特許第5,474,796号; Schena, M. 他 (1996) Proc. Natl. Acad. Sci. 93:10614-10619; Baldeschweiler 他 (1995) PCT出願番号WO95/251116; Shalon, D. 他 (1995) PCT出願番号WO95/35505; Heller, R. A. 他 (1997) Proc. Natl. Acad. Sci. 94:2150-2155; 及び Heller, M. J. 他 (199 50

7) 米国特許第 5, 605, 662 号を参照)。様々なタイプのマイクロアレイが周知であり、詳細については、DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London に記載されている。また、この文献を引用することを以って本明細書の一部とする。

【0300】

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。コーディング配列または非コーディング配列の何れかを用いることができるが、或る例では、コーディング配列より非コード配列が好ましい。例えば、多重遺伝子ファミリーのメンバー間にコーディング配列が保存されていることにより、染色体マッピング時に望ましくない交差ハイブリダイゼーションが生じる可能性がある。この配列は、特定の染色体、染色体の特定領域または人工の染色体、例えば、ヒト人工染色体 (HAC)、酵母人工染色体 (YAC)、細菌人工染色体 (BAC)、細菌 P1 産物、或いは単一染色体 cDNA ライブラリに対してマッピングされる (Harrington, J. J. ら (1997) Nat Genet. 15:345-355、Price, C. M. (1993) Blood Rev. 7:127-134、Trask, B. J. (1991) Trends Genet. 7:149-154 等を参照)。一度マッピングすると、本発明の核酸配列を用いて、例えば病状の遺伝と特定の染色体領域やまたは制限断片長多型 (RFLP) の遺伝とが相関するような遺伝子連鎖地図を作成可能である (Land 10
er, E. S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357 を参照)。

【0301】

in situ 蛍光ハイブリダイゼーション (FISH) は、他の物理的及び遺伝子地図データと相関し得る (例えば、Heinz-Ulrich, 他による (1995) in Meyers, 前出, pp. 965-968 を参照)。遺伝子地図データの例は、種々の科学誌あるいは Online Mendelian Inheritance in Man (OMIM) のワールドワイドウェブのサイトで見付けることができる。物理的な染色体地図上の 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 30
を決定するクローニングが行われる。

【0302】

染色体標本の in situ ハイブリダイゼーション、及び確定した染色体マーカーを用いた結合分析などの物理的マッピング技術を用いて、遺伝子地図を拡張することもできる。マウスなどの別の哺乳動物の染色体上に遺伝子を配置させることにより、たとえ正確なヒト染色体の位置が分かていなくても、関連するマーカーが明らかになる場合が多い。この情報は、位置クローニング或いは別の遺伝子発見技術を用いて遺伝的疾患の研究をしている研究者にとって価値がある。疾患や症候群に關与する 1 つ或いは複数の遺伝子の位置が、例えば血管拡張性失調症の 11q22-23 などの特定の遺伝子領域に遺伝子結合によって大まかに決定されると、その領域に対するどの配列マッピングも、さらなる調査の 40
ための関連する遺伝子或いは調節遺伝子を表す (例えば、Gatti, R. A. 他による (1988) Nature 336:577-580 を参照)。また、目的の本発明のスクレオチド配列を用いて、正常者、保有者、即ち感染者の間の、転位置、反転などによる染色体位置の違いを検出することもある。

【0303】

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。このようなスクリーニングに用いる断片は、溶液に遊離、固体支持物に固定、細胞の表面上に保持、或いは細胞内に存在する。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。

【0304】

薬剤スクリーニングに用いる別の方法は、目的のタンパク質に対して、好適な結合親和性を有する化合物のスクリーニング処理能力を高めるために用いられる（例えば、Gey sen, 他による（1984）PCT出願番号WO84/03564を参照）。この方法では、相当な数の異なる小さな試験用化合物が、プラスチックピン或いは他の基板の上に合成される。試験用化合物は、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、ペプチドを捕らえ、固体支持物に固定することもできる。

10

【0305】

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。

【0306】

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

【0307】

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

20

【0308】

前述した及び以下に記載する全ての特許出願、特許、刊行物、特に米国特許出願第60/197,590号、同第60/198,403号、同第60/200,185号、同第60/202,234号、および同第60/203,509号に言及することをもって本明細書の一部とする。

【0309】

（実施例）

1 cDNAライブラリの作製

30

インサイトcDNAはLIFESEQ GOLD データベース（Incyte Genomics, Palo Alto CA）に含まれているcDNAライブラリに由来し、表4の列5に示されている。ある組織をグアニジニウムイソチオシアネート溶液においてホモジナイズして溶解する一方、別の組織を、フェノールにおいて、或いはグアニジニウムイソチオシアネート及びフェノールの単相溶液であるTRIZOL（Life Technologies）などの変性剤の好適な混合液においてホモジナイズして溶解した。この溶解物を塩化セシウムにおいて遠心分離によって、或いはクロロホルムで抽出した。RNAは、イソプロパノール或いは酢酸ナトリウムとエタノールのどちらか、或いは別の一般的な方法でこの溶解物から沈殿させた。

【0310】

40

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

【0311】

ある場合には、Stratagene社にRNAを提供し、Stratagene社が対応するcDNAライブラリを作製した。そうでない場合は、UNIZAPベクターシステ

50

ム (Stratagene) または SUPERSCRIPT プラスミドシステム (Life Technologies) を用いて当分野で周知の推奨方法または類似の方法で cDNA を合成して cDNA ライブラリを作製した。(例えば、Ausubel, 1997, 前出, ユニット 5. 1-6. 6 を参照)。逆転写は、オリゴ d (T) またはランダムプライマーを用いて開始させた。合成オリゴヌクレオチドアダプターを二本鎖 cDNA に結合させてから、好適な 1 或いは複数の制限酵素で cDNA を消化した。殆どのライブラリでは、SEPHACRYL S1000、SEPHAROSE CL2B、SEPHAROSE CL4B カラムクロマトグラフィー (Amersham Pharmacia Biotech)、またはアガロースゲル電気泳動法によって cDNA の大きさ (300 ~ 1000 bp) を選択した。PBLUESCRIPT プラスミド (Stratagene)、pSPORT1 プラスミド (Life Technologies)、pcDNA 2. 1 プラスミド (Invitrogen Carlsbad CA)、PBK-CMV プラスミド (Stratagene)、pINCY プラスミド (Incyte Pharmaceuticals, Palo Alto CA)、またはそれらの誘導体などの好適なプラスミドのポリリンカーの適合性制限酵素部位に cDNA を結合させた。この組換えプラスミドを、Stratagene 社の XL1-Blue, XL1-Blue MRF、SOLR、または Life Technologies 社の DH5α または DH10B、ELECTROMAX DH10B を含むコンピテント大腸菌細胞に導入し形質転換した。

10

20

30

40

50

【0312】

2 cDNA クローンの単離

上記実施例 1 に記載したように得たプラスミドは、UNIZAP ベクターシステム (Stratagene) 或いは細胞溶解を利用した in vivo 切除によって宿主細胞から回収した。プラスミドの精製には、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 Plus、QIAWELL 8 Plus Plasmid、または QIAWELL 8 Ultra Plasmid 精製システム、または REAL Prep 96 プラスミドキットの内の少なくとも 1 つを用いた。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しないで 4 °C で保管した。

【0313】

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

【0314】

3 シークエンシング及び分析

実施例 2 に記載したようにプラスミドから回収したインサイト cDNA を、以下に示すようにシークエンシングした。cDNA のシークエンシング反応は標準的な方法で行うか、または HYDRA マイクロディスペンサー (Robbins Scientific) 或いは MICROLAB 2200 (Hamilton) 液体移送装置と共に ABI CALYST 800 (PE Biosystems) サーマルサイクラー或いは PTC-200 thermal cycler (MJ Research) などのハイスルーブット装置を用いて行った。cDNA のシークエンシング反応は、Amersham Pharmacia Biotech 社の試薬、または ABI PRISM BIGDYE Terminator cycle sequencing ready reaction キット (PE Biosystems) などの ABI シークエンシングキットに含まれる試薬

を用いて行った。cDNAシーケンシングの反応物の電気泳動的による分離及び標識したポリヌクレオチドの検出は、MEGABACE 1000 DNAシーケンシングシステム (Molecular Dynamics)、標準ABIプロトコル及び塩基対呼び出しソフトウェアを用いるABI PRISM 373または377シーケンシングシステム (PE Biosystems)、または当分野で周知のその他の配列解析システムを用いて行った。cDNA配列内の読み枠は、標準的な方法 (Ausubel, 1997, 前出, unit 7.7) を用いて決定した。cDNA配列の幾つかを選択して、実施例 8に記載した方法で配列を伸長した。

【0315】

インサイトcDNAに由来する本ポリヌクレオチド配列の確認は、BLAST、動的プロ
10 グラミング、およびジヌクレオチドの分布による解析 (dinucleotide nearest neighbor analysis) に基づいたプログラム及びアルゴリズムを用いて、ベクター、リンカー、およびポリA配列を取り除き、更にあいまいな塩基対をマスクすることで行った。次に、インサイトcDNA配列およびそれらの翻訳を、公共のデータベースであるGenBankの霊長類、げっ歯類、哺乳類、脊椎動物、および真核生物のデータベース、およびBLOCKS、PRINTS、DOMO、PRODOM、およびPFAMなどの隠れマルコフモデル (HMM) を基にしたタンパク質ファミリーのデータベースから選択した配列に対して問合せた (HMMは、遺伝子ファミリーのコンセンサス主構造を分析する確率的手法である。例えば、Eddy, S. R. (1996)
20) Curr. Opin. Struct. Biol. 6: 361-365を参照)。このような問合せは、BLAST、FASTA、BLIMPS、およびHMMRに基づいたプログラムを用いて行った。インサイトcDNA配列を組み立てて、完全長ポリヌクレオチド配列を作製した。或いは、GenBank cDNAs、GenBank EST、ステッチ配列 (stitched sequence)、ストレッチ配列 (stretched sequences)、またはGenscan-推定コード配列 (実施例 4および5を参照) を用いて、インサイトcDNA群を完全長の配列に伸長した。配列の組み立ては、Phred、Phrap、およびConsedに基づいたプログラムを用いて行い、GeneMark、BLAST、およびFASTAに基づいたプログラムを用いて、オープンリーディングフレームを決定するべくcDNA群をスクリーニングした。完全長のポリヌクレオチド配列を翻訳して対応する完全長ポリペプチド配列を得た。別法では、
30 本発明のポリヌクレオチドは、完全長翻訳ポリペプチドの任意のメチオニン残基から始まり得る。次に、完全長ポリペプチド配列をGenBankタンパク質データベース (genpept)、SwissProt、BLOCKS、PRINTS、DOMO、PRODOM、Prosite、およびPFAMなどの隠れマルコフモデル (HMM) に基づいたタンパク質ファミリーデータベースに対して問合せて分析した。これらの完全長ポリヌクレオチド配列はまた、MACDNASIS PRO ソフトウェア (Hitachi Software Engineering, South San Francisco CA) およびLASERGENEソフトウェア (DNASTAR) を用いて分析した。ポリヌクレオチド配列およびポリペプチド配列のアラインメントを、アラインメントした配列間のパーセント同一性も計算するMEGALIGNマルチシーケンスアラインメントプログラム (DNASTAR) に組み込まれたCLUSTALアルゴリズムによって指定された
40 デフォルトパラメータを用いて作成した。

【0316】

表7は、インサイトcDNAおよび完全長配列の組み立て、および組み立てた配列の分析に利用したツール、プログラム、およびアルゴリズム、並びにそれらの説明、引用文献、閾値パラメータを簡単に示す。表7の列1は用いたツール、プログラム、およびアルゴリズム、列2はそれらの簡単な説明、列3は引用することで本明細書の一部とした引用文献、列4の記載されている部分は2つの配列の一致の程度を評価するために用いたスコア、確率値、およびその他のパラメータを示す (スコアが高くなれば高くなるほど即ち確率値が低ければ低いほど、配列間の相同性が高くなる)。

【0317】

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: 11-20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。

【0318】

4 ゲノムDNAからコード配列の同定および編集

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、様々な生物に由来するゲノムDNA配列を分析するための汎用遺伝子同定プログラムである（Burge, C. および S. Karlin (1997) J. Mol. Biol. 268 : 78-94、Burge, C. および S. Karlin (1998) Curr. Opin. Struct. Biol. 8 : 346-354を参照）。このプログラムは推定エキソンを連結して、メチオニンから停止コドンまで伸長した組み立てcDNA配列を構築する。Genscanにより得られる配列は、FASTAデータベースのポリヌクレオチド配列およびポリペプチド配列になる。Genscanによって一回で解析できる配列の最大長さは30kbに設定されている。これらのGenscan推定cDNA配列の内、どの配列が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。次に、これらの選択されたGenscan推定配列を、BLAST解析を用いてgeneptおよびgbpri公共データベースの配列と比較した。必要に応じて、Genscan推定cDNA配列を、geneptにおいてBLASTで最もヒットした配列と比較して、Genscan推定配列における余分なエキソンや省いてしまったエキソンなどのエラーを修正し、編集した。BLAST解析を用いてGenscan推定cDNA配列を含むインサイトcDNAまたは公共のcDNAを見つけ出すことにより、転写の証拠が得られる。インサイトcDNAがGenscan推定cDNA配列を含む場合、この情報を用いてGenscan推定配列を修正或いは確認できる。完全長ポリヌクレオチド配列は、実施例3に説明した組み立て方法でGenscan推定コード配列とインサイトcDNAおよび／または公共のcDNA配列を組み立てて作製した。別法では、完全長ポリヌクレオチド配列は、その全てが編集した或いは未編集のGenscan推定コード配列から作製した。

【0319】

5 ゲノム配列データとcDNA配列データとの組み立て

ステッチ配列 (Stitched Sequence)

部分的なcDNA配列を、実施例4に記載したGenscan遺伝子同定プログラムによって推定されたエキソンで伸長した。実施例3に記載されたように組み立てられた部分的なcDNAをゲノムDNAにマッピングし、関連するcDNAおよび1或いは複数のゲノム配列に由来する関連する推定Genscanエキソンを含む複数のクラスターに入れた。各クラスターを、グラフ理論および動的計画法に基づいたアルゴリズムを用いて、cDNAおよびゲノム情報を統合して分析し、後に確認される潜在的なスプライスバリエーションを生成し、編集或いは伸長して完全長の配列を作製した。或るクラスターの2つ以上の配列に或る区間の全長が存在する配列区間を同定し、推移 (transitivity) により同定した区間を同等と考える。例えば、或る区間がcDNAおよび2つのゲノム配列のそれぞれに存在する場合、これら3つ全ての区間を同等と考える。この方法によって、関連しないが連続するゲノム配列をcDNA配列によって繋ぎ1つにする。このようにして同定された区間を、親配列 (parent sequence) に沿って現われるようにステッチアルゴリズムで縫い合わせ、可能な最も長い配列および変異配列を作製する。或るタイプ (cDNAとcDNA、またはゲノム配列とゲノム配列) の親配列に沿って連

結される区間と区間との繋ぎ合わせは、親配列のタイプが異なる（cDNAとゲノム配列）連結より好ましい。得られたステッチ配列を翻訳し、BLAST解析でgenpeptおよびgbpr i 公共データベースにおける配列と比較した。Genscanによって推定された不適当なエキソンを、geneptにおいてBLASTで最もヒットした配列と比較して修正する。このような配列を更なるcDNA配列で伸長し、必要に応じてゲノムDNAで検査した。

【0320】

ストレッチ配列 (Stretched Sequence)

部分的なDNA配列をBLAST解析に基づいたアルゴリズムで完全長に伸長した。まず、実施例3に記載したように組み立てた部分的なcDNAを、BLASTプログラムを用いてGenBankの霊長類、げっ歯類、哺乳類、脊椎動物、および真核生物のデータベースなどの公共のデータベースに対して問い合わせた。次に、GenBankの相同性の最も高いタンパク質を、実施例4に記載したインサイトcDNA或いはGenScanエキソン推定配列の何れかと比較した。得られた複数の高スコアのセグメント対(HSP)を用いてキメラタンパク質を作製し、GenBankの相同タンパク質上に翻訳した配列をマッピングした。元のGenBankの相同タンパク質に対して、キメラタンパク質に挿入や欠失が起こり得る。公共のヒトゲノムデータベースから相同ゲノム配列を探し出すために、GenBankの相同タンパク質およびキメラタンパク質の両方をプローブとして用いた。このようにして、部分的なDNA配列を相同ゲノム配列の付加によりストレッチすなわち伸長した。完全な遺伝子を含んでいるか得られたストレッチ配列を検査した。

【0321】

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: 11-20を組み立てるために用いた配列を、BLAST及びSmith-Watermanアルゴリズムを用いて、インサイトLIFESEQデータベース及び公共のドメインデータベースの配列と比較した。SEQ ID NO: 11-20と一致するこれらのデータベースの配列を、Phrap (表7) などの構築アルゴリズムを使用して、連続及び重複した配列のクラスターに組み入れた。Stanford Human Genomese Center (SHGC)、Whitehead Institute for Genome Research (WIGR) 及びGenethonなどの公共の情報源から入手できる放射線ハイブリッド (radiation hybrid) 及び遺伝子マッピングのデータを用いて、クラスター化した配列がすでにマッピングされているかを調べる。クラスターにマッピングされた配列が含まれている場合は、そのクラスターの全ての配列 (特定のSEQ ID NOを含む) をそのマッピング位置に割り当てた。

【0322】

遺伝子地図の位置は、範囲、区間、またはヒト染色体によって表される。センチモルガンで示したマッピング位置の範囲は、染色体の短腕 (p) の末端から測定した (センチモルガン (cM) は、同一染色体上の遺伝子間の乗換え率に基づいた距離を表す単位である。平均すると、1 cMはヒトの染色体の1メガベースに概ね等しいが、組換え率の高い部分と低い部分があるため、大きく変化し得る)。距離cMは、配列がそれぞれのクラスターに含まれている放射線ハイブリッドマーカーの境界を検出できるGenethonによってマッピングされた遺伝子マーカーに基づいている。NCBI「GeneMap99」 (<http://www.ncbi.nlm.nih.gov/genemap>) などの公衆が入手可能なヒト遺伝子マップおよびその他の情報源を用いて、上記した区間が既に同定されている疾患遺伝子マップ内若しくは近傍に位置するかを決定できる。

【0323】

7 ポリヌクレオチド発現の分析

ノーザン分析は、遺伝子の転写物の存在を検出するために用いられる実験用技術であり、特定の細胞種或いは組織からのRNAが結合されている膜への標識されたヌクレオチド配列のハイブリダイゼーションを伴う (例えば、Sambrook, 前出, 7章; 及び

Ausubel, F. M. 他、前出, 4章及び16章を参照)。

【0324】

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

【0325】

【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

として定義される積スコアである。積スコアは、0～100の標準化された値であり、以下のように求める。BLASTスコアにヌクレオチド配列の一致率を乗じ、その積を2つの配列の短い方の長さの5倍で除す。高スコアのセグメントの対(HSP)において一致する各塩基に+5のスコアを割り当て、各不適性塩基対に-4を割り当てることにより、BLASTスコアを計算する。2つの配列は、2以上のHSPを共有し得る(ギャップにより離隔される)。2以上のHSPがある場合には、最高BLASTスコアの塩基対を用いて積スコアを計算する。積スコアは、BLASTアラインメントの断片的重複と質とのバランスを表す。例えば積スコア100は、比較した2つの配列の短い方の長さ全体にわたって100%一致する場合にのみ得られる。積スコア70は、100%の同一性で重畳が70%であるか、或いは88%の同一性で重畳が100%であるかの何れかの場合である。積スコア50は、100%の同一性で重畳が50%であるか、或いは79%の同一性で重畳が100%であるかの何れかの場合である。

【0326】

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、ある完全長の配列は、少なくとも部分的にインサイトcDNA配列をオーバーラップさせて組み立てられる(実施例3を参照)。各cDNA配列は、ヒト組織から作製されたcDNAライブラリに由来する。各ヒト組織は、心血管系、結合組織、消化系、胚構造、内分泌系、外分泌腺、女性生殖器、男性生殖器、生殖細胞、血液および免疫系、肺、筋骨格系、神経系、脾臓、呼吸器系、感覚器官、皮膚、顎口腔系、分類不能/混合、または尿管などの1つの生物/組織のカテゴリーに分類される。各カテゴリーにおけるライブラリの数をカウントし、その合計数を全カテゴリーのライブラリ数で除す。同様に、各ヒト組織は、癌、細胞系、発生、炎症、神経、外傷、心血管、プール(pool)などの1つの疾患/症状のカテゴリーに分類され、各カテゴリーにおけるライブラリの数をカウントし、その合計数を全カテゴリーのライブラリ数で除す。得られるパーセンテージは、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。

【0327】

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は、完全長分子の好適な断片から設計したオリゴヌクレオチドプライマーを用いてその完全長分子の好適な断片を伸長して作製した。一方のプライマーは既知の断片の5'の伸長を開始するために合成し、他方のプライマーは既知の断片の3'の伸長を開始するために合成した。開始プライマーは、OLIGO 4.06ソフトウェア(National Biosciences) 或いは他の適切なプログラムを用いて、約22個から約30個のヌクレオチドの長さで約50%以上のGC含量を有し、かつ約68～72℃の温度で標的配列にアニールするように設計した。ヘアピン構造及び

10

20

30

40

50

プライマー-プライマー二量体が生じないようにヌクレオチドを伸長した。

【0328】

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

【0329】

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

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

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

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

【0330】

各ウェルのDNA濃度は、1X TE及び0.5 μ lの希釈していないPCR産物に溶解した100 μ lのPICOGREEN定量試薬 (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) を不透明な蛍光光度計プレート (Coming Costar, Acton MA) の各ウェルに分配してDNAが試薬と結合できるようにして測定する。このプレートをFluoroskan II (Labsystems Oy, Helsinki, Finland) でスキャンして、サンプルの蛍光を計測してDNAの濃度を定量化する。反応混合物の5~10 μ lのアリコットを1%のアガロースミニゲル上での電気泳動によって解析し、何れの反応物が配列を伸長することに成功したかを決定する。

【0331】

伸長したヌクレオチドを脱塩及び濃縮してから384ウェルプレートに移し、CviJI コレラウイルスエンドヌクレアーゼ (Molecular Biology Research, Madison WI) で消化し、pUC 18ベクター (Amersham Pharmacia Biotech) に再連結する前に音波処理またはせん断を行った。ショットガンシーケンシングのために、消化したヌクレオチドを低濃度 (0.6~0.8%) のアガロースゲル上に分離して断片を切断し、寒天をAgar ACE (Promega) で消化した。T4リガーゼ (New England Biolabs, Beverly MA) を用いて伸長したクローンをpUC 18ベクター (Amersham Pharmacia Biotech) に再連結し、Pfu DNAポリメラーゼ (Stratagene) で制限部位の延び出しを処理してコンピテント大腸菌細胞に形質移入した。形質移入した細胞を選択して抗生物質を含む培地に移し、それぞれのコロニーを切りと

ってLB／2Xカルベニシリン培養液の384ウェルプレートに37℃で一晩培養した。

【0332】

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

- ステップ1 94℃で3分間
- ステップ2 94℃で15秒
- ステップ3 60℃で1分間
- ステップ4 72℃で2分間
- ステップ5 ステップ2、3、及び4を29回繰り返す
- ステップ6 72℃で5分間
- ステップ7 4℃で保管。

10

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

20

【0333】

同様に上述の手順で、完全長のポリヌクレオチド配列を検査したり、或いは完全長のポリヌクレオチド配列を利用して、この伸長のために設計したオリゴヌクレオチドと好適なゲノムライブラリを用いて5'調節配列を得た。

【0334】

9 個々のハイブリダイゼーションプローブの標識化及び使用法

SEQ ID NO: 11-20から導き出されたハイブリダイゼーションプローブを用いて、cDNA、mRNA、またはゲノムDNAをスクリーニングする。約20塩基対からなるオリゴヌクレオチドの標識について特に記すが、より大きなcDNAフラグメントの場合でも基本的に同じ手順を用いる。オリゴヌクレオチドを、OLIGO4.06ソフトウェア (National Bioscience) のような最新式のソフトウェアを用いてデザインし、50 pmolの各オリゴマーと、250 µCiの[γ-³²P]アデノシン三リン酸 (Amersham, Chicago, IL) 及びT4ポリヌクレオチドキナーゼ (DuPont NEN, Boston MA) とを組み合わせて用いることにより標識する。標識されたオリゴヌクレオチドを、SEPHADEX G-25超精細排除デキストランビードカラム (Amersham Pharmacia Biotech) を用いて実質的に精製する。毎分10⁷ カウントの標識されたプローブを含むアリコットを、次のエンドヌクレアーゼ、Ase I、Bgl II、Eco RI、Pst I、Xba I 或いはPvu II (DuPont NEN) の1つを用いて切断したヒトゲノムDNAの典型的な膜ベースのハイブリダイゼーション解析において用いる。

30

【0335】

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

40

【0336】

10 マイクロアレイ

マイクロアレイ上のアレイエレメントの連結または合成は、フォトリソグラフィ、ピエゾ

50

プリント（インクジェットプリンター、前出のBaldeschweiler等を参照）、機械的マイクロスポッティング技術及びこれらから派生したものをを用いて達成することが可能である。上記各技術において基板は、均一な非多孔性の固体とするべきである（Schena（1999）：前出）。推奨する基板には、シリコン、シリカ、スライドガラス、ガラスチップ及びシリコンウエハがある。別法では、ドットプロット法またはスロットプロット法に類似のアレイを利用して、熱や紫外線、または化学的或いは機械的な結合手段で基板の表面にエレメントを配置して結合させることができる。通常のアレイは利用可能な方法や機械を用いて作製でき、任意の適正な数のエレメントを含めることができる（Schena, M. 他（1995）Science 270:467-470、Shalon, D. 他（1996）Genome Res. 6:639-645、Marshall, A. and J. Hodgson（1998）Nat. Biotechnol. 16:27-31. を参照）。 10

【0337】

完全長cDNA、発現遺伝子配列断片（EST）、或いはそれらの断片やオリゴマーが、マイクロアレイのエレメントとなり得る。ハイブリダイゼーションに好適な断片やオリゴマーを、LASERGENEソフトウェア（DNASTAR）などの当分野で周知のソフトウェアを用いて選択することが可能である。このアレイエレメントを、生体サンプル中のポリヌクレオチドとハイブリダイズさせる。生体サンプル中のポリヌクレオチドは、検出を容易にするために蛍光標識またはその他の分子タグに結合する。ハイブリダイゼーションの後、生体サンプルからハイブリダイズしなかったヌクレオチドを除去し、蛍光スキャナを用いて各アレイエレメントにおけるハイブリダイゼーションを検出する。別法では、レーザー脱離及び質量スペクトロメトリーを用いてもハイブリダイゼーションを検出し得る。マイクロアレイ上のエレメントにハイブリダイズする各ポリヌクレオチドの相補性の程度及び相対的存在量は、算定することができる。一実施例におけるマイクロアレイの調整及び使用について、以下に詳述する。 20

【0338】

組織または細胞サンプルの調製

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

【0339】

マイクロアレイの準備

本発明の配列を用いて、アレイエレメントを作製する。各アレイエレメントは、クローン化cDNA挿入断片を含むベクターを含有する細菌性細胞から増幅する。PCR増幅は、 50

cDNA挿入断片に隣接するベクター配列に相補的なプライマーを用いる。30サイクルのPCRによって、1~2 ngの初期量から5 µgを超える最終量までアレイエレメントを増幅する。増幅されたアレイエレメントは、SEPHACRYL-400 (Amersham Pharmacia Biotech) を用いて精製する。

【0340】

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

10

【0341】

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

20

【0342】

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

【0343】

ハイブリダイゼーション

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

30

【0344】

検出

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

40

【0345】

50

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

【0346】

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

【0347】

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、リニア20色変換を用いてシグナル強度が青色 (低シグナル) から赤色 (高シグナル) までの擬似カラー範囲にマッピングされるイメージとして表示される。データはまた、定量的に分析される。2つの異なる蛍光体を同時に励起して測定する場合には、各蛍光体の発光スペクトルを用いて、先ずデータは蛍光体間の光学的漏話 (重複発光スペクトルに起因する) に対して補正される。

【0348】

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

【0349】

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

【0350】

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をサブクロニングする。このようなプロモーターには、lacオペレーター調節エレメントに関連するT5またはT7バクテリオファージプロモーター及びtrp-lac (tac) ハイブリッドプロモーターが含まれるが、これらに限定されるものではない。組換えベクターを、BL21 (DE3

10

20

30

40

50

）などの好適な細菌宿主に形質転換する。抗生物質耐性をもつ細菌が、イソプロピル β -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、強いポリヘドリンプロモータによって高いレベルのcDNAの転写が行われる。組換えバキュロウイルスは、多くの場合は Spodoptera frugiperda (Sf9) 昆虫細胞に感染に用いられるが、ヒト肝細胞の感染にも用いられることもある。後者の感染の場合は、バキュロウイルスの更なる遺伝的変更が必要になる。(例えば、Engelhard, E. K. 他 (1994) Proc. Natl. Acad. Sci. USA 91: 3224-3227; Sandig, V. 他 (1996) Hum. Gene Ther. 7: 1937-1945. を参照)。

【0351】

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。 Schistosoma japonicum からの26キロダルトンの酵素GSTによって、タンパク質の活性及び抗原性を維持した状態で固定されたグルタチオンで融合タンパク質の精製が可能となる (Amersham Pharmacia Biotech)。精製の後、GST部分を特定の操作部位でDMEからタンパク質分解的に切断できる。アミノ酸8個のペプチドであるFLAGで、市販のモノクローナル及びポリクローナル抗FLAG抗体 (Eastman Kodak) を用いた免疫親和性の精製が可能となる。6個の連続するヒスチジン残基のストレッチである6-Hisによって、金属キレート樹脂 (QIAGEN) で精製が可能となる。タンパク質の発現及び精製の方法は、Ausubel (1995, 前出, ch 10, 16) に記載されている。これらの方法で精製した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。

【0352】

1.3 機能のアッセイ

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。このようなベクターには、pCMV SPORTTM (Life Technologies.) 及びpCR3.1 (Invitrogen, Carlsbad, CA) が含まれ、どちらもサイトメガロウイルスプロモーターを含んでいる。5~10 μ gの組換えベクターを、例えば内皮由来か造血由来のヒト細胞株にリポソーム製剤或いは電気穿孔法によって一時的に形質移入する。更に、標識タンパク質をコードする配列を含む1~2 μ gのプラスミドを同時に形質移入する。標識タンパク質の発現により、形質移入された細胞と形質移入されていない細胞とを区別できる。また、標識タンパク質の発現によって、cDNAの組換えベクターからの発現を正確に予想できる。このような標識タンパク質には、緑色蛍光タンパク質 (GFP; Clontech)、及びCD64またはCD64-GFP融合タンパク質が含まれる。レーザー光学に基づいた技術を利用した自動流動細胞計測法 (FCM) を用いて、GFPまたはCD64-GFPを発現する形質移入された細胞を同定し、その細胞のアポトーシス状態や他の細胞特性を評価する。また、FCMで、先行した或いは同時の細胞死の現象を診断する蛍光分子の取り込みを検出して計量する。これらの現象には、プロピジウムヨウ化物でのDNAの染色によって計測される核DNA内容物の変化と、プロモデオキシウリジンの取り込み量の低下によって計測されるDNA合成の下方調節と、特異的な抗体との反応性によって計測される細胞表面及び細胞内のタンパク質の発

現の変化と、蛍光複合アネキシンVタンパク質の細胞表面への結合によって計測される原形質膜組成の変化とが含まれる。流動細胞計測法は、Ormerod, M. G. による (1994) Flow Cytometry Oxford, New York, NY. に記載されている。

【0353】

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) の保存された領域と結合する。形質転換された細胞と形質転換されない細胞とは、ヒトIgG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

【0354】

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

【0355】

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

【0356】

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

【0357】

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。結合の後、そのレジンを製造者の使用説明書に従ってブロッキング処理し洗浄する。

40

【0358】

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、pH 2~3のバッファー、或いは高濃度の尿素またはチオシアン酸塩イオンのようなカオトロピ

50

ックイオンで) 溶出させ、DMEを回収する。

【0359】

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なその断片を、¹²⁵I ボルトンハンター試薬 (例えば、Boltton A. E. 及びW. M. Hunter (1973) Biochem. J. 133:529を参照) で標識する。マルチウェルプレートに予め配列しておいた候補の分子を、標識した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。

【0360】

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。

【0361】

DMEはまた、ハイスルーブット型の酵母2-ハイブリッドシステムを使用するPATHC ALLING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)。

【0362】

17 DME活性の実証

DMEのチトクロームP450活性を、アニリンの4-ヒドロキシ化を利用して測定する。アニリンは酵素によって4-アミノフェノールに変換され、630nmにおいて光吸収率が最大である (GibsonおよびSkett, 前出)。このアッセイは便利な方法であるが、ヒドロキシ化が2位および3位でも起こるため全体のヒドロキシ化が過少評価される。アッセイは37℃で行い、反応バッファには酵素のアリコートおよび好適な量のアニリン (最大2mM) が含まれている。バッファは、反応のためにNADPHまたはNADPH生成補助因子系を含まなければならない。この反応バッファのある調整では、85mM Tris pH 7.4、15mM MgCl₂、50mMニコチンアミド、40mg trisodium isocitrate、および2単位のイソクエン酸デヒドロゲナーゼを含め、アッセイの前に8mg NADP⁺を10mLの反応バッファストックに加える。反応は蛍光キュベットで行い、630nmで吸光度を測定する。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。既知の濃度の4-アミノフェノールを用いて標準的な曲線を作成することができる。

【0363】

DMEの1 α , 25-ジヒドロキシビタミンD 24-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 α , 25-ジヒドロキシビタミンD (1 α , 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1 μ gの1 α , 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 α , 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Dye

10

20

30

40

50

r, W. J. (1959) Can. J. Biochem. Physiol. 37: 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 α , 25-ジヒドロキシビタミンDまたは24, 25-ジヒドロキシビタミンD (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

【0364】

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生成補助因子系 (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。酸素の取り込み割合が酵素活性に比例する。

20

30

【0365】

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に、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。

40

【0366】

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GS

50

T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。

【0367】

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 (choilyl-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。

10

【0368】

DMEのN-アセチ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)の反応に基づいた分光光度アッセイが用いられ得る。N-アセチ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。

20

【0369】

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が、様々な時間間隔で、37℃で測定される。S-アデノシル-L-[メチル-³H]メチオニン ([³H] AdoMet; 比活性 p = 75 Ci/mmol; NEN Life Science Products) は、メチルドナー基質として用いられている。有用なメチル受容基質は、グルタチオンS-トランスフェラーゼ核小体蛋白質 (fibrillarin) glycine-アルギニンドメイン融合タンパク質 (GST-GAR)、非均質核リボ核タンパク質 (hnRNP)、またはアデノシンジアルデヒド処置細胞よりライセート中に存在する低メチル化タンパク質を含む。メチル化反応は、SDS-PAGEサンプルバッファを添加することで停止される。反応生成物は、SDS-PAGEで分解され、フルオログラフィーで示される。³H-ラベルされたメチルドナー基質は、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を示している (Tang, J. ら、(2000) J. Biol. Chem. 275:7723-7730 及び Tang, J. ら (2000) J. Biol. Chem. 275:19866-19876)。

30

40

【0370】

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 μM SAM (S-アデノシル-L-メチオニン) ヨウ化物 (0.5 μCiのメチル-[H³]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 μg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

50

n-ヘプタンで抽出する。次に10分間1000×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)。

【0371】

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 μM 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 μM 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)。

【0372】

DMEのアルド／ケト還元酵素活性を、NADPHが消費される時の340 nm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、135 mMのリン酸ナトリウムバッファ (酵素によりpH 6.2~7.2の範囲)、0.2 mM NADPH、0.3 M硫酸リチウム、0.5~2.5 μg 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1 μg 当たり消費されるNADPHのモル数として計算される。

【0373】

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340 nm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、50 mMのリン酸ナトリウム、pH 7.5、および0.25 mM 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1 μg 当たり生成されたNADHのモル数として計算する。

【0374】

DMEのカルボキシルエステラーゼ活性を、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10 μl のDME含有サンプルを、0.5 mM 4-酢酸メチルウンベルフェリルを含む1 mlの反応バッファ (90 mM KH₂PO₄, 40 mM KCl, pH 7.3) 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 ($\epsilon_{350} = 12,200 \text{ M}^{-1} \cdot \text{cm}^{-1}$) で1分30秒モニタリングする。酵素活性は、1分間に1 mgのタンパク質で生成され生成物の μmol として表すことができ、サンプルにおけるDME活性に非例する (Evgenia, V. 他 (1997) J. Biol. Chem. 272:14769-14775)。

【0375】

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、約0.1 mlの酵素および3.3 mM コカインを、1 mM benzamidine および1 mM EDTAを含む反応バッファ (50 mM NaH₂PO₄, pH 7.4) で37℃でインキュベートして測定する。全量で0.4 mlの反応液を1時間インキュベートし、等量の5%トリクロロ酢酸で終了させる。0.1 mlの内部標準3,4-ジメチル安息香酸 (10 $\mu\text{g}/\text{ml}$) 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、0.4 ml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、100 ml当たり8 μl のジエチルアミンを含む14%アセトニトリル、250 mM KH₂PO₄、pH 4.0に再懸濁してから

10

20

30

40

50

、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)。

【0376】

別法では、水溶性基質である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法によって測定する。この方法では、DME含有サンプルを、6mMタウロコール酸塩の存在下で0.5M Tris-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)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、405nm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)。

10

【0377】

DMEのスルホトランスフェラーゼ活性を、 $[^3\ ^5\text{S}]$ PAPSからフェノールなどのモデル基質内への $^3\ ^5\text{S}$ 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1mLの10mMリン酸バッファ、pH6.4、50μMフェノール、および0.4~4.0μM $[^3\ ^5\text{S}]$ 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、0.2mLの0.1M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0.2mLの0.1M Ba(OH)₂を加え、その後0.2mLのZnSO₄を加える。円心分離して上澄みをはっきりさせ、タンパク質および未反応の $[^3\ ^5\text{S}]$ 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。

20

【0378】

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、2.5μmolイミダゾールHCl (pH6.8)、3.75μgの塩化プロタミン(protoamine chloride)、25nmol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50pmol (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50μlで37℃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。0.1μmol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1pmol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)。

30

40

【0379】

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(0.05M Tris-HCl, pH8.0)で洗浄し、3~5mmのセグメントに切断し、0.15M NaClを含む同じバッファを100μlで4℃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. 270:41

50

72-4179)。

【0380】

別法では、DMEのスルホトランスフェラーゼ活性を、 $[^{35}\text{S}]$ PAPSから固定されたペプチドへの $[^{35}\text{S}]$ 硫酸の転移を測定して決定する。この固定されたペプチドは、C末端システイン残基が付加された成熟P-セレクトリン糖タンパク質リガンド-1ポリペプチドのN末端の15残基である。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(1ml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、140mM Pipes (pH 6.8)、0.3M NaCl、20mM MnCl_2 、50mM NaF、1% Triton X-100、および1mM 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含有サンプルを結合させて行った。このアッセイは、0.5 μCi の $[^{35}\text{S}]$ PAPS (1.7 μM ; 1 Ci = 37 GBq)を加えて開始させた。37℃で30分経過した後に、反応ビーズを65℃で6M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移送された $[^{35}\text{S}]$ 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)。

【0381】

別法では、DMEスルホトランスフェラーゼのアッセイを、50mM Hepes-NaOH (pH 7.0)、250mM スクロース、1mM ジチオスレイトール、14 μM $[^{35}\text{S}]$ PAPS (15 Ci/mmol)、およびドーパミン (25 μM)、 p -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 $[^{35}\text{S}]$ 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、100℃で3分間加熱して停止させる。生成された沈降物を遠心分離によって除去する。次に ^{35}S 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。 ^{35}S 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、反応性生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)。

【0382】

DMEのスクアレノ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 (pH 7.5)、0.01mM FAD、0.2単位のNADPH-チトクロームC (P-450) レダクターゼ、0.01 mM $[^{14}\text{C}]$ スクアレノ (20 μl 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)-オキシドスクアレノの存在を、好適な脂質標準を用いて確認する(Sakakibara, J. 他 (1995) 270:17-20)。

【0383】

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 (pH 8.0)、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、isoprene monoxide、エピクロロヒドリン、epibromohydrin、epifl 40

uorohydrin、グリシドール、1, 2-epoxybutane、1, 2-epoxyhexane、または1, 2-epoxyoctane)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1 ml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0.2 mm×25 mm C₁₈-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1 μmol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)。

10

【0384】

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、1 mM L-キヌレニンおよび1 mM 2-オキシグルタル酸の存在下で、最終容量200 μlの70 μM PLPを含む150 mM Tris 酢酸バッファ(pH 8.0)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330 nm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用い、キサンツレン酸の生成を340 nm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)。

20

【0385】

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ピリドキサル5'-リン酸(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シ酸基質に対する活性をして決定して行う。この反応は、9 μM 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シ酸基質)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。アミノ酸からオキシ酸への半反応の後に、酵素結合PLPのピリドキサミン5'-リン酸(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)。

30

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 μM 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比重走査により行う(Hartth, G. およびHorwitz, M. A. (1999) J. Biol. Chem. 274:4281-4292)。

40

【0386】

18 DMEインヒビターの同定

検査する化合物を、実施例17のアッセイで記載したように、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入れる。DME活性をそれぞれのウェルについて測定し、それぞれの化合物のDME活性を阻害する能力を決定して用量反応

50

曲線を作成する。このアッセイを用いて D M E 活性を促進する分子を同定することが可能であろう。

【0387】

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

【0388】

(表の簡単な説明)

10

表1は、本発明の完全長ポリヌクレオチド配列及びポリペプチド配列に対する系統的な名称を示す。

【0389】

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

【0390】

表3は、推定上のモチーフおよびドメインを含む本発明のポリペプチド配列の構造的な特徴、並びに本発明のポリペプチドの分析に用いた方法、アルゴリズム、および検索可能なデータベースを示す。

20

【0391】

表4は、本発明のポリヌクレオチド配列の組み立てに用いた c D N A 断片およびゲノム D N A 断片のリスト、並びに本発明のポリヌクレオチド配列の選択された断片のリストを示す。

【0392】

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

【0393】

表6は、表5に示す c D N A ライブラリの作製に用いた組織およびベクターを示す付録である。

【0394】

30

表7は、本発明のポリヌクレオチド及びポリペプチドの分析に用いたツール、プログラム、及びアルゴリズム、並びにその説明、引用文献、閾値パラメータを示す。

【表1】

表1

インサイトプロジェクトID	ポリペプチド SEQ ID NO:	インサイト ポリペプチドID	ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチドID
2434655	1	2434655CD1	11	2434655CB1
2516747	2	2516747CD1	12	2516747CB1
7472775	3	7472775CD1	13	7472775CB1
7473323	4	7473323CD1	14	7473323CB1
7472777	5	7472777CD1	15	7472777CB1
1272843	6	1272843CD1	16	1272843CB1
7472790	7	7472790CD1	17	7472790CB1
7473944	8	7473944CD1	18	7473944CB1
2244136	9	2244136CD1	19	2244136CB1
7474327	10	7474327CD1	20	7474327CB1

10

20

30

40

表2

ポリペプチド SEQ ID NO:	インサイト ポリペプチドID	GenBank ID NO:	確率スコア	GenBank 相合体
1	2434655CD1	g5802604	4.70E-210	UDPグルクロン酸トランスフェラーゼ [UGT2A3 [モルモット] (Smith, S.A.他 (1998) Mol. Genet. Metab. 68:68-77)]
2	2516747CD1	g1246787	1.50E-69	UDPガラクトースセラムシドガラクトシルトランスフェラーゼ [ハツカネズミ] (Bosio, A.他 (1996) Genomics 35:223-226)
3	7472775CD1	g2660716	2.10E-40	HNK-1スルホトランスフェラーゼ [ドブネズミ] (Bakker, H.他 (1997) J. Biol. Chem. 272:29942-29946)
		g12711481	0	N-アセチルガラクトサミン4-O-スルホトランスフェラーゼ2 GalNAc4ST-2 [ヒト]
4	7473323CD1	g2660716	1.50E-37	HNK-1スルホトランスフェラーゼ [ドブネズミ] (Bakker, H.他 (1997) J. Biol. Chem. 272:29942-29946)
		g11990885	0	GalNAc 4-スルホトランスフェラーゼ [ヒト] (Okuda, T.他 (2000) J. Biol. Chem. 275:40605-40613)
5	7472777CD1	g5917706	3.70E-178	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 [ヒト] (Lee, J.K.他 (1999) Biochem. Biophys. Res. Commun. 263:543-549)
6	1272843CD1	g402843	2.70E-134	シトクロムP450 2B-Bx, フェノバルビタール誘発性[カイウサギ] (Ryan, R.他 (1993) Arch. Biochem. Biophys. 304:454-463)
7	7472790CD1	g12653085	1.00E-104	N-アセチルトランスフェラーゼ, <i>S. cerevisiae</i> と相同 ARD1 [ヒト]
		g1302661	3.40E-97	ARD1 N-アセチルトランスフェラーゼ反復タンパク質 [ヒト] (Brenner, V.他 (1997) Genomics 44:8-14)
8	7473944CD1	g7340847	0	コンドロイチン4-スルホトランスフェラーゼ [ハツカネズミ] (Yamauchi, S.他 (2000) J. Biol. Chem. 275:8975-8981)
		g2921306	3.10E-39	HNK-1スルホトランスフェラーゼ [ヒト]
9	2244136CD1	g6651438	6.50E-23	推定N-アセチルトランスフェラーゼカメロ2 [ヒト]
10	7474327CD1	g178489	7.80E-153	アルドースレダクターゼ [ヒト] (Chung, S.及び S. LaMendola (1999) J. Biol. Chem. 264:14775-14777)

【表3】

10

20

30

40

表3-1

SEQ ID NO:	インサート ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列, ドメイン及びモチーフ	分析方法及び データベース
1	2434655CD1	527	S3 T120 T169 T200 S435 T520 S70 S343 S62 T141 T204 T243 T258 S296 S297 T419 Y130 Y143	N131 N313 N518	シグナルペプチドM1-C23 シグナル分割M1-C23 膜貫通ドメイン: Y91-N111, L492-F512 UDPグルクロン酸トランスフェラーゼ: G24-K525 UDPグルコシルトランスフェラーゼBL00375: A:S34-L56; B:C125-P165; C:P188-N211; D:I253-C280; E:F293-P342; F:N348-P392; G:H447-H486 UDPグルクロン酸及びUDPグルコシル トランスフェラーゼシグネチャN376-T417 UDPグルクロン酸トランスフェラーゼ (PD000190): G24-K338, V386-E527 UDPグルクロン酸及びグルコシル トランスフェラーゼ: DM00367(P36510)176-459; M178-F460 UDPグルクロン酸トランスフェラーゼ シグネチャW354-Q397 シグナルペプチドM1-A22 シグナル分割M1-L19 膜貫通ドメイン:Y484-C503 UDPグルクロン酸及びUDPグルコシル トランスフェラーゼA23-K521 UDPグルクロン酸及びUDPグルコシル トランスフェラーゼシグネチャN373-T416 UDPグルコシルトランスフェラーゼBL00375: A:S34-L56; B:C121-P161; C:P180-N203; D:L246-E273; F:N345-P389; G:L444-Q483 UDPグルクロン酸トランスフェラーゼ: PD000190: K24-G291, L246-D430, L422-K522 UDPグルクロン酸及びUDPグルコシル トランスフェラーゼ: DM00367(P36513)188-462L177-V456	HMMER SPScan HMMER HMMER-PRAM ProfileScan BLIMPS- BLOCKS ProfileScan BLIMPS- BLOCKS BLAST- PRODOM BLAST-DOMO MOTIFS HMMER SPScan HMMER HMMER-PRAM ProfileScan BLIMPS- BLOCKS BLAST- PRODOM BLAST-DOMO
2	2516747CD1	523	S107 S81 S95 T99 S126 S213 S219 T224 S374 S59 S133 S362 S409 T423 Y74 Y433	N52		

【表 4】

10

20

30

40

表3-2

SEQ ID NO.	インサイトポリペプチド ID	アミノ酸残基数	潜在的リン酸化部位	潜在的グリコシル化部位	シグネチャ(signature)配列, ドメイン及びモチーフ	分析方法及びデータベース
3	7472775CD1	358	S95 S42 S16 S19 T287 S315 T333 T44 S255 T335 Y251	N74 N158 N239 M352	HNK1 スルホトランスフェラーゼ: PD041629N65-L358	BLAST-PRODOM
4	7473323CD1	396	T44 S143 S144 T149 T356 S124 S190 S322 Y286 Y380	N100 N266 N339 N387	HNK1 スルホトランスフェラーゼ: PD041629L86-P391	BLAST-PRODOM
5	7472777CD1	395	S48 T214 T257 S322 T231 S363	N116 N229 N305 N328	シグナルペプチドM1-A35 シグナル分割IM1-S32 スルホトランスフェラーゼPD042480: V242-V375	HMMS SPScan BLAST-PRODOM
6	1272843CD1	504	T131 S119 S192 S343 S61 S99 T248 T288 T378 S473		シトクロームP450: DM00022(S3566)49-475L53-G480 シトクロームP450:P33-R493 シトクロームP450ステインヘム鉄リガンドシグネチャF412-S462 シトクロームP450 PD000021:P33-Q150 シトクロームP450:F433-G442 EクラスP450グループI, PR00463: A:S61-V80; B:A86-L107; C:A177-D195; D:N291-S308; E:V311-G337; F:D354-P372; G:L395-D419; H:F430-Q440; I:C440-L463 EクラスP450グループIV, PR00465: B:S55-V78; C:L293-M319; D:L350-L366; E:F482-D486; F:L400-L418; G:F424-C440; H:C440-L458 P450スーパーファミリー-シグネチャ PR00385: A:A302-M319; B:K320-N333; C:A355-L366; D:L431-C440; E:C440-F451 シグナルペプチドM1-A28 シグナル分割IM1-A28 膜貫通ドメイン: W7-T26, I297-L315	BLAST-DOMO HMMS-PFAM ProfileScan BLAST-PRODOM MOTIFS BLIMPS-PRINTS BLIMPS-PRINTS BLIMPS-PRINTS HMMS SPScan HMMS

10

20

30

40

【表 5】

表3-3

SEQ ID NO:	インサイト ポリペプチド ID	アミノ酸 残基数	潜在的リン酸化部位	潜在的 グリコシル化部位	シグネチャ(signature)配列, ドメイン及びモチーフ	分析方法及び データベース
7	7472790CD1	229	S80 S114 S131 T184 S186 S202		アセチルトランスフェラーゼ(GNAT)ファミリー: I3-Q129 アセチルトランスフェラーゼ: PD071891:1130-E179 アセチルトランスフェラーゼ: DM04629(P41227):1-193:M1-C192 シグナル分割:M1-R38 膜貫通ドメイン:C15-130 スルホトランスフェラーゼ: PD041629:L72-L344	HMME-PFAM BLAST- PRODOM BLAST-DOMO SPScan HMME BLAST- PRODOM
8	7473944CD1	347	T83 T90 S91 T134 T196 T244 S302 T98 S163 T244 T276 T305 S51 Y238	N200 N218 N316 N337	シグナル分割:M1-Q63 アセチルトランスフェラーゼ:L43-R181 アルドノケト レダクターゼファミリー: L7-R299 アルドノケト レダクターゼファミリー: BL00798A:L7-W21 BL00798B:A35-I59 BL00798C:V68-W80 BL00798D:A92-Y108 BL00798E:K179-S216 BL00798F:T240-M288	BLIMPS- BLOCKS
9	2244136CD1	218	S69 S197	N156	アルドノケト レダクターゼ: PR00069A:A35-I59, PR00069B:K95-P113, PR00069C:M147-F164, PR00069D:V183-Y212, PR00069E:L230-F254	BLIMPS- PRINTS
10	7474327CD1	318	T76 S115 T240 S308 S216 S229	N9 N244	アルドノケト レダクターゼ: PD000288:V28-N287 アルドノケト レダクターゼファミリー: DM00192(P07943):1-297:A2-C301 アルドノケト レダクターゼ:M147-F164 アルドノケト レダクターゼファミリー: シグネチャ:P124-Q186, I236-R299	BLAST- PRODOM BLAST-DOMO Motifs ProfileScan

【表 6】

10

20

30

40

表4-1

ポリスクレオチド SEQ ID NO:	インサイト ポリスクレオチドID	配列長さ	選択された断片	配列断片	5'位置	3'位置
11	2434655CB1	1636	1-192, 196-294, 1531-1636, 1271-1328	71447910V1	400	1040
				3700065F6 (SININOT05)	1	615
				2096523R6 (BRAITUT02)	1257	1636
				1500810F6 (SINPEST01)	879	1429
12	2516747CB1	2086	964-1669, 2039-2086	SXEC00120V1	1545	2086
				SXBC01648V1	1189	1753
				SXBC00083V1	1040	1610
				SCSA02735V1	782	1141
13	7472775CB1	1814	1795-1814	SXBC01451V1	1	544
				SXBC01528V1	283	863
				2859356F6 (SININOT03)	611	1039
				3549975H1 (BRNDIT01)	170	416
				GNN:g6094626_000014_002:1-78,79- 1170	94	1185
				4189221H1 (BRSTNOT31)	517	845
				g2077168	1526	1814
				60208402U1	357	812
14	7473323CB1	1650	1-664	4166935T6 (PANCMOT21)	1151	1797
				g5663459	1395	1806
				g880305	1	513
				g6699998	827	1283
				FL152824_00001	1	1637
				6197818H1 (PITUNON01)	1011	1650

【表 7】

10

20

30

40

表4-2

ポリヌクレオチド SEQ ID NO:	インサイト ポリヌクレオチドID	配列長さ	選択された断片	配列断片	5'位置	3'位置
15	7472777CB1	1647	1-120, 802-1295	1593910F6 (BRAINOT14)	1190	1647
				6869651H1 (BRAGNON02)	859	1301
				GNN.g7025744_000055_002	1	1188
16	1272843CB1	2620	1-552, 1122-1286, 1-1345	70792863V1	1150	1832
				5056523H1 (COLATMT01)	2413	2620
				1272843T6 (TESTTUT02)	1796	2594
				70567140V1	991	1560
				6576957H1 (COLHTUS02)	1836	2598
				6820526H1 (SINTNOR01)	209	831
				6823903H1 (SINTNOR01)	1	285
				6820526J1 (SINTNOR01)	377	1060
17	7472790CB1	690	498-690	GNN.g7024025_000010_002	1	690
18	7473944CB1	1510	1213-1510	7391722H1 (LIVRFE02)	517	1113
				966072R1 (BRSTNOT05)	1015	1481
				6304381H2 (NE3DITD03)	1181	1510
				7321512H1 (SPINNOE01)	1	465
				6512789H1 (THYMDIT01)	315	922
19	2244136CB1	3701	354-3701, 1-788, 2902-2953, 694-1425, 1677-1926, 2902-3701	5961077F8 (BRATNOT05)	1259	1997
				7325258H1 (TESTTUE02)	395	997
				7170810H2 (BRSTTMC01)	3023	3701
				5973836H1 (BRAZNOT01)	2649	3198
				7074871H1 (BRAUTDR04)	2145	2755
				1699688F6 (BLADTUT05)	1732	2352
				6637768H1 (BRACDIR02)	666	1352
				688956J1 (BRAITDR03)	1	649
20	7474327CB1	960	1-22	GNN.g7407895_000004_002	1	960

【表 8】

10

20

30

40

表5

ポリヌクレオチド SEQ ID NO:	インサイト プロジェクトID	代表的ライブラリ
11	2434655CB1	SINITMC01
12	2516747CB1	LIVRTUT04
13	7472775CB1	PANCNOT21
14	7473323CB1	PITUNON01
15	7472777CB1	BRAINOT14
16	1272843CB1	SINTNOR01
18	7473944CB1	BRAINOT09
19	2244136CB1	BRAITDR03

【表 9】

10

20

30

40

表 6 - 1

ライブラリ	ベクター	ライブラリの説明
BRAINOT09	pINCY	ライブラリは、妊娠期間 23 週間目に死亡した白人の男胎児の脳組織から単離した RNA を用いて作製した。
BRAINOT14	pINCY	ライブラリは、年齢 40 歳の白人女性の脳腫瘍病巣の切除の際に左前葉から採取された脳組織から単離した RNA を用いて作製した。関連腫瘍組織の病変は、グレード 4 大円形細胞性星状膠腫を示していた。
BRAITDR03	PCDNA2.1	ライブラリは、胆管細胞癌で死亡した年齢 55 歳の白人女性から採取した異種皮質、後部帯状組織から単離した RNA を用いて作成した。病変は主として円蓋上の軽度の髄膜繊維症、帯状皮質及び視床の白質中に散在する球状軸索、及び臭内皮質と水道周囲灰白質の中のいくつかの神経原線維濃縮体を示した。関連腫瘍組織は、残留若しくは再発性の腫瘍を伴った肝臓の高分化肝芽腫を示した。家族歴は、肝芽腫、手術後の Budd-Chiari 症候群、胆管の腹水、水胸、脱水、栄養失調、乏尿、及び急性腎不全を含む。過去の手術歴には肝臓の 85% の切除、胆嚢摘除術を含む。
LIVRTUT04	pINCY	ライブラリは、50 歳の白人男性の部分肝切除の際に採取した腎腫瘍素組織から単離した RNA を用いて作製した。病理学的には、グレード 3 のヘパトームであって、腫瘤を形成していた。患者の既往歴には、良性高血圧症、及び肝炎があった。B 型肝炎コア抗原及び B 型肝炎表面抗原が患者に現われていた。
PANCNOT21	pINCY	ライブラリは、酸素欠乏症で死亡した 8 歳の黒人男児の脾臓組織より単離した RNA を用いて作製された。
SINTNOR01	PCDNA2.1	ランダムプライマー法によるライブラリは、31 歳の白人女性からルー-Y 胃バイパスの際に採取された小腸組織から単離した RNA を用いて作製した。患者の病歴には、臨床的な肥満が含まれる。

【表 10】

10

20

30

40

ライブラリ	ベクター	ライブラリの説明
PITUNO01	pINCY	この標準化された下垂体組織ライブラリは、ある下垂体ライブラリからの 6.92×10^6 個の独立したクローンから作成した。開始 RNA は、慢性閉塞性肺疾患で死亡した 55 歳の男性から採取した下垂体組織から作成された。神経病理学的には、軽度の脳室拡張以外は大きな異常は見られなかった。特に、前頭葉において、錐体細胞染色で銀色の陽性がいくつか見られる以外は、試験した新皮質領域に明らかな微視的な異常は何もなかった。この有意性は未決定である。他の微視的な異常としては、海馬の前部及び後部の CA3 領域中にいくつかの腫脹した軸索を伴った、銀染色が目立つくらいである。小脳の微視的部分には、プルキンエ細胞中に軽度のベルグマン神経膠症が見られる。患者歴は、精神分裂病を含む。このライブラリは、著しく長い (48 時間/回) アンダーリングハイブリダイゼーションが用いられたの以外は Soares 他 PNAS (1994) 91:9228 及び Bonaldo 他 Genome Research (1996) 6:791 の条件に従って 2 回標準化した。
SIN1TMC01	pINCY	この大きくサイズ分割されたライブラリは、2 人のドナーからのプールされた cDNA を用いて作製された。cDNA は、部分的結腸切除、開腹での肝臓の生検、及び永久的人工肛門形成の際に 30 歳の白人女性 (ドナー A) から採取された回腸組織と；右側結腸半切除、開腹での肝臓の生検、S 状結腸鏡検査、大腸鏡検査、及び永久的人工肛門形成の際に 70 歳の白人女性 (ドナー B) から採取された回腸組織とから単離された mRNA を用いて作製された。関連腫瘍組織 (ドナー A) の病変は、回腸末端部に生じたカルチノイド腫瘍 (グレード 1 の神経内分泌癌) を示した。この組織は、回腸壁を通り、腸間膜脂肪の中へと浸潤し、癒着盲腸にまで延在しており、そこでは腫瘍は腸壁を通して延在して粘膜表面にまで達していた。複数のリンパ節が腫瘍に対して陽性であった。更にリンパ節 (2) には直接腫瘍の拡張部分が含まれているものもあった。病理学的には、ドナー B では腫瘍状でない回腸の縁を示した。関連腫瘍組織 (ドナー B) の病変は、回盲弁の末端部に位置し、潰瘍性の塊を形成するグレード 2 の浸潤性線癌を示した。腫瘍は、筋層の固有層を通して、漿膜層の脂肪組織の中まで浸潤していた。ある局所性リンパ節は、転移線癌の微小焦点に陽性であった。ドナー A には、潮紅及び不特定の腹部/骨盤の症状があった。患者の病歴は、子宮内膜症、及びニコチン及びアルコール中毒を含む。ドナー B の病歴は、胸部の悪性腫瘍、II 型糖尿病、高脂血症、ウイルス性肝炎、不特定甲状腺障害、変形性関節症、及び皮膚の悪性腫瘍を含む。ドナー B の投薬には tamoxifen を含む。

【表 1 1】

表7-1

プログラム名	説明	引用文献	パラメーター閾値
ABI FACTURA	核酸配列においてベクター配列を除去して不定の塩基をマスクするプログラム。	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	Fast Data Finder は、アミノ酸または核酸配列の比較及び注釈付け (annotation) に有用である。	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	不一致50%
ABI AutoAssembler	核酸配列を構築するプログラム。	Applied Biosystems, Foster City, CA.	
BLAST	Basic Local Alignment Search Tool は、アミノ酸及び核酸配列の配列類似性検索に有用であり、blastp 及び blastn, blastx, tblastn, tblastx の5つのファンクションがある。	Altschul, S.F. 他 (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. 他 (1997) Nucleic Acids Res. 25: 3389-3402.	ESTs: 確率値=1.0E-8 以下 完全長配列: 確率値=1.0E-10 以下
FASTA	Pearson 及び Lipman アルゴリズムは、問合わせの配列と同種の配列群との類似性を検索する。FASTA は、fasta 及び tfasta、fastx、tfastx、ssearch の少なくとも5つのファンクションを含む。	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; Smith, J.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E 値=1.06E-6 構築された ESTs: fasta 同一性=95%以上、一致長さ=200 塩基以上、 fastx E 値=1.0E-8 以下 完全長配列: fastx スコア=100 以上
BLIMPS	Blocks IMPROVED Searcher は、BLOCKS 及び PRINTS、DOMO、PRODOM、PFAM データベースにおける配列に対して、ある配列の一致性を調べ、遺伝子ファミリー及び配列相同性、構造的フィンガープリント領域を探索する。	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.C. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; Artwood, T.K. 他 (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	確率値=1.0E-3 以下
HMMER	PFAM などのタンパク質ファミリーコンセンサス配列の隠れマルコフモデル (HMM) に基づいたデータベースに対して問合せ配列を検索するアルゴリズム。	Krogh, A. 他 (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. 他 (1998) Nucleic Acids Res. 26:320-322; Durbin, R. 他 (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	PFAM ヒット: 確率値=1.0E-3 以下 シグナルペプチドヒット: スコア=0 以上

10

20

30

40

【表 1 2】

表7-2

プログラム名	説明	引用文献	パラメーター値
ProfileScan	Prosite で定義された配列パターンと一致するタンパク質配列における構造及び配列のモチーフを検索するアルゴリズム。	Gribkov, M. 他 (1988) CABIOS 4:61-66; Gribkov, 他 (1989) Methods Enzymol. 183:146-159; Bairoch, A. 他 (1997) Nucleic Acids Res. 25:217-221	標準化された質のスコア≧特定の Prosite モチーフに対する GCG 指定 "HIGRP" 値 通常、スコア=1.4-2.1
Phred	高い感度及び確率で自動配列決定機のトレースを調べる塩基読出しアルゴリズムである。	Ewing, B. 他 (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	Smith-Waterman アルゴリズムの効率的なインプリメンテーションに基づく SWAT や CrossMatch を含む Phil's Revised Assembly プログラムであって、配列相同性の検索及び DNA 配列の構築に有用である。	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; Green, P., University of Washington, Seattle, WA.	スコア=120 以上 一致長さ=56 以上
Consed	Phrap で構築したものの表示及び編集をするためのグラフィックツールである。	Gordon, D. 他 (1998) Genome Res. 8:195-202.	
SPScan	タンパク質配列をスキヤンして分泌シグナルベプチドの存在を調べる重み付けマトリクス解析プログラムである。	Nielson, H. 他 (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	スコア=3.5 以上
TMAP	蛋白配列での膜貫通セグメントの描写及び配向の決定のための、加重マトリクスを用いたプログラムである。	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMM	蛋白配列での膜貫通セグメントの描写及び配向の決定のための、隠された Markov モデル (HMM) を用いたプログラムである。	Sonnhammer, E.L. 他 (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et 他., es., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	Prosite で定義された配列と一致したパターンについてアミノ酸配列を検索するプログラムである。	Bairoch 他, (1997) Nucleic Acid Res. 25: 217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

【国際公開パンフレット】

WO 01/79468

PCT/US01/11869

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number
WO 01/79468 A2

(51) International Patent Classification: C12N 9/00

(21) International Application Number: PCT/US01/11869

(22) International Filing Date: 12 April 2001 (12.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/197,590	13 April 2000 (13.04.2000)	US
60/198,403	19 April 2000 (19.04.2000)	US
60/200,185	28 April 2000 (28.04.2000)	US
60/202,234	5 May 2000 (05.05.2000)	US
60/203,509	11 May 2000 (11.05.2000)	US

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

US	60/198,403 (CIP)
Filed on	19 April 2000 (19.04.2000)
US	60/197,590 (CIP)
Filed on	13 April 2000 (13.04.2000)
US	60/200,185 (CIP)
Filed on	28 April 2000 (28.04.2000)
US	60/202,234 (CIP)
Filed on	5 May 2000 (05.05.2000)
US	60/203,509 (CIP)
Filed on	11 May 2000 (11.05.2000)

(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): POLICKY, Jennifer, L. [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US). IIAFALLA, April [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). BURFORD, Neil [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US). BING, Huijun, Z. [US/US]; 625 Orange Avenue, Los Altos, CA 94023 (US). LAI, Freeth [DN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). TRIBOULEY, Catherine, M. [FR/US]; 1121 Tennessee Street, #5, San Francisco, CA 94107 (US). YAO, Montque, G. [US/US];

111 Frederick Court, Mountain View, CA 94013 (US). YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). PATTERSON, Chandru [US/US]; 460 Sherwood Way #1, Menlo Park, CA 94025 (US). DAS, Debopriya [IN/US]; Apartment 3, 1179 Ramita Avenue, Mountain View, CA 94034 (US). SANJANWALA, Mndhu, S. [US/US]; 210 Sylvia Court, Los Altos, CA 94024 (US). GANDHI, Ameena, R. [US/US]; 837 Robie Avenue, #1, Menlo Park, CA 94025 (US). REDDY, Roopa [US/US]; 1233 W. McKinley Avenue, #3, Sunnyvale, CA 94086 (US). KHAN, Farrah, A. [IN/US]; 333 Escuelin Avenue, #221, Mountain View, CA 94040 (US). BAUGHIN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). RAMKUMAR, Jayn [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). GRIFFIN, Jennifer, A. [US/US]; 181 Irene Court, #11, Belmont, CA 94002 (US). AU-YOUNG, Janice [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US).

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

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

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

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.

(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 01/79468 A2

WO 01/79468

PCT/US01/11869

DRUG METABOLIZING ENZYMES

TECHNICAL FIELD

5 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.

10

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). Additionally, some nontoxic compounds (e.g. aflatoxin, benzo[a]pyrene) are metabolized to toxic

35

WO 01/79468

PCT/US01/11869

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., Andur, 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 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 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 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 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, 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. 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

WO 01/79468

PCT/US01/11869

signature; Graham-Lorence, S. and Peterson, J.A. (1996) *FASEB J.* 10:206-214.)

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).

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 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 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 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 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-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

WO 01/79468

PCT/US01/11869

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; Kitanaka, 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.* 462:283-288) 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 (ferrohemoglobin) 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).

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 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,

WO 01/79468

PCT/US01/11869

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 (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).

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, 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 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 cholesterol in the Achilles' tendons, brain, lungs, and many other tissues. The disease presents with progressive neurologic dysfunction, including postpubescent cerebellar ataxia, atherosclerosis, and cataracts. 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; Gambini, G.T. et al. (1985) J. Clin. Invest. 75:954-960; and W.L. and Portale, A.A. *supra*).

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 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. 267:632-639).

WO 01/79468

PCT/US01/11869

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.

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 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-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.

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 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

WO 01/79468

PCT/US01/11869

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.

- 5 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
- 10 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

- 15 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 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
- 25 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
- 30 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).

Aldo/keto reductases

- Aldo/keto reductases are monomeric NADPH-dependent oxidoreductases with broad substrate specificities (Bohren, K.M. et al. (1989) J. Biol. Chem. 264:9547-9551). These enzymes
- 35 catalyze the reduction of carbonyl-containing compounds, including carbonyl-containing sugars and

WO 01/79468

PCT/US01/11869

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. 273:11429-11435).

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, are unable to oxidize methanol, and are not inhibited by pyrazole. Class III isozymes prefer even longer chain aliphatic alcohols (five 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).

WO 01/79468

PCT/US01/11869

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
 5 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
 10 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
 15 signature).

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
 20 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 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. ST substrates are predominantly phenols and aliphatic alcohols, but also include aromatic amines and
 30 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
 35 characterized STs fall into five classes based on their substrate preference: arylsulfotransferase,

WO 01/79468

PCT/US01/11869

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 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 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-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. 259:13751-13757; OMIM *217800 Macular 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 Hennet, 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

WO 01/79468

PCT/US01/11869

sequence motif (Hennet, T. *supra*). Recent work suggests that brainiac protein is a β 1,3-galactosyltransferase (Yuan, Y. et al. (1997) *Cell* 88:9-11; and Hennet, T. *supra*).

UDP-Gal:GlcNAc-1,4-galactosyltransferase (β 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, 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 $\sim$ 7.5-9.0 (Alpha), and pI $\sim$ 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.* 274:549-555).

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

WO 01/79468

PCT/US01/11869

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 90:8567-8580). The mutagenicity of ethylene dibromide and ethylene dichloride is increased in
 5 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 14:1371-1376). 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
 10 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 then
 15 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: 6215-6220). Thus control of GST activity in cancerous tissues may be useful in treating MDR in cancer patients.

Gamma-glutamyl transpeptidase

20 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
 25 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; Tamiguchi, 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. 122:367-380).

Acyltransferase

30 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

WO 01/79468

PCT/US01/11869

metabolic 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 (Palany, C.N. et al. (1994) J. Biol. Chem. 269:19375-19379; Johnson, M.R. et al. (1991) J. Biol. Chem. 266:10227-10233). BAT is also useful as a predictive indicator for prognosis of hepatocellular carcinoma patients after partial hepatectomy (Furutani, M. et al. (1996) Hepatology 24:1441-1445).

10 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 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, TFIIIF, 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/itbm-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>).

30 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,

WO 01/79468

PCT/US01/11869

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, procinamide, dapsone, 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 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

WO 01/79468

PCT/US01/11869

metabolite L-kynurenine to form kynurenic acid. The enzyme may also catalyzes 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).

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 the 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 S_N2 -like methylation reaction requires Mg^{++} and is inhibited by Ca^{++} . The binding of the donor and substrate to COMT occurs sequentially. AdoMet first binds 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-propiothetropolone) 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 -methyldopa. 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) Pharmacol. Rev. 51:593-628).

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 O_2 and H_2O_2 . The rate of 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

WO 01/79468

PCT/US01/11869

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

5 (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.

15 273:22921-22928).

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).

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, 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

35 topoisomerase I and DNA (Pouliot, J.J. et al. (1999) Science 286:552-555; Yang, S.-W. (1996) Proc.

WO 01/79468

PCT/US01/11869

Natl. Acad. Sci. USA 93:11534-11539).

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).

Glycerophosphoryl diester phosphodiesterase (also known as glycerophosphodiester phosphodiesterase) is a phosphodiesterase which hydrolyzes deacetylated phospholipid glycerophosphodiesters 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).

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) Am. J. Resp. Crit. Care Med. 157:351-370).

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) Prog. Nucleic Acid Res. Mol. Biol. 63:1-38). The existence of multiple PDE families, 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

WO 01/79468

PCT/US01/11869

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.* 47:893-906). 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 (Juilfs, 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 flxazinone. 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 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 (McAlister-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

WO 01/79468

PCT/US01/11869

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-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 dipyridamole (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).

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, dipyridamole, or IBMX (3-isobutyl-1-methylxanthine), but they are sensitive to the PDE5 inhibitor zaprinast (Fisher, D.A. et al.

WO 01/79468

PCT/US01/11869

(1998) *J. Biol. Chem.* 273:15559-15564; Soderling, S.H. et al. (1998) *J. Biol. Chem.* 273:15553-15558).

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;

5 Fujishige, K. et al. (1999) *J. Biol. Chem.* 274:18438-18445; Loughney, K. et al (1999) *Gene* 234:109-117).

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.* 10 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 15 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 (Turko, L.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 20 extensive similarity (>60%) outside the catalytic domain; while across families, there is little or no sequence similarity outside this domain.

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 25 with decreased levels of cyclic nucleotides. For example, a form of diabetes insipidus in mice has 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.* 30 *Crit. Care Med.* 157:351-370). PDE3 inhibitors are being developed as antithrombotic agents, antihypertensive agents, and as cardiotonic 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 35 in vitro. Therefore, rolipram may inhibit HIV-1 replication (Angel, J.B. et al. (1995) *AIDS*

WO 01/79468

PCT/US01/11869

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).

5 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
10 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
15 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
20 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).

Phosphotriesterases

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

35 Thioesterases

WO 01/79468

PCT/US01/11869

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 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).

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

20 Carboxylesterases

Mammalian carboxylesterases constitute a multigene family expressed in a variety of tissues 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, dipivefrin hydrochloride, cocaine, salicylates, capsaicin, palmitoyl-coenzyme A, imidapril, haloperidol, pyroglutidine alkaloids, steroids, p-nitrophenyl acetate, malathion, butanilicaine, 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).

35 Neuroligins are a class of molecules that (i) have N-terminal signal sequences, (ii) resemble

WO 01/79468

PCT/US01/11869

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, neuroligins 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).

5 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

10 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):

15 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

20 (e.g., coronary arteries) results in decreased blood flow and potential necrosis of the tissues deprived of adequate 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

25 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).

30 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

35 hydrolases from *Pseudomonas putida*, and haloalkane dehalogenase from *Xanthobacter*

WO 01/79468

PCT/US01/11869

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

5 cis-9,10-epoxyoctadec-9(Z)-enoic acid (leukotoxin) to form its corresponding diol, threo-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 threo-12,13-dihydroxyoctadec-9(Z)-enoic acid (isoleukotoxin diol). Leukotoxins alter membrane permeability and ion transport and cause inflammatory responses. In addition, epoxide carcinogens

10 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

15 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).

Enzymes involved in tyrosine catalysis

20 The degradation of the amino acid tyrosine to either succinate and pyruvate or fumarate and acetoacetate, requires a large number of enzymes and generates a large number of intermediate compounds. In addition, many xenobiotic compounds may be metabolized using one or more reactions that are part of the tyrosine catabolic pathway. While the pathway has been studied primarily in bacteria, tyrosine degradation is known to occur in a variety of organisms and is likely to

25 involve many of the same biological reactions.

The enzymes involved in the degradation of tyrosine to succinate and pyruvate (e.g., in *Arthrobacter* species) include 4-hydroxyphenylpyruvate oxidase, 4-hydroxyphenylacetate 3-hydroxylase, 3,4-dihydroxyphenylacetate 2,3-dioxygenase, 5-carboxymethyl-2-hydroxymuconic semialdehyde dehydrogenase, *trans,cis*-5-carboxymethyl-2-hydroxymuconate isomerase,

30 homoprotocatechuate isomerase/decarboxylase, *cis*-2-oxohept-3-ene-1,7-dioate hydratase, 2,4-dihydroxyhept-*trans*-2-ene-1,7-dioate aldolase, and succinic semialdehyde dehydrogenase.

The enzymes involved in the degradation of tyrosine to fumarate and acetoacetate (e.g., in *Pseudomonas* species) include 4-hydroxyphenylpyruvate dioxygenase, homogentisate 1,2-dioxygenase, maleylacetoacetate isomerase, and fumarylacetoacetase. 4-hydroxyphenylacetate

35 1-hydroxylase may also be involved if intermediates from the succinate/pyruvate pathway are

WO 01/79468

PCT/US01/11869

accepted.

Additional enzymes associated with tyrosine metabolism in different organisms include 4-chlorophenylacetate-3,4-dioxygenase, aromatic aminotransferase, 5-oxopent-3-ene-1,2,5-tricarboxylate decarboxylase, 2-oxo-hept-3-ene-1,7-dioate hydratase, and 5-carboxymethyl-2-hydroxymuconate isomerase (Ellis, L.B.M. et al. (1999) *Nucleic Acids Res.* 27:373-376; Wackett, L.P. and Ellis, L.B.M. (1996) *J. Microbiol. Meth.* 25:91-93; and Schmidt, M. (1996) *Amer. Soc. Microbiol. News* 62:102).

In humans, acquired or inherited genetic defects in enzymes of the tyrosine degradation pathway may result in hereditary tyrosinemia. One form of this disease, hereditary tyrosinemia I (HTI) is caused by a deficiency in the enzyme fumarylacetoacetate hydrolase, the last enzyme in the pathway in organisms that metabolize tyrosine to fumarate and acetoacetate. HTI is characterized by progressive liver damage beginning at infancy, and increased risk for liver cancer (Endo, F. et al. (1997) *J. Biol. Chem.* 272:24426-24432).

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.

20

SUMMARY OF THE INVENTION

The invention features purified polypeptides, drug metabolizing enzymes, referred to collectively as "DME" and individually as "DME-1," "DME-2," "DME-3," "DME-4," "DME-5," "DME-6," "DME-7," "DME-8," "DME-9," and "DME-10." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-10.

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:1-10, b) a naturally occurring polypeptide comprising an amino acid

35

WO 01/79468

PCT/US01/11869

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

- 5 In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-10. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:11-20.

- Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

- The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10. 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:1-10, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10.

WO 01/79468

PCT/US01/11869

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: 11-20, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO: 11-20, 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: 11-20, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO: 11-20, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO: 11-20, b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO: 11-20, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1-10, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence

WO 01/79468

PCT/US01/11869

selected from the group consisting of SEQ ID NO:1-10, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, and a pharmaceutically acceptable excipient. In one
 5 embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-10. 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 composition.

The invention also provides a method for screening a compound for effectiveness as an
 10 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:1-10, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10,
 15 and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of
 20 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:1-10, b) a naturally occurring
 25 polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10. The method comprises a) exposing a sample comprising the
 30 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.

35 The invention further provides a method of screening for a compound that specifically binds

WO 01/79468

PCT/US01/11869

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: 1-10, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO: 1-10, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 1-10, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 1-10. 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: 1-10, b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO: 1-10, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 1-10, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO: 1-10. 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 sequence selected from the group consisting of SEQ ID NO: 11-20, 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: 11-20, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO: 11-20, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the

WO 01/79468

PCT/US01/11869

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:11-20, ii) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:11-20, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

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

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

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

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

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

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

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

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is

WO 01/79468

PCT/US01/11869

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

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

10 Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and
15 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
20 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
25 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
30 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
35 polypeptide with at least one functional characteristic of DME. Included within this definition are

WO 01/79468

PCT/US01/11869

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

5 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

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

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide,

15 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.

20 "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,

25 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.

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

30 determinant. 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,

35 thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to

WO 01/79468

PCT/US01/11869

immunize the animal.

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

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

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic 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.).

WO 01/79468

PCT/US01/11869

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

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

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

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

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

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by

WO 01/79468

PCT/US01/11869

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

5 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.

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, 10 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 15 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:11-20 comprises a region of unique polynucleotide sequence that 20 specifically identifies SEQ ID NO:11-20, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:11-20 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:11-20 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:11-20 and the region of SEQ ID NO:11-20 to which the fragment corresponds are routinely 25 determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-10 is encoded by a fragment of SEQ ID NO:11-20. A fragment of SEQ ID NO:1-10 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-10. For example, a fragment of SEQ ID NO:1-10 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1- 30 10. The precise length of a fragment of SEQ ID NO:1-10 and the region of SEQ ID NO:1-10 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "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 35 "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

WO 01/79468

PCT/US01/11869

"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 3.12c sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. 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:

30 Matrix: BLOSUM62
 Reward for match: 1
 Penalty for mismatch: -2
 Open Gap: 5 and Extension Gap: 2 penalties
 Gap x drop-off: 50
 35 Expect: 10

WO 01/79468

PCT/US01/11869

*Word Size: 11**Filter: on*

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

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

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

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

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

*Matrix: BLOSUM62**Open Gap: 11 and Extension Gap: 1 penalties**Gap x drop-off: 50**Expect: 10*

Word Size: 3

WO 01/79468

PCT/US01/11869

Filter: on

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

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

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

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

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

WO 01/79468

PCT/US01/11869

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%.

- 5 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
- 10 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

- 15 hybridization complex may be formed in solution (e.g., C₀t or R₀t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

- The words "insertion" and "addition" refer to changes in an amino acid or nucleotide
- 20 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.

- 25 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 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.

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

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

- The term "modulate" refers to a change in the activity of DME. For example, modulation
- 35 may cause an increase or a decrease in protein activity, binding characteristics, or any other

WO 01/79468

PCT/US01/11869

biological, functional, or immunological properties of DME.

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

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

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition.

PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an 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.

"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 DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

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

Methods for preparing and using probes and primers are described in the references, for

WO 01/79468

PCT/US01/11869

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
 5 can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to
 10 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from
 15 megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection
 20 programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both
 25 unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence
 30 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
 35 been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently,

WO 01/79468

PCT/US01/11869

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 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.

WO 01/79468

PCT/US01/11869

"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.

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

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

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

30 A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least

WO 01/79468

PCT/US01/11869

99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

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

THE INVENTION

The invention is based on the discovery of new human 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.

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

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the

WO 01/79468

PCT/US01/11869

polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank homolog. Column 5 shows the annotation of the GenBank homolog along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

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

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are drug metabolizing enzymes. For example, SEQ ID NO:9 is 32% identical to human putative N-acetyltransferase Camello 2 (GenBank ID g6651438) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 6.50E-23, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:9 also contains a signal peptide signature sequence as determined by the SPScan program, and an acetyltransferase domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. The probability value of the HMM comparison to acetyltransferase domain is 3.2e-12. (See Table 3.) Based on BLAST and HMM analyses, the protein of SEQ ID NO:9 is an N-acetyltransferase which N-acetylates aromatic amines and hydrazine-containing compounds including but not limited to para-aminobenzoic acid para-

WO 01/79468

PCT/US01/11869

database (See Table 7 for descriptions). All of these signature sequence identifications are highly statistically significant (see Table 7 for threshold values); for example, the probability value of the HMM comparison to the aldo-keto reductase family signature sequence is 5.6e-170. Based on BLAST, HMM, BLIMPS, MOTIFS, and ProfileScan analyses, the protein of SEQ ID NO:10 is an aldo/keto reductase which reduces carbonyl-containing sugars and aromatic compounds, including, but not limited to, glucose, and carbonyl-containing drug molecules and xenobiotics. SEQ ID NO:1-8 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-10 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:11-20 or that distinguish between SEQ ID NO:11-20 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 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, 3700065F6 is the identification number of an Incyte cDNA sequence, and SININOT05 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., 71447910V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g5663459) which contributed to the assembly of the full length polynucleotide sequences. Alternatively, the identification numbers in column 5 may refer to coding regions predicted by Genscan analysis of genomic DNA. For example, GNN.g7025744_000055_002 is the identification number of a Genscan-predicted coding sequence, with g7025744 being the GenBank identification number of the sequence to which Genscan was applied. The Genscan-predicted coding sequences may have been edited prior to assembly. (See Example IV.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For

WO 01/79468

PCT/US01/11869

example, FL152824_00001 represents a "stitched" sequence in which 152824 is the identification number of the cluster of sequences to which the algorithm was applied, and 00001 is the number of the prediction generated by the algorithm. (See Example V.) Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon-stretching" algorithm. (See Example V.) In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

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

The invention also encompasses 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.

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:11-20, which encodes DME. The polynucleotide sequences of SEQ ID NO:11-20, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding 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 polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:11-20 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:11-20. 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.

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

WO 01/79468

PCT/US01/11869

produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring 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 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:11-20 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

WO 01/79468

PCT/US01/11869

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 sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) *PCR Methods Applic.* 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) *Nucleic Acids Res.* 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) *PCR Methods Applic.* 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) *Nucleic Acids Res.* 19:3055-3060.) Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

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

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the

WO 01/79468

PCT/US01/11869

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
5 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
10 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
15 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
20 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
25 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
30 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.

35 In another embodiment, sequences encoding DME may be synthesized, in whole or in part,

WO 01/79468

PCT/US01/11869

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.

- 5 (See, e.g., Creighton, T. (1984) *Proteins, Structures and Molecular Properties*, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of DME, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide
10 or a polypeptide having a sequence of a naturally occurring polypeptide.

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

- 15 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 inducible promoters, and 5' and 3' untranslated regions in the vector and in
20 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 the appropriate expression vector, no additional transcriptional
25 or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See,
30 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 vivo* genetic recombination. (See, e.g., Sambrook, J. et al. (1989)
35 *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and

WO 01/79468

PCT/US01/11869

16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding DME. These include, but are not limited to, microorganisms such as bacteria
 5 transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., T1 or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke,
 10 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
 15 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.
 20 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
 25 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
 30 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
 35 promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such

WO 01/79468

PCT/US01/11869

vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, *supra*; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; and Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184.)

5 Plant systems may also be used for expression of DME. Transcription of sequences encoding 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. 10 (1984) *Science* 224:838-843; and Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., *The McGraw Hill Yearbook of Science and Technology* (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases 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 20 virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino 25 polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of 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 30 expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be 35 propagated using tissue culture techniques appropriate to the cell type.

WO 01/79468

PCT/US01/11869

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 *aprt*⁻ cells, respectively. (See, e.g., Wigler, M. et al. (1977) *Cell* 11:223-232; Lowy, I. et al. (1980) *Cell* 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) *Proc. Natl. Acad. Sci. USA* 77:3567-3570; Colbere-Garapin, F. et al. (1981) *J. Mol. Biol.* 150:1-14.) Additional selectable genes have been described, e.g., *tpb* 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 membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of 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 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

WO 01/79468

PCT/US01/11869

MN, Sect. IV; Coligan, J.E. et al. (1997) *Current Protocols in Immunology*, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) *Immunochemical Protocols*, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and
 5 may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding DME include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the 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
 10 commercially available, and may be used to synthesize RNA probes *in vitro* by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent,
 15 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 produced by a transformed cell may be secreted or retained intracellularly depending on the
 20 sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides 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 inserted sequences or to process the expressed protein in the desired fashion. Such modifications of
 25 the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the
 30 American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding 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
 35 containing a heterologous moiety that can be recognized by a commercially available antibody may

WO 01/79468

PCT/US01/11869

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., Coligan, J.F. 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. 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.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is

WO 01/79468

PCT/US01/11869

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. 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 as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding DME may also be manipulated *in vitro* in ES cells derived from

WO 01/79468

PCT/US01/11869

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).

5 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
10 treated with 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 (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists
15 between regions of DME and drug metabolizing enzymes. In addition, the expression of DME is closely associated with a liver tumor cell line, liver tumor tissue, pancreatic tissue, pituitary gland tissue, brain tissue, small intestine tissue, fetal brain tissue, and allocortex brain tissue. 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
20 of disorders associated with increased DME expression or activity, it is desirable to 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
25 activity of DME. Examples of such disorders include, but are not limited to, an autoimmune/inflammatory disorder, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED),
30 bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis,
35 polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome,

WO 01/79468

PCT/US01/11869

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, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a developmental disorder, such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucopolysaccharidosis, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea and cerebral palsy, spina bifida, anencephaly, craniorachischisis, congenital glaucoma, cataract, and sensorineural hearing loss; 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, 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) 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 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-Gordon syndrome, pheochromocytoma tumors, and Addison's disease; disorders associated with gonadal steroid

WO 01/79468

PCT/US01/11869

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, retinoblastic tumor, 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, hyperadrenism, hypoadrenism, 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; hypocalcemia, hypophosphatemia, and postpubescent cerebellar ataxia, tyrosinemia, 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, 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, proclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas.

In another embodiment, a vector capable of expressing DME or a fragment or derivative

WO 01/79468

PCT/US01/11869

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.

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.

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.

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. 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 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

WO 01/79468

PCT/US01/11869

which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and *Corynebacterium parvum* are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to DME 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.

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 Cole, S.P. et al. (1984) *Mol. Cell Biol.* 62:109-120.)

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

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population 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

WO 01/79468

PCT/US01/11869

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 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 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, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene

WO 01/79468

PCT/US01/11869

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.)

- 5 In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 10 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, 15 R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

- In another embodiment of the invention, polynucleotides encoding 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 20 immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, 25 familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) 30 (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 DME from an appropriate population of transduced cells may alleviate the 35 clinical manifestations caused by the genetic deficiency.

WO 01/79468

PCT/US01/11869

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

Expression vectors that may be effective for the expression of DME include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, 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 sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. *supra*), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding DME from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to 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.*

WO 01/79468

PCT/US01/11869

- 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).
- 15 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"), hereby incorporated by reference. For adenoviral vectors, see also Antinuzzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 389:239-242, both incorporated by reference herein.
- 25 In another alternative, a herpes-based, gene therapy delivery system is used to deliver 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 been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a
- 35

WO 01/79468

PCT/US01/11869

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
 5 cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to
 10 deliver polynucleotides encoding DME to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA,
 15 resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for 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
 20 persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of 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
 25 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,
 30 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.F. and B.I. Carr, *Molecular and Immunologic Approaches*, Futura Publishing, Mt. Kisco NY, pp. 163-
 35 177.) A complementary sequence or antisense molecule may also be designed to block translation of

WO 01/79468

PCT/US01/11869

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 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

WO 01/79468

PCT/US01/11869

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 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). 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

WO 01/79468

PCT/US01/11869

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

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

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

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

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

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

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising 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-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

WO 01/79468

PCT/US01/11869

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

5 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 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
10 lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD₅₀/ED₅₀ ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED₅₀ with little or no
15 toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include
20 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
25 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.

30 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
35 assays for DME include methods which utilize the antibody and a label to detect DME in human

WO 01/79468

PCT/US01/11869

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.

5 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
10 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,
15 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.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide
20 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 probe identifies only naturally occurring sequences encoding DME, allelic variants, or related
25 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:11-20 or from genomic sequences including promoters, enhancers, and introns of the DME gene.

30 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 polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a
35 variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels,

WO 01/79468

PCT/US01/11869

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 autoimmune/inflammatory disorder, such as acquired immunodeficiency syndrome (AIDS),

- 5 Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis,
- 10 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 syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura,
- 15 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, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers
- 20 including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; a developmental disorder, such as renal tubular acidosis, anemia, Cushing's syndrome,
- 25 achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucopolysaccharidosis, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Sydenham's chorea
- 30 and cerebral palsy, spina bifida, anencephaly, cranio rachischisis, 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, vascular malformations, thrombosis, infections, immunological disorders, and complications due to head trauma; disorders associated with
- 35 hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease,

WO 01/79468

PCT/US01/11869

Hand-Schüller-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) 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 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-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, 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, retinobulbar tumor, 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, McKus syndrome, occipital horn syndrome, mannosidosis, neuraminidase deficiency, obesity, pentosuria, phenylketonuria, pseudovitamin D-deficiency rickets; hypocalcemia, hypophosphatemia, and postpubescent cerebellar ataxia, tyrosinemia, 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

WO 01/79468

PCT/US01/11869

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, 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. 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 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 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 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. 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.

WO 01/79468

PCT/US01/11869

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

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

Additional diagnostic uses for oligonucleotides designed from the sequences encoding 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
15 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
20 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
25 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 amplimers in high-
30 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
35 sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass

WO 01/79468

PCT/US01/11869

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 quantitation.

10 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
15 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
20 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.

25 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
30 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
35 resultant transcript image would provide a profile of gene activity.

WO 01/79468

PCT/US01/11869

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

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

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

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance

WO 01/79468

PCT/US01/11869

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

A proteomic profile may also be generated using antibodies specific for 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

WO 01/79468

PCT/US01/11869

each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

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

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

In another embodiment of the invention, nucleic acid sequences encoding 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 P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Landor, 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-Urich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of

WO 01/79468

PCT/US01/11869

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 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.

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

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

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with DME, or fragments thereof,
25 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, 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
30 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.

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
35 properties of nucleotide sequences that are currently known, including, but not limited to, such

WO 01/79468

PCT/US01/11869

properties as the triplet genetic code and specific base pair interactions.

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

The disclosures of all patents, applications and publications, mentioned above and below, including U.S. Ser. No. 60/197,590, U.S. Ser. No. 60/198,403, U.S. Ser. No. 60/200,185, U.S. Ser. No. 60/202,234, and U.S. Ser. No. 60/203,509, are expressly incorporated by reference herein.

10

EXAMPLES

I. Construction of cDNA Libraries

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

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 POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

25 In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic
30 oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g.,
35 PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid

WO 01/79468

PCT/US01/11869

(Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF⁺, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

5 II. Isolation of cDNA Clones

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

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

III. Sequencing and Analysis

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

The polynucleotide sequences derived from Incyte cDNAs were validated by removing
30 vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and

WO 01/79468

PCT/US01/11869

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

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

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:11-20. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization

WO 01/79468

PCT/US01/11869

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

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

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

WO 01/79468

PCT/US01/11869

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

"Stretched" Sequences

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

VI. Chromosomal Mapping of DME Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:11-20 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:11-20 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

WO 01/79468

PCT/US01/11869

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 G  n  thon 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.

10 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.)

15 Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIPESQ (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:

20

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

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

WO 01/79468

PCT/US01/11869

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 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

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

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PC1 A and PC1 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

WO 01/79468

PCT/US01/11869

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

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

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

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

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

IX. Labeling and Use of Individual Hybridization Probes

35 Hybridization probes derived from SEQ ID NO:11-20 are employed to screen cDNAs,

WO 01/79468

PCT/US01/11869

genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μ Ci of [5'-³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10⁷ counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases:

10 Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

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

15 Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Skena (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

25 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., Skena, 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

30 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,

35 and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser

WO 01/79468

PCT/US01/11869

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.

5 Tissue or Cell Sample Preparation

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

20 Microarray Preparation

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

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US

WO 01/79468

PCT/US01/11869

Patent No. 5,807,522, incorporated herein by reference. 1 μ l of the array element DNA, at an average concentration of 100 ng/ μ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

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

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

Hybridization

- 10 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 larger than a microscope slide. The chamber is kept at 100% humidity internally by the
15 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.

Detection

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

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores.

- 30 Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

WO 01/79468

PCT/US01/11869

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
 5 from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital
 10 (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
 15 emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

20 XI. Complementary Polynucleotides

Sequences complementary to the 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
 25 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.

30 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
 35 promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory

WO 01/79468

PCT/US01/11869

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 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.

25 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;

WO 01/79468

PCT/US01/11869

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 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.

20 XIV. Production of DME Specific Antibodies

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-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

WO 01/79468

PCT/US01/11869

goat anti-rabbit 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
 5 covalently coupling anti-DME antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing 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
 10 antibody/DME binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and DME is collected.

XVI. Identification of Molecules Which Interact with DME

DME, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent.
 15 (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
 25 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
 30 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 MgCl₂, 50 mM
 35 nicotinamide, 40 mg trisodium isocitrate, and 2 units isocitrate dehydrogenase, with 8 mg NADP⁺

WO 01/79468

PCT/US01/11869

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.

- 5 1 α ,25-dihydroxyvitamin D 24-hydroxylase activity of DME is determined by monitoring the conversion of ³H-labeled 1 α ,25-dihydroxyvitamin D (1 α ,25(OH)₂D) to 24,25-dihydroxyvitamin D (24,25(OH)₂D) in transgenic rats expressing DME. 1 μ g of 1 α ,25(OH)₂D 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
- 10 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 homogenate is then incubated with 0.25 nM 1 α ,25(OH)₂[1-³H]D, with a specific activity of approximately 3.5 GBq/mmol, for 15 min at 37 °C under oxygen with constant shaking.
- 15 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 FINEPAK 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
- 20 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 α ,25-dihydroxyvitamin D or 24,25-dihydroxyvitamin D (24,25(OH)₂D), with the chromatograms of the reaction products, the relative mobilities of the substrate (1 α ,25(OH)₂[1-³H]D) and product
- 25 (24,25(OH)₂[1-³H]D) are determined and correlated with the fractions collected. The amount of 24,25(OH)₂[1-³H]D produced in control rats is subtracted from that of transgenic rats expressing DME. The difference in the production of 24,25(OH)₂[1-³H]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
- 30 (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.
- 35 Alternatively, activity is measured by monitoring oxygen uptake using a Clark-type electrode. For

WO 01/79468

PCT/US01/11869

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 MgCl₂, 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 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 sulfate, 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.

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 (CoASH) described below.

WO 01/79468

PCT/US01/11869

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 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 fibrillarin glycine-arginine domain fusion protein (GST-GAR), heterogeneous nuclear ribonucleoprotein (hnRNP), or hypomethylated proteins present in lysates from adenovirus 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).

Catechol-O-methyltransferase activity of DME is measured in a reaction mixture consisting 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 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 catechol-O-methyltransferase activity of DME (Zhu, B.T. and J.G. Ljeher (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

WO 01/79468

PCT/US01/11869

Iwakura, M. (1999) *J. Biol. Chem.* 274:19041-19047).

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 mg 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 / mg 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 / mg of enzyme.

Carboxylesterase activity of 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 ($\epsilon_{330} = 12.2 \text{ mM}^{-1} \text{ cm}^{-1}$) 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 incubating approximately 0.1 ml of enzyme and 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 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 injected onto a C18 reverse-phase HPLC column for separation. The column eluate is monitored at 235 nm. DME activity is quantified by comparing peak area ratios of the analyte to the internal standard. A standard curve is 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 (pH 7.4 or 8.0) or sodium acetate (pH 5.0) in the presence of 6 mM taurocholate. The assay is

WO 01/79468

PCT/US01/11869

initiated by adding a freshly prepared para-nitrophenyl butyric acid solution (100 µg/ml in sodium acetate, pH 5.0). Carboxyl esterase activity is then monitored and compared with control autohydrolysis of the substrate using a spectrophotometer set at 405 nm (Wan, L. et al. (2000) J. Biol. Chem. 275:10041-10046).

5 Sulfotransferase activity of DME is measured using the incorporation of ³⁵S from [³⁵S]PAPS into a model substrate such as phenol (Folds, A. and J.L. Meek (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 mM phenol, and 0.4-4.0 mM [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
10 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.

Heparan sulfate 6-sulfotransferase activity of DME is measured *in vitro* by incubating a
15 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 x 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 1 min. 0.1 µmol (as glucuronic acid) of chondroitin sulfate A is added to the reaction
20 mixture as a carrier. ³⁵S-Labeled polysaccharides are precipitated with 3 volumes of cold ethanol containing 1.3% 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., determined by the amount of [³⁵S]PAPS incorporated into the precipitated polysaccharides (Habuchi, H. et al. (1995) J. Biol.
25 Chem. 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
30 buffer (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 the same buffer containing 0.15 M NaCl for 48 h. The eluted enzyme is collected by centrifugation and assayed for the sulfotransferase activity as described 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 that represents the N-terminal 15 residues
35 of the mature P-selectin glycoprotein ligand-1 polypeptide to which a C-terminal cysteine residue is

WO 01/79468

PCT/US01/11869

- added. 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 $\mu\text{mol peptide/ml}$ 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 MnCl_2 , 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 [^{35}S]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 [^{35}S]sulfate to the head-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 [^{35}S]PAPS as the sulfate donor in a final volume of 30 μl , containing 50 mM Hepes-NaOH (pH 7.0), 250 mM sucrose, 1 mM dithiothreitol, 14 μM [^{35}S]PAPS (15 Ci/nmol), 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 centrifugation. The supernatants are then subjected to the analysis of ^{35}S -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 ^{35}S -sulfated products and determine the enzyme specificity of the DME-containing samples based on relative rates of 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 [^{14}C]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(5)-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,

WO 01/79468

PCT/US01/11869

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

5 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).

10 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,

15 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 and xanthurenic acid, respectively, is indicative of aminotransferase activity (Buchli, R. et al. (1995) *J. Biol. Chem.* 270:29330-29335).

In another alternative, aminotransferase activity of DME is measured by determining the
20 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, pyridoxal 5'-phosphate (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 substrate to be tested (amino and oxo acid substrates). The half-reaction from amino
25 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 pyridoxamine 5' phosphate (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).

Superoxide dismutase activity of DME is assayed from cell pellets, culture supernatants, or
30 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 background, following illumination of the gels on a lightbox. Quantitation of superoxide
35 dismutase activity is performed by densitometric scanning of the activity gels using the appropriate

WO 01/79468

PCT/US01/11869

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).

XVIII. Identification of DME Inhibitors

5 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.

10

 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
15 embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

WO 01/79468

PCT/US01/11869

Table 1

Incyte Project ID	Polypeptide Seq. ID No:	Incyte Polynucleotide ID	Polynucleotide Seq. ID No:	Incyte Polynucleotide ID
243655	1	243655CD1	11	243655CB1
2516747	2	2516747CD1	12	2516747CB1
7472775	3	7472775CD1	13	7472775CB1
7473323	4	7473323CD1	14	7473323CB1
7473777	5	7473777CD1	15	7473777CB1
1272843	6	1272843CD1	16	1272843CB1
7473790	7	7473790CD1	17	7473790CB1
7473944	8	7473944CD1	18	7473944CB1
2246136	9	2246136CD1	19	2246136CB1
7474327	10	7474327CD1	20	7474327CB1

WO 01/79468

PCT/US01/11869

Table 2

Polypeptide Seq ID NO:	Incyte Polypeptide Seq ID NO:	GenBank ID NO:	Probability Score	GenBank Homolog
1	2434655CD1	g5802604	4.70E-210	UDP-glucuronosyltransferase UGT2A3 [Cavia porcellus] (Smith S. A. et al. (1999). Mol. Genet. Metab. 68:68-72)
2	2516741CD1	g1246787	1.50E-69	UDP-galactose 4-epimerase [Mus musculus] (Bostel A. et al. (1995). Genomics 15:223-226)
3	7472775CD1	g2660716	2.10E-40	HKK-1 sulfotransferase [Rattus norvegicus] (Bakker, R. et al. (1997). C. Biol. Chem. 272:2942-2946)
4	7473323CD1	g12711481	0	N-acetylgalactosamine 4-O-sulfotransferase 2 GalNAc4ST-2 [Homo sapiens]
		g2660716	1.50E-37	HKK-1 sulfotransferase [Rattus norvegicus] (Bakker, R. et al. (1997). J. Biol. Chem. 272:2942-2946)
		g11990885	0	GalNAc 4-sulfotransferase [Homo sapiens] (Okuda, T. et al. (2000). J. Biol. Chem. 275:40605-40613)
5	7472777CD1	g5917706	3.70E-178	N-acetylgalactosamine 6-O-sulfotransferase [Homo sapiens] (Lee, J. K. et al. (1999). Biochem. Biophys. Res. Commun. 263:543-549)
6	1272843CD1	g402843	2.70E-134	Cytochrome P450 2B-Bx, phenobarbital-inducible [Cytolagus cuniculus] (Ryan, R. et al. (1993). Arch. Biochem. Biophys. 304:454-463)
7	7472790CD1	g12651085	1.00E-104	N-acetylttransferase, homolog of S. cerevisiae ARD1 [Homo sapiens]
		g1302661	3.40E-97	ARD1 N-acetyl transferase related protein [Homo sapiens] (Brenner, V. et al. (1997). Genomics 44:8-14)
8	7473944CD1	g7340847	0	Chondroitin 4-sulfotransferase [Mus musculus] Yamauchi, S. et al. (2000). J. Biol. Chem. 275:8975-8981)
9	2244136CD1	g2921306	3.10E-39	HKK-1 sulfotransferase [Homo sapiens]
10	7474327CD1	g5651438	6.50E-23	Putative N-acetylttransferase Camello 2 [Homo sapiens]
		g178489	7.80E-153	Aldose reductase [Homo sapiens] (Chung, S. and S. LaMendola (1989). J. Biol. Chem. 264:14775-14777)

WO 01/79468

PCT/US01/11869

Table 3

SEQ ID NO.	Incyte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Software Packages
1	243465CD1	527	S3 T30 T169 T200 S435 T320 S70 S343 S52 T141 T204 T243 T258 S296 S297 T419 Y130 Y143	N131 N313 S518	Signal Peptide: M1-Q23 Signal Cleavage: M1-Q23 Transmembrane domains: V91-M111, I492-P512 UDP-glucoronosyltransferase: G24-K325 UDP-glucosyltransferases BL00375: A: S34-L56; B: C125-P165; C: F188-N211; D: I253-C280; E: F297-S322; F: W348-F394; G: H447-H482 UDP-glucoronosyl and UDP-glucosyl transferases signature: N376-T417 UDP-glucoronosyltransferase (PD000190): G24-K338, V166-E527 UDP-glucoronosyl and glucosyl transferases: G167/P165(I)I26-459, M178-P460 UDP-glucosyltransferase signature: W164-Q197	HMER SPScan HMER HMER-PFAM BLIMPS-BLOCKS ProfileScan BLAST-PRODOM BLAST-DOMO NCI/IF5
2	2516747CD1	523	S107 S81 S95 T59 S126 S213 S219 T224 S374 S59 S133 S362 S409 T423 Y74 Y433	N52	Signal Peptide: M1-A22 Transmembrane domain: Y484-C503 UDP-glucoronosyl and UDP-glucosyl transferase: A23-K521 UDP-glucoronosyl and UDP-glucosyl transferases signature: N373-T416 UDP-glucosyltransferases BL00375: A: S84-S107; B: C124-P163; C: F180-N203; D: I246-L273; E: N145-P189; G: L444-Q483 UDP-glucoronosyltransferase (PD000190): K24-G291, L246-D430, L422-K522 UDP-glucoronosyl and UDP-glucosyl transferases: DM00357/P165(I)I88-462; L177-Y455	HMER SPScan HMER HMER-PFAM ProfileScan BLIMPS-BLOCKS BLAST-PRODOM BLAST-DOMO

WO 01/79468

PCT/US01/11869

Table 3 (cont.)

SEQ ID NO	Inverte Polypeptide	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and
3	747275CD1	356	S95 S12 S16 S19 T287 S315 T333 T44 S255 T335 Y251	N74 N158 N239 N352	HNK1 Sulfotransferase: PD041629; N65-L358	BLAST-PRODOM
4	7473323CD1	396	T44 S143 S164 T149 T356 S124 S190 S322 Y286	N100 N266 N339 N387	HNK1 Sulfotransferase: PD041629; L86-P391	BLAST-PRODOM
5	747277CD1	395	S380 T214 T237 S322 T231 S563	N116 N229 N305 N328	Signal peptide: M1-A15 Signal cleavage: M1-S32 Sulfochromase PD043460; V242-V375	HMMER SPScan BLAST-PRODOM
6	1272843CD1	504	T131 S119 S192 S343 S61 S99 T248 T288 T378 S473		Cytochrome P450: DN00022 [S5566] 49-475; L53-G480 Cytochrome P450: F33-R493 Cytochrome P450 cysteine heme-iron ligand signature: F412-S463 Cytochrome P450 PD00021: P33-Q150 Cytochrome P450: F433-G432 E-class P450 group IV, PR00465: A: S861-V80; B: R86; L1107; C: A177-D195; D: R291-S108; E: V311-G337; F: D354-P372; G: L395-D419; H: F430-C440; I: C440-L463 E-class P450 group IV, PR00465: B: S55-V78; C: L293-M19; D: L350-L566; E: F462-D496; F: L300-L448; G: F424-C440; H: C440-L463 P450 superfamily signature PR00385; A: A102-M19; B: K320-N333; C: A355-L366; D: L431-C440; E: C440-F451 Signal peptide: M1-A28 Signal cleavage: M1-A28 Transmembrane domains: M1-228, 1297-1315	BLAST-PRODOM HMMER SPScan HMMER

WO 01/79468

PCT/US01/11869

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 Software
7	747276CD1	229	S80 S114 S131 T184 S186 S202		Acetyltransferase (GNAT) family: I3-Q129 Acetyltransferase: PD071591: I130-R179 PRODOM Acetyltransferase: PD04529 [P41227] 1-191: M1-C192 Signal cleavage: M1-R38 Transmembrane domain: C15-I30 Sulfotransferase: PD041029: L72-L344	HMER-PFAM BLAST- PRODOM BLAST- SFS- HMER- BLAST- PRODOM
8	747394CD1	347	T83 T90 S91 T134 T186 T244 T244 T296 S305 S51 Y238	N200 N218 N218 N337	Signal cleavage: M1-G63 Acetyltransferase: L43-R181 Aldo/keto reductase family: L7-R299	SFS- HMER-PFAM HMER-PFAM
9	2244136CD1	218	S69 S197	N156	Aldo/keto reductase family: BL00798A: L7-W21 BL00798B: A35-I39 BL00798C: A39-Q106 BL00798D: A32-Y108 BL00798E: K179-S216 BL00798F: T240-M266	BLIMPS- BLOCKS
10	7474327CD1	318	T76 S115 T240 S308 S216 S229	N9 N244	Aldo-keto reductase signature: PR00069A: A35-I59, PR00069B: X35-P112, PR00069C: M147-F164, PR00069D: L133-L242 PR00069E: L233-L254 Aldose reductase: PD000288: V28-N297	BLIMPS- PRINTS
					Aldo/keto reductase family: DN00192 [P07943] 1-297: A2-C301 Aldo-keto reductase 2: M147-F164	BLAST- PRODOM BLAST- DOMO
					signatures: P124-Q186, T235-R299	ProfileScan

WO 01/79468

PCT/US01/11869

Table 4

Polynucleotide Seq ID NO:	Incyte Polynucleotide	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
11	2434655CB1	1636	1-192	71447910V1	400	1040
			196-294	370006576 (SINNOT05)	1	615
			1631-1636	208652386 (BRATUT02)	1357	1636
			1201-1328	130051078 (SINTEST01)	879	1489
			964-1669	SXEC00120V1	1345	2086
12	2516747CB1	2086	2039-2086	SXEC01545V1	1189	1753
				SXEC00083V1	1040	1610
				SCSA02735V1	782	1141
				SXEC01451V1	1	544
				SXEC01528V1	283	863
13	7472775CB1	1814	1795-1814	285935576 (SINNOT03)	611	1039
				3549975H1 (BRONDIT01)	170	415
				GN1:45094626_000014_002.1-79, 79-1170	94	1185
				4189221H1 (BRSTNOT01)	517	845
				92077169	1526	1814
14	7473233CB1	1650	1-664	6020840201	357	812
				4-6693576 (PAKNOT21)	1151	1797
				95663459	1395	1806
				9860305	1	513
				96699998	827	1283
14	7473233CB1	1650	1-664	FL152824_00001	1	1637
				6197816H1 (PITUNON01)	1011	1650

WO 01/79468

PCT/US01/11869

Table 4 (cont.)

Polynucleotide Seq ID NO:	Incyte Polynucleotide	Sequence Length	Selected Fragments	Sequence Fragments	5' Position	3' Position
15	7472777CB1	1647	1-120, 802-1295	1593910F6 (BRATN0014), 6869551H1 (BRACD0002), GMR_27025744_000055_002	1190 859 1	1647 1301 1188
16	1272843CB1	2620	1-552, 1122-1286, 1-1345	70793833V1 (COLAT0001), 5085632H1 (COLAT0001), 1272843T6 (TEST0002), 7056740V1 (COLAT0002), 6176957H1 (COLAT0002), 682026H1 (SINW0001), 682390H1 (SINW0001), 6820265T1 (SINW0001), GMR_27024095_000010_002	1150 2413 1796 991 1836 209 1 377 1	1632 2620 2594 2560 2598 831 285 1060 650
17	7472790CB1	690	498-690	7391722H1 (CTVRF002), 966072R1 (BRST0005), 6304381H2 (NERD0003), 7321512H1 (SPJAN0801), 6512789H1 (THYAD0101), 5961077F8 (BRATN0005), 7170810H2 (BRST0001), 7325258H1 (TEST0002), 7170810H2 (BRST0001), 5973316H1 (BRATN0001), 7074871H1 (BRATN0004), 1677-1926, 2902-3701	517 1015 1181 1 315 1259 395 3023 2649 2145 1732 666 1	1113 1481 1510 465 922 1937 997 3701 3198 2755 2352 1352 949
18	7473944CB1	1510	1213-1510	GMR_27024095_000010_002	1	650
19	2244136CB1	3701	3544-3701, 1-788, 2902-2953, 694-1425, 1677-1926, 2902-3701	7391722H1 (CTVRF002), 966072R1 (BRST0005), 6304381H2 (NERD0003), 7321512H1 (SPJAN0801), 6512789H1 (THYAD0101), 5961077F8 (BRATN0005), 7170810H2 (BRST0001), 7325258H1 (TEST0002), 7170810H2 (BRST0001), 5973316H1 (BRATN0001), 7074871H1 (BRATN0004), 1677-1926, 2902-3701	517 1015 1181 1 315 1259 395 3023 2649 2145 1732 666 1	1113 1481 1510 465 922 1937 997 3701 3198 2755 2352 1352 949
20	7474327CB1	960	1-22	GMR_27407895_000004_002	1	960

WO 01/79468

PCT/US01/11869

Table 5

Polynucleotide Seq ID No:	Incyte Project ID	Representative Library
11	243455CB1	SINHW01
12	251577CB1	243570704
13	747277CB1	PANC0721
14	747333CB1	PITUR0801
15	747377CB1	BRAIN0714
16	137284CB1	SINTN0801
18	747394CB1	BRAIN0709
19	224413CB1	BRAIDR03

WO 01/79468

PCT/US01/11869

Table 6

Library	Vector	Library Description
BRAINOT09	PINCY	Library was constructed using RNA isolated from brain tissue removed from a Caucasian male fetus, who died at 23 weeks' gestation.
BRAINOT14	PINCY	Library was constructed using RNA isolated from brain tissue removed from the left frontal lobe of a 40-year-old Caucasian female during excision of a cerebral meningioma lesion. Pathology for the associated tumor tissue indicated grade 4 gemistocytic astrocytoma.
BRAITDR03	PCDMA2.1	This random primed library was constructed using RNA isolated from allocortex, cingulate posterior tissue removed from a 55-year-old Caucasian female who died from cholangiocarcinoma. Pathology indicated mild meningioma, predominantly over the convexities, scattered axonal spheroids in the white matter of the cingulate cortex and the thalamus, and a few scattered neurofibrillary tangles in the entorhinal cortex and the parietal/occipital gray region. Pathology for the associated tumor tissue indicated well-differentiated cholangiocarcinoma of the liver with residual or relapsed tumor. Patient history included cholangiocarcinoma, post-operative Budd-Chiari syndrome, biliary ascites, hydrothorax, dehydration, malnutrition, oliguria and acute renal failure. Previous surgeries included cholecystectomy and resection of 85% of the liver.
LIVRPT04	PINCY	Library was constructed using RNA isolated from liver tumor tissue removed from a 50-year-old Caucasian male during a partial hepatectomy. Pathology indicated a grade 3-4 hepatoma, forming a mass. Patient history included benign hypertension and hepatitis. Hepatitis B core antigen and hepatitis B surface antigen was present in the patient.
PANENOT21	PINCY	Library was constructed using RNA isolated from pancreatic tissue removed from an 8-year-old Black male who died from anoxia.
SINTKOR01	PCDMA2.1	This random primed library was constructed using RNA isolated from small intestine tissue removed from a 31-year-old Caucasian female during Roux-en-Y gastric bypass. Patient history included clinical obesity.

WO 01/79468

PCT/US01/11869

Table 6 (cont.)

Library	Vector	Library Description
FTUN0001	PICV	This normalized pituitary gland tissue library was constructed from 6.92 million independent clones from a pituitary gland tissue library. Starting RNA was made from pituitary gland tissue removed from a 55-year-old male who died from chronic obstructive pulmonary disease. Neuropathology indicated there were no gross abnormalities, other than mild ventricular enlargement. There was no apparent microscopic abnormality in any of the neocortical areas examined, except for a number of silver positive neurons with apical dendrite staining, particularly in the frontal lobe. The significance of this was undetermined. The only other microscopic abnormality was that there was prominent silver staining with some swollen axons in the CA3 region of the anterior and posterior hippocampus. Microscopic sections of the cerebellum revealed mild Bergmann's gliosis in the Purkinje cell layer. Patient history included schizophrenia. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
SIN2W001	PICV	This large size-fractionated library was constructed using pooled cDNA from two donors. cDNA was generated using mRNA isolated from ileum tissue removed from a 30-year-old Caucasian female (donor A) during partial colectomy, open liver biopsy, and permanent colostomy, and from ileum tissue removed from a 70-year-old Caucasian female (donor B) during right hemicolectomy, open liver biopsy, sigmoidoscopy, colonoscopy, and permanent colostomy. Pathology for the matched tumor tissue (donor A) indicated carcinoid tumor (grade 1 neuroendocrine carcinoma) arising in the terminal ileum. The tumor permeated through the ileal wall into the mesenteric fat and extended into the adherent cecum, where tumor extended through the bowel wall up to the mucosal surface. Multiple lymph nodes were positive for tumor. Additional (2) lymph nodes were also involved by direct tumor extension. Pathology for donor B indicated a non-tumorous margin of ileum. Pathology for the matched tumor (donor B) indicated invasive grade 2 adenocarcinoma forming an ulcerated mass, situated distal to the ileocecal valve. The tumor invaded through the muscularis propria just into the serosal adipose tissue. One regional lymph node was positive for a microfocus of metastatic adenocarcinoma. Donor A presented with flushing and unspecified abdominal/pelvic symptoms. Patient history included endometriosis, and tobacco and alcohol abuse. Donor B's history included a malignant breast neoplasm, type II diabetes, hyperlipidemia, viral hepatitis, an unspecified thyroid disorder, osteoarthritis, and a malignant skin neoplasm. Donor B's medication included tamoxifen.

WO 01/79468

PCT/US01/11869

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABIFARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Perkin Elmer, Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA compares at least five functions: fasta, fasta, fasta, fasta, and search.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.0E-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 sequences: fasta score=100 or greater
BLIMPS	A BLOCKS IMPROVED Searcher that matches a sequence against those in BLOCKS, PRINTS, DDB, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Atwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1998) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Naustell, Cambridge Univ. Press, pp. 1-350.	PFAM hits: Probability value= 1.0E-3 or less Signal peptide hits: Score= 0 or greater

WO 01/79468

PCT/US01/11869

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.	Gribisov, M. et al. (1988) CABIOS 4:61-66; Gribisov, M. et al. (1989) Methods Enzymol. 183:146-159; Bauroch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality scores GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phila Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 36 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Clavette, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TM/AP	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.	Sonhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bauroch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

WO 01/79468

PCT/US01/11869

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:1-10,
 - b) a naturally occurring polypeptide comprising an amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-10,
 - c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, and
 - 10 d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10.
2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-10.
- 15 3. An isolated polynucleotide encoding a polypeptide of claim 1.
4. An isolated polynucleotide encoding a polypeptide of claim 2.
- 20 5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:11-20.
6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.
- 25 7. A cell transformed with a recombinant polynucleotide of claim 6.
8. A transgenic organism comprising a recombinant polynucleotide of claim 6.
- 30 9. A method for producing a polypeptide of claim 1, the method comprising:
 - a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

WO 01/79468

PCT/US01/11869

b) recovering the polypeptide so expressed.

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

5 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:11-20,
- b) a naturally occurring polynucleotide comprising a polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:11-20,
- 10 c) a polynucleotide complementary to a polynucleotide of a),
- d) a polynucleotide complementary to a polynucleotide of b), and
- e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a
15 polynucleotide of claim 11.

13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides
20 comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and
- b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

25 14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- 30 a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
- b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

WO 01/79468

PCT/US01/11869

16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence
5 selected from the group consisting of SEQ ID NO:1-10.

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 for 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.

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 for 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.

WO 01/79468

PCT/US01/11869

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:
- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
 - 5 b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.
26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:
- 10 a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
 - b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
 - c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound
 - 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 for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:
- 20 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
 - 25 c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.
28. A method for assessing toxicity of a test compound, said method comprising:
- a) treating a biological sample containing nucleic acids with the test compound;
 - 30 b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;
 - 35 c) quantifying the amount of hybridization complex; and

WO 01/79468

PCT/US01/11869

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 comprising the steps of:

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 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from

WO 01/79468

PCT/US01/11869

the group consisting of SEQ ID NO:1-10, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

- b) isolating antibodies from said animal; and
- c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10.

36. An antibody produced by a method of claim 35.

37. A composition comprising the antibody of claim 36 and a suitable carrier.

38. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising:

- a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10, or an immunogenic fragment thereof, under conditions to elicit an antibody response;
- b) isolating antibody producing cells from the animal;
- c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;
- d) culturing the hybridoma cells; and
- e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10.

39. A monoclonal antibody produced by a method of claim 38.

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.

42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant immunoglobulin library.

43. A method for detecting a polypeptide having an amino acid sequence selected from the

WO 01/79468

PCT/US01/11869

group consisting of SEQ ID NO:1-10 in a sample, comprising the steps of:

- 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
- 5 polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10 in the sample.

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10 from a sample, the method comprising:

- 10 a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and
- b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-10.

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.

20 47. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

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 01/79468

PCT/US01/11869

55. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:11.

56. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
5 NO:12.

57. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:13.

58. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
10 NO:14.

59. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:15.

60. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
15 NO:16.

61. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
20 NO:17.

62. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:18.

63. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
25 NO:19.

64. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID
NO:20.

WO 01/79468

1

PCT/US01/11869

<110> INCYTE GENOMICS, INC.
 POLICKY, Jennifer L.
 HAFALIA, April
 BURFORD, Neil
 RING, Huijun Z.
 LAL, Preeti
 TRIBOULEY, Catherine M.
 YAO, Monique G.
 YUE, Henry
 TANG, Y. Tom
 PATTERSON, Chandra
 DAS, Debopriya
 SANJANWALA, Madhu S.
 GANDHI, Ameena R.
 REDDY, Roopa
 KHM, Farrah A.
 BAUGHN, Mariah R.
 RAMKUMAR, Jaya
 GRIFFIN, Jennifer A.
 AU-YOUNG, Janice

<120> DRUG METABOLIZING ENZYMES
 <130> PI-0070 PCT
 <140> To Be Assigned
 <141> Herewith
 <150> 60/197,590; 60/198,403; 60/200,185; 60/202,234; 60/203,509
 <151> 2000-04-13; 2000-04-19; 2000-04-28; 2000-05-05; 2000-05-11
 <160> 20
 <170> PERL Program
 <210> 1
 <211> 527
 <212> PRT
 <213> Homo sapiens
 <220>
 <221> misc.feature
 <223> Incyte ID No: 2434655CD1
 <400> 1
 Met Arg Ser Asp Lys Ser Ala Leu Val Phe Leu Leu Leu Gln Leu
 1 5 10 15
 Phe Cys Val Gly Cys Gly Phe Cys Gly Lys Val Leu Val Trp Pro
 20 25 30
 Cys Asp Met Ser His Trp Leu Asn Val Lys Val Ile Leu Glu Glu
 35 40 45
 Leu Ile Val Arg Gly His Glu Val Thr Val Leu Thr His Ser Lys
 50 55 60
 Pro Ser Leu Ile Asp Tyr Arg Lys Pro Ser Ala Leu Lys Phe Glu
 65 70 75
 Val Val His Met Pro Gln Asp Arg Thr Glu Glu Asn Glu Ile Phe
 80 85 90
 Val Asp Leu Ala Leu Asn Val Leu Pro Gly Leu Ser Thr Trp Gln
 95 100 105
 Ser Val Ile Lys Leu Asn Asp Phe Phe Val Glu Ile Arg Gly Thr
 110 115 120
 Leu Lys Met Met Cys Glu Ser Phe Ile Tyr Asn Gln Thr Leu Met
 125 130 135

WO 01/79468

2

PCT/US01/11869

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

<210> 2

<211> 523

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2516747CD1

WO 01/79468

3

PCT/US01/11869

```

<400> 2
Met Val Gly Gln Arg Val Leu Leu Leu Val Ala Phe Leu Leu Ser
1      5      10      15
Gly Val Leu Leu Ser Glu Ala Ala Lys Ile Leu Thr Ile Ser Thr
20     25     30
Leu Gly Gly Ser His Tyr Leu Leu Leu Asp Arg Val Ser Gln Ile
35     40     45
Leu Gln Glu His Gly His Asn Val Thr Met Leu His Gln Ser Gly
50     55     60
Lys Phe Leu Ile Pro Asp Ile Lys Glu Glu Glu Lys Ser Tyr Gln
65     70     75
Val Ile Arg Trp Phe Ser Pro Glu Asp His Gln Lys Arg Ile Lys
80     85     90
Lys His Phe Asp Ser Tyr Ile Glu Thr Ala Leu Asp Gly Arg Lys
95     100    105
Glu Ser Glu Ala Leu Val Lys Leu Met Glu Ile Phe Gly Thr Gln
110    115    120
Cys Ser Tyr Leu Leu Ser Arg Lys Asp Ile Met Asp Ser Leu Lys
125    130    135
Asn Glu Asn Tyr Asp Leu Val Phe Val Glu Ala Phe Asp Phe Cys
140    145    150
Ser Phe Leu Ile Ala Glu Lys Leu Val Lys Pro Phe Val Ala Ile
155    160    165
Leu Pro Thr Thr Phe Gly Ser Leu Asp Phe Gly Leu Pro Ser Pro
170    175    180
Leu Ser Tyr Val Pro Val Phe Pro Ser Leu Leu Thr Asp His Met
185    190    195
Asp Phe Trp Gly Arg Val Lys Asn Phe Leu Met Phe Phe Ser Phe
200    205    210
Ser Arg Ser Gln Trp Asp Met Gln Ser Thr Phe Asp Asn Thr Ile
215    220    225
Lys Glu His Phe Pro Glu Gly Ser Arg Pro Val Leu Ser His Leu
230    235    240
Leu Leu Lys Ala Glu Leu Trp Phe Val Asn Ser Asp Phe Ala Phe
245    250    255
Asp Phe Ala Arg Pro Leu Leu Pro Asn Thr Val Tyr Ile Gly Gly
260    265    270
Leu Met Glu Lys Pro Ile Lys Pro Val Pro Gln Asp Leu Asp Asn
275    280    285
Phe Ile Ala Asn Phe Gly Asp Ala Gly Phe Val Leu Val Ala Phe
290    295    300
Gly Ser Met Leu Asn Thr His Gln Ser Gln Glu Val Leu Lys Lys
305    310    315
Met His Asn Ala Phe Ala His Leu Pro Gln Gly Val Ile Trp Thr
320    325    330
Cys Gln Ser Ser His Trp Pro Arg Asp Val His Leu Ala Thr Asn
335    340    345
Val Lys Ile Val Asp Trp Leu Pro Arg Ser Asp Leu Leu Ala His
350    355    360
Pro Ser Ile Arg Leu Phe Val Thr His Gly Gly Gln Asn Ser Val
365    370    375
Met Glu Ala Ile Arg His Gly Val Pro Met Val Gly Leu Pro Val
380    385    390
Asn Gly Asp Gln His Gly Asn Met Val Arg Val Val Ala Lys Asn
395    400    405
Tyr Gly Val Ser Ile Arg Leu Asn Gln Val Thr Ala Asp Thr Leu
410    415    420
Thr Leu Thr Met Lys Gln Val Ile Glu Asp Lys Arg Tyr Lys Ser
425    430    435
Ala Val Val Ala Ala Ser Val Ile Leu His Ser Gln Pro Leu Ser
440    445    450
Pro Ala Gln Arg Leu Val Gly Trp Ile Asp His Ile Leu Gln Thr
455    460

```

WO 01/79468

PCT/US01/11869

4

Gly Gly Ala Thr His Leu Lys Pro Tyr Ala Phe Gln Gln Pro Trp
 470 475 480
 His Glu Gln Tyr Leu Ile Asp Val Phe Val Phe Leu Leu Gly Leu
 485 490 495
 Thr Leu Gly Thr Met Trp Leu Cys Gly Lys Leu Leu Gly Val Val
 500 505 510
 Ala Arg Trp Leu Arg Gly Ala Arg Lys Val Lys Lys Thr
 515 520

<210> 3
 <211> 358
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7472775CD1

<400> 3
 Met Pro Glu Asp Val Arg Glu Lys Lys Glu Asn Leu Leu Leu Asn
 1 5 10 15
 Ser Glu Arg Ser Thr Arg Leu Leu Thr Lys Thr Ser His Ser Gln
 20 25 30
 Gly Gly Asp Gln Ala Leu Ser Lys Ser Thr Gly Ser Pro Thr Glu
 35 40 45
 Lys Leu Ile Glu Lys Arg Gln Gly Ala Lys Thr Val Phe Asn Lys
 50 55 60
 Phe Ser Asn Met Asn Trp Pro Val Asp Ile His Pro Leu Asn Lys
 65 70 75
 Ser Leu Val Lys Asp Asn Lys Trp Lys Lys Thr Glu Glu Thr Gln
 80 85 90
 Glu Lys Arg Arg Ser Phe Leu Gln Glu Phe Cys Lys Lys Tyr Gly
 95 100 105
 Gly Val Ser His His Gln Ser His Leu Phe His Thr Val Ser Arg
 110 115 120
 Ile Tyr Val Glu Asp Lys His Lys Ile Leu Tyr Cys Glu Val Pro
 125 130 135
 Lys Ala Gly Cys Ser Asn Trp Lys Arg Ile Leu Met Val Leu Asn
 140 145 150
 Gly Leu Ala Ser Ser Ala Tyr Asn Ile Ser His Asn Ala Val His
 155 160 165
 Tyr Gly Lys His Leu Lys Lys Leu Asp Ser Phe Asp Leu Lys Gly
 170 175 180
 Ile Tyr Thr Arg Leu Asn Thr Tyr Thr Lys Ala Val Phe Val Arg
 185 190 195
 Asp Pro Met Glu Arg Leu Val Ser Ala Phe Arg Asp Lys Phe Glu
 200 205 210
 His Pro Asn Ser Tyr Tyr His Pro Val Phe Gly Lys Ala Ile Ile
 215 220 225
 Lys Lys Tyr Arg Pro Asn Ala Cys Glu Glu Ala Leu Ile Asn Gly
 230 235 240
 Ser Gly Val Lys Phe Lys Glu Phe Ile His Tyr Leu Leu Asp Ser
 245 250 255
 His Arg Pro Val Gly Met Asp Ile His Trp Glu Lys Val Ser Lys
 260 265 270
 Leu Cys Tyr Pro Cys Leu Ile Asn Tyr Asp Phe Val Gly Lys Phe
 275 280 285
 Glu Thr Leu Glu Glu Asp Ala Asn Tyr Phe Leu Gln Met Ile Gly
 290 295 300
 Ala Pro Lys Glu Leu Lys Phe Pro Asn Phe Lys Asp Arg His Ser
 305 310 315
 Ser Asp Glu Arg Thr Asn Ala Gln Val Val Arg Gln Tyr Leu Lys
 320 325 330

WO 01/79468

5

PCT/US01/11869

Asp Leu Thr Arg Thr Glu Arg Gln Leu Ile Tyr Asp Phe Tyr Tyr
 335 340 345
 Leu Asp Tyr Leu Met Phe Asn Tyr Thr Thr Pro Phe Leu
 350 355

<210> 4

<211> 396

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7473323CD1

<400> 4

Met Ala Ile Asp Ala Leu Val Ser Leu Cys Leu Pro Glu Val Ile
 1 5 10 15
 Arg Ile Lys Phe Asn Ile Arg Pro Arg Gln Pro His His Asp Leu
 20 25 30
 Pro Pro Gly Gly Ser Gln Asp Gly Asp Leu Lys Glu Pro Thr Glu
 35 40 45
 Arg Val Thr Arg Asp Leu Ser Ser Gly Ala Pro Arg Gly Arg Asn
 50 55 60
 Leu Pro Ala Pro Asp Gln Pro Gln Pro Pro Leu Gln Arg Gly Thr
 65 70 75
 Arg Leu Arg Leu Arg Gln Arg Arg Arg Arg Leu Ile Lys Lys
 80 85 90
 Met Pro Ala Ala Ala Thr Ile Pro Ala Asn Ser Ser Asp Ala Pro
 95 100 105
 Phe Ile Arg Pro Gly Pro Gly Thr Leu Asp Gly Arg Trp Val Ser
 110 115 120
 Leu His Arg Ser Gln Gln Glu Arg Lys Arg Val Met Gln Glu Ala
 125 130 135
 Cys Ala Lys Tyr Arg Ala Ser Ser Ser Arg Arg Ala Val Thr Pro
 140 145 150
 Arg His Val Ser Arg Ile Phe Val Glu Asp Arg His Arg Val Leu
 155 160 165
 Tyr Cys Glu Val Pro Lys Ala Gly Cys Ser Asn Trp Lys Arg Val
 170 175 180
 Leu Met Val Leu Ala Gly Leu Ala Ser Ser Thr Ala Asp Ile Gln
 185 190 195
 His Asn Thr Val His Tyr Gly Ser Ala Leu Lys Arg Leu Asp Thr
 200 205 210
 Phe Asp Arg Gln Gly Ile Leu His Arg Leu Ser Thr Tyr Thr Lys
 215 220 225
 Met Leu Phe Val Arg Glu Pro Phe Glu Arg Leu Val Ser Ala Phe
 230 235 240
 Arg Asp Lys Phe Glu His Pro Asn Ser Tyr Tyr His Pro Val Phe
 245 250 255
 Gly Lys Ala Ile Leu Ala Arg Tyr Arg Ala Asn Ala Ser Arg Glu
 260 265 270
 Ala Leu Arg Thr Gly Ser Gly Val Arg Phe Pro Glu Phe Val Gln
 275 280 285
 Tyr Leu Leu Asp Val His Arg Pro Val Gly Met Asp Ile His Trp
 290 295 300
 Asp His Val Ser Arg Leu Cys Ser Pro Cys Leu Ile Asp Tyr Asp
 305 310 315
 Phe Val Gly Lys Phe Glu Ser Met Glu Asp Asp Ala Asn Phe Phe
 320 325 330
 Leu Ser Leu Ile Arg Ala Pro Arg Asn Leu Thr Phe Pro Arg Phe
 335 340 345
 Lys Asp Arg His Ser Gln Glu Ala Arg Thr Thr Ala Arg Ile Ala
 350 355 360

WO 01/79468

6

PCT/US01/11869

His Gln Tyr Phe Ala Gln Leu Ser Ala Leu Gln Arg Gln Arg Thr
 365 370 375
 Tyr Asp Phe Tyr Tyr Met Asp Tyr Leu Met Phe Asn Tyr Ser Lys
 380 385 390
 Pro Phe Ala Asp Leu Tyr
 395

<210> 5

<211> 395

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7472777CD1

<400> 5

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

WO 01/79468

PCT/US01/11869

7

Gln Glu Leu Cys Ala Gly Ala Leu Gln Leu Leu Gly Tyr Arg Pro
 350 355 360
 Val Tyr Ser Glu Asp Glu Gln Arg Asn Leu Ala Leu Asp Leu Val
 365 370 375
 Leu Pro Arg Gly Leu Asn Gly Phe Thr Trp Ala Ser Ser Thr Ala
 380 385 390
 Ser His Pro Arg Asn
 395

<210> 6

<211> 504

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1272843CD1

<400> 6

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

WO 01/79468

PCT/US01/11869

8

Glu Leu Asn Arg Glu Leu Gly Ala Gly Gln Ala Pro Ser Leu Gly
 335 340 345
 Asp Arg Thr Arg Leu Pro Tyr Thr Asp Ala Val Leu His Glu Ala
 350 355 360
 Gln Arg Leu Leu Ala Leu Val Pro Met Gly Ile Pro Arg Thr Leu
 365 370 375
 Met Arg Thr Thr Arg Phe Arg Gly Tyr Thr Leu Pro Gln Gly Thr
 380 385 390
 Glu Val Phe Pro Leu Leu Gly Ser Ile Leu His Asp Pro Asn Ile
 395 400 405
 Phe Lys His Pro Glu Glu Phe Asn Pro Asp Arg Phe Leu Asp Ala
 410 415 420
 Asp Gly Arg Phe Arg Lys His Glu Ala Phe Leu Pro Phe Ser Leu
 425 430 435
 Gly Lys Arg Val Cys Leu Gly Glu Gly Leu Ala Lys Ala Glu Leu
 440 445 450
 Phe Leu Phe Phe Thr Thr Ile Leu Gln Ala Phe Ser Leu Glu Ser
 455 460 465
 Pro Cys Pro Pro Asp Thr Leu Ser Leu Lys Pro Thr Val Ser Gly
 470 475 480
 Leu Phe Asn Ile Pro Pro Ala Phe Gln Leu Gln Val Arg Pro Thr
 485 490 495
 Asp Leu His Ser Thr Thr Gln Thr Arg
 500

<210> 7
 <211> 229
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7472790CD1

<400> 7
 Met Asn Ile Arg Asn Ala Gln Pro Asp Asp Leu Met Asn Met Gln
 1 5 10 15
 His Cys Asn Leu Leu Cys Leu Pro Glu Asn Tyr Gln Met Lys Tyr
 20 25 30
 Tyr Leu Tyr His Gly Leu Ser Trp Pro Gln Leu Ser Tyr Ile Ala
 35 40 45
 Glu Asp Glu Asp Gly Lys Ile Val Gly Tyr Val Leu Ala Lys Met
 50 55 60
 Glu Glu Glu Pro Asp Asp Val Pro His Gly His Ile Thr Ser Leu
 65 70 75
 Ala Val Lys Arg Ser His Arg Arg Leu Gly Leu Ala Gln Lys Leu
 80 85 90
 Met Asp Gln Ala Ser Arg Ala Met Ile Glu Asn Phe Asn Ala Lys
 95 100 105
 Tyr Val Ser Leu His Val Arg Lys Ser Asn Arg Pro Ala Leu His
 110 115 120
 Leu Tyr Ser Asn Thr Leu Asn Phe Gln Ile Ser Glu Val Glu Pro
 125 130 135
 Lys Tyr Tyr Ala Asp Gly Glu Asp Ala Tyr Ala Met Lys Arg Asp
 140 145 150
 Leu Ser Gln Met Ala Asp Glu Leu Arg Arg Gln Met Asp Leu Lys
 155 160 165
 Lys Gly Gly Tyr Val Val Leu Gly Ser Arg Glu Asn Gln Glu Thr
 170 175 180
 Gln Gly Ser Thr Leu Ser Asp Ser Glu Glu Ala Cys Gln Gln Lys
 185 190 195
 Asn Pro Ala Thr Glu Glu Ser Gly Ser Asp Ser Lys Glu Pro Lys
 200 205 210

WO 01/79468

PCT/US01/11869

9

Glu Ser Val Glu Ser Thr Asn Val Gln Asp Ser Ser Glu Ser Ser
 215 220 225
 Asp Ser Thr Ser

<210> 8
 <211> 347
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7473944CD1

<400> 8
 Met Lys Pro Ala Leu Leu Glu Val Met Arg Met Asn Arg Ile Cys
 1 5 10 15
 Arg Met Val Leu Ala Thr Cys Leu Gly Ser Phe Ile Leu Val Ile
 20 25 30
 Phe Tyr Phe Gln Ile Met Arg Arg Asn Pro Phe Gly Val Asp Ile
 35 40 45
 Cys Cys Arg Lys Gly Ser Arg Ser Pro Leu Gln Glu Leu Tyr Asn
 50 55 60
 Pro Ile Gln Leu Glu Leu Ser Asn Thr Ala Val Leu His Gln Met
 65 70 75
 Arg Arg Asp Gln Val Thr Asp Thr Cys Arg Ala Asn Ser Ala Thr
 80 85 90
 Ser Arg Lys Arg Arg Val Leu Thr Pro Asn Asp Leu Lys His Leu
 95 100 105
 Val Val Asp Glu Asp His Glu Leu Ile Tyr Cys Tyr Val Pro Lys
 110 115 120
 Val Ala Cys Thr Asn Trp Lys Arg Leu Met Met Val Leu Thr Gly
 125 130 135
 Arg Gly Lys Tyr Ser Asp Pro Met Glu Ile Pro Ala Asn Glu Ala
 140 145 150
 His Val Ser Ala Asn Leu Lys Thr Leu Asn Gln Tyr Ser Ile Pro
 155 160 165
 Glu Ile Asn His Arg Leu Lys Ser Tyr Met Lys Phe Leu Phe Val
 170 175 180
 Arg Glu Pro Phe Glu Arg Leu Val Ser Ala Tyr Arg Asn Lys Phe
 185 190 195
 Thr Gln Lys Tyr Asn Ile Ser Phe His Lys Arg Tyr Gly Thr Lys
 200 205 210
 Ile Ile Lys Arg Gln Arg Lys Asn Ala Thr Gln Glu Ala Leu Arg
 215 220 225
 Lys Gly Asp Asp Val Lys Phe Glu Glu Phe Val Ala Tyr Leu Ile
 230 235 240
 Asp Pro His Thr Gln Arg Glu Glu Pro Phe Asn Glu His Trp Gln
 245 250 255
 Thr Val Tyr Ser Leu Cys His Pro Cys His Ile His Tyr Asp Leu
 260 265 270
 Val Gly Lys Tyr Glu Thr Leu Glu Glu Asp Ser Asn Tyr Val Leu
 275 280 285
 Gln Leu Ala Gly Val Gly Ser Tyr Leu Lys Phe Pro Thr Tyr Ala
 290 295 300
 Lys Ser Thr Arg Thr Thr Asp Glu Met Thr Thr Glu Phe Phe Gln
 305 310 315
 Asn Ile Ser Ser Glu His Gln Thr Gln Leu Tyr Glu Val Tyr Lys
 320 325 330
 Leu Asp Phe Leu Met Phe Asn Tyr Ser Val Pro Ser Tyr Leu Lys
 335 340 345
 Leu Glu

WO 01/79468

10

PCT/US01/11869

```

<210> 9
<211> 218
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2244136CD1

<400> 9
Met Thr Pro Ala Pro Pro Pro Gly Ala Arg Pro Gly Ala Ala Ser
1      5      10      15
Leu Ala Gly Phe Ala Gly Val Ala Ser Leu Gly Pro Gly Asp Pro
20     25     30
Arg Arg Ala Ala Asp Pro Arg Pro Leu Pro Pro Ala Leu Cys Phe
35     40     45
Ala Val Ser Arg Ser Leu Leu Leu Thr Cys Leu Val Pro Ala Ala
50     55     60
Leu Leu Gly Leu Arg Tyr Tyr Tyr Ser Arg Lys Val Ile Arg Ala
65     70     75
Tyr Leu Glu Cys Ala Leu His Thr Asp Met Ala Asp Ile Glu Gln
80     85     90
Tyr Tyr Met Lys Pro Pro Gly Ser Cys Phe Trp Val Ala Val Leu
95     100    105
Asp Gly Asn Val Val Gly Ile Val Ala Ala Arg Ala His Glu Glu
110    115    120
Asp Asn Thr Val Glu Leu Leu Arg Met Ser Val Asp Ser Arg Phe
125    130    135
Arg Gly Lys Gly Ile Ala Lys Ala Leu Gly Arg Lys Val Leu Glu
140    145    150
Phe Ala Val Val His Asn Tyr Ser Ala Val Val Leu Gly Thr Thr
155    160    165
Ala Val Lys Val Ala Ala His Lys Leu Tyr Glu Ser Leu Gly Phe
170    175    180
Arg His Met Gly Ala Ser Asp His Tyr Val Leu Pro Gly Met Thr
185    190    195
Leu Ser Leu Ala Glu Arg Leu Phe Phe Gln Val Arg Tyr His Arg
200    205    210
Tyr Arg Leu Gln Leu Arg Glu Glu
215

```

```

<210> 10
<211> 318
<212> PRT
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 7474327CD1

```

```

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

```

WO 01/79468

PCT/US01/11869

11

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

<210> 11

<211> 1636

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2434655CB1

<400> 11

gacacgggtg tgagggaact gccatcatga ggtctgacaa gtcagctltg gtattttctgc 60
 tccctcagct ctctctgttt ggcctctggat tctgtgggaa agtccctgtg tggccctgtg 120
 acatgagcca ttggcttaat gtcaaggcca ttctagaaga gctcatagtg agaggccatg 180
 aggtaacagt attgactcac tcaaaagcctt cgttaattga ctacaggaag cctttctgat 240
 tgaatttga ggtgttccat atgcccengg acagaacaga agaaatgaa atattgtgtg 300
 acctagctct gaatgtcttg ccaggcttat caacctggca atcagttata aaattaaatg 360
 atttttttgt tgaataaaga ggaactttta aatgatgtg tgagagcttt atctacaatc 420
 agacgcttat gaagaagcta caggaaacaa actacgatgt aatgcttata gacctgtga 480
 ttccctgtgg agacctgatg gctgagttgc ttgcagtcctc ttttctctc acacttagaa 540
 ttctgtagg aggaatata gagggaagct gtgggaact tccagctcca ctltctatg 600
 taactgtgac tatgaaggc ctacacagaa gaatgaact tctgggaaga gtaaaaaa 660
 caatgtcttc agttttgttc cacttctgga ttcagabta cgaatataat ttttgggaag 720
 agttttatag taaggcaatb ggaaggccca ctacattatg tgagactgtg ggaagaagctg 780
 agatattggt aatagcaaca tatgtggatt ttgaatttcc tcaaccatac caacctaaat 840
 ttgagtttgt tggaggattg cactgtaaac ctgccaagc tttctctaaag gaaatggaaa 900
 attttgtcca gagttcagg gaagatggta ttgtgtgtgt ttctctgggg tcaactgttc 960
 aaaaatttac agaagaag gctaatatca ttgttccag ccttgcacag atcccacaga 1020
 aggtgttatg gaggtacaaa ggaaaaaaac catccacatt aggagccaat actcggctgt 1080
 atgattggat accccagaat gatcttcttg gtcacccaa aaccaagct tttatcactc 1140
 atggtggat gaatgggac tatgaagcta ttaccatgg ggtccctatg gtgggagttc 1200

WO 01/79468

12

PCT/US01/11869

```

ccatattttg  tgcacagctt  gataacatag  ctccacatgaa  ggccaaagga  gcagctgtag  1260
aaataaactt  caaaactatg  acaagcgaag  atttactgag  ggctttgaga  acagtccatta  1320
cgatttcctc  ttataaagag  aatgctatga  gattatcaag  aattcccat  gatcaacctg  1380
taagccccct  agatcgagca  gtctcttggg  tcyagtttgt  catgcccac  aaaggagcca  1440
agcacttgcg  atcagctgce  ctggaactca  cctggttcca  gcaactactc  atagatgtga  1500
ttgggttctc  gctgactgtc  gtggcaactg  ctctatctct  gtccacaaa  tgttttttat  1560
tttcttgcca  aaatctaat  aaaactagaa  agatagaaaa  gaggaatag  atctttccaa  1620
attcaagaaa  gacctg      1636

```

<210> 12

<211> 2086

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2516747CB1

<400> 12

```

ctacatctctt  cctaggccca  cagccagtc  ctttggagta  ctgaggcgcg  cacagagtcc  60
ttagcccgcg  gcaggcgcg  cagcccgcg  tgagatccgc  tctctctgtg  gaagtggaga  120
tggttgggca  cgggtgctg  cttctagtgg  ccttctctct  ttctgggggc  ctgctctcag  180
aggtgcacca  aatcttgaca  atatctacac  tgggtgggaag  ccattacctc  ctgttggaac  240
gggtgtctca  gattctcaa  gacgatggtc  ataatgtgac  tatgttcat  cagagtggaa  300
agtttttgat  ccagatat  aaagaggagg  aaaaatcata  ccagttatc  aggtgttttt  360
caactgaaga  tcatcaaaaa  agaattaaga  agcattttga  tagctacata  gaaacagcat  420
tggatggcag  aaagaaatct  gaagcccttg  taaagctaat  ggaatatatt  gggactcaat  480
gtagtattct  gctaaagcga  aaggatataa  tggattctct  aaagaatgag  aactatgalc  540
tggattttgt  tgaagcattt  gattctcttt  ctttctcgat  tgcctgagaag  ctgtgaaaaa  600
catttgggc  caattctccc  accacattcg  gctctttgga  ttctgggcta  ccaggccctc  660
tgttttatgt  tccagtatct  cttctcttgc  tgaactgata  calggacttc  tggggcgag  720
tgaagaattt  tctgatgttc  tttagtctct  ccaggagcca  atgggacatg  cagtctacat  780
ttgacaacac  catcaaggag  catttcccag  aagcctctag  gccagttttg  tctcatcttc  840
tactgaagcg  agagctgtgg  ttgttaact  ctgattttgc  ctttattttt  gcccgccccc  900
tgcctcccaa  cactgtttat  attggaggct  tgatggaaaa  acctattaaa  ccagtlaccac  960
aagctctgga  caacttcatt  gccacttttg  gggatgcagg  gtttctctct  gtggcctttg  1020
gtctcatgtt  gaacacccat  cagtcccgag  aagtcctcaa  gaagatgcac  aatgcctttg  1080
cccactctcc  tcaaggagtg  atatggacat  gtccagtltc  tcaattgccc  agagatgttc  1140
atttggccac  aaatgtgaaa  attgtggact  ggtctctctg  gagtgacctc  ctggctcaac  1200
ccagcatccg  tctttctgtc  actcatggtg  ggcagaacag  cgtaatggag  gccatccgtc  1260
atgggtgtgc  catggtggga  ttaccagtca  atgggagcca  gcattggaaa  atagtcagag  1320
tactagccaa  aaattatggt  gtctctatcc  ggttgaaaca  ggtcacagcc  gacacactga  1380
catttaacat  gaacacagtc  atagaagaca  agaggtacac  gtccgagtg  gtggcagcca  1440
gtgtcatctc  gcactctcag  cccctgagcc  ccgcacacgc  cctgtgtggc  tgaatcgacc  1500
aacctctcca  gactggggga  gcgagccacc  tcaagcccta  tgctctccag  cagccttggc  1560
atgagcagta  cctcattgat  gtctttgtgt  ttctgtctgg  gctcactctg  ggcactatgt  1620
ggctttgtgg  gaagctgctg  ggtgtgtgtg  ccagctggct  cgttggggcc  aggaaggtga  1680
agaagacatg  aggttaggtg  tagccttggg  tgaggggagg  gcctccctgg  tctttgaag  1740
gttctctcca  cccagcaca  cgcacccct  ctgtctctct  ttccagctca  cctgcacctg  1800
atctcgcaac  ttgtctcttt  ctattctctg  cctctgttta  gaaatcttca  cacaccactg  1860
aggtctcttg  acttgcacct  tgtgacttga  attccagctc  cagatacaaa  ttttccactg  1920
ccagcctgac  ctctctcttt  ctctcttttc  ctgacacacg  gacttgaca  acttctactc  1980
ccttgtttag  atgacttccc  agtttccagt  cccattttct  cttctatca  cttttctaaa  2040
aaaaactcag  gaaatatttg  acatctcttc  catttcaaat  tcttcc      2086

```

<210> 13

<211> 1814

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7472775CB1

WO 01/79468

13

PCT/US01/11869

```

<400> 13
cttggagcag ggggagagtg gagaagagaa gagaacaaaa agtaacttca ggaatggggac 60
cagtgaagta cttggggcct gtacccagaa tcaaacccca agtttccat gctgaggat 120
gtacgagaaa aaaagggaaa tcttctactc aatttctaga gatctactag gctcttaaca 180
aagaccagtc attcacaagg agggggatcaa gctttaagla agtccacagg gtacccaaaa 240
gagaagttag ttgaaaaaccg tcaaggagct aagaactgtt ttaacaagtt cagcaacatg 300
aattggccag tggacattca ccttttaaac aaaagttag tcaaatgaa taatgggag 360
aaactggagg agacccaaga gaaacgaagg tctttctctc agggatttgg caagaataac 420
gttggggatga gtcatcatca gtacacattt tccatcacag tatccagant ctatgtagaa 480
gataaacaca aaattctata ttgtggagta cctaaaggct gctgttccaa ttggaaaaga 540
attctgatgg tactaaatgg attggttctc tctgcataca acatctccca caatgctgc 600
cactacggga agcattttgan gaagctagat agctttgacc taaaagggat atataccgc 660
ttaaatactt acaccaaaagc tgtgtttgtt cgtgatccca tggaaagatt agtatcagcc 720
tttagggaca aatttgaaca ccccaatagt tattaccatc cagtattcgg aaaggcaatt 780
atcaagaat atcgacaaaa tgcctgtgan gaagcattaa ttaattggatc tggagtcaag 840
ttcaaaagat tbatccacta ctgtctggat tcccaccgtc cagttagaat ggaactaac 900
tgggaaaagg tcagcaaaat ctgctatccg tgttttgaca acctatgatt tggagggaaa 960
tttgacttt tgggaagaga tgcaaatatc ttittacaga tgaatcgpgc tccaaaagg 1020
ctgaatttcc ccaactttaa ggttaggcac tcttcagatg aagaaccaa tgcctcagtc 1080
gtgagacagt atttaaaaga tctgactaga actgagagac aattaatcta tgacttttat 1140
taattggact atttaattgt taattatata actccatttt tgaatttgc attcatttcc 1200
taaaacccctg tatatactta atgatgataa gttcaaatca gctgtaattt ttctataatt 1260
ctctgataga cagaatttta accaagtcca gttgtcttga tttaattgat atttttacca 1320
aatgatatga caccaattgg cacaaagtta taggaaaatc acctacagga gatgtaaaaa 1380
acttggatgg cctcaaaatg tttggaaag agtctgtttt gcaattatga tatatlggt 1440
gaacaataaa cctagccagc tgttgcatta gctaaagcag cctcttgcaa tggtaggaaa 1500
aaaggtcttc aaatagcatg agtctatgtc tatatccctga aattttatgt ctaaaaatgc 1560
tgaatatatt tttagcagtc tgtggcatat taatcaaat gttgaattgt ttctttacac 1620
cctggaaacc ttcttatcaa ctacaatgat aaatccattc tgaattgata ttttgactt 1680
agcatttcta ctttagattg gaaggcatta ttgatattac aatatgaaa tatagcaga 1740
aaaccagatg aggtgtggc tttttatatt caacagccaa taaaaaatgc acaactgct 1800
aagctcaag caaa 1840

```

<210> 14

<211> 1650

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7473323CB1

<400> 14

```

atggccattg atgcactggc ctctctgtgt cttcctgagg tcatcagaat aaagttraac 60
atcaggccaa ggcagcccca ccacgacctc ccaccaggcg gctccaggga tggtagaattg 120
aaggaaccca cagagagggc cactcgggac ttatccagtg gggcccgag gggccgcaac 180
ctgccagcgc ctgaccagcc tcaacccccc ctgacagggg gaaccgctct gggcctccgc 240
cagcggcgcgc gccgtctgct catcaagaaa atgccagctg cggcgaccat ccggcccaac 300
agctcggagc gcccttctat ccggccggga cccgggagcg tggatggcgc ctgggtcagc 360
ctgccccgga gccagcagga gccgaagcgg gtgatgcagg aggcctggcg caagtaccgg 420
gcgagcagca gccgccggcg cgtcacgccc cggccagctg cccgtatctt cgtggaggac 480
cgccaccgcg tgcctactcy cgaagtgccc aaggccgggt gctccaattg gaagcgggtg 540
ctcatggctg tggccggcct ggcctctgtc actgcgcaca tccagcacaa caccgtccac 600
tatggcagcg ctcccaagcg cctggcaacc ttcgaccgcc agggatctct gaccgtctc 660
agaccttaca ccaagatgct cttgtccgc gaggccttc agaggctggg gtcgcgttc 720
cgcgcgaagt ttgagcaacc caacagctac tatcaaccgg tcttcggcaa ggcacactgc 780
gcccggtccc gccccaatgc ctctcgggag cccctgcgga cgggctctgg ggtgcgtttt 840
cccgagttcg tccagtacct gctggagcgt caccggcccg tggggatgga cattcaactg 900
gcccatgtca gccggtcttg cagcccttgc ctcatcgact acgatttctg aggcagttc 960
gagacatgag aggaagatgc caactcttct ctgagcctca tccgcgcgcg gccgaacctg 1020
accttccccc ggttcaagga ccggcactcg caggagcgcg gaccacagc gaggatcgcc 1080
caccagtact tgcgccaaat ctggccctcy caaaggcagc gacccataga ctcttactac 1140
atggattacc cgaatgtcaa ctattccaa ccccttgccg atctgtactg aggggcgcgc 1200

```

WO 01/79468

14

PCT/US01/11869

```

cagctggccc gggccgccc gcccgggcca ctccactgtg ctccggggca tccctctgtc 1360
cctgggctcct catctcgaga gcaacagggc tctggaggag tggggaggca tggctgtggg 1370
aggcagcagg cctcggttgg ggggcagagg cgtccagcct tggatgvyva ccccgagccc 1380
tggctgtgac ctgtttcttc attcctctggc tggaggagag gctggagact gggcagacac 1440
cctggagact cagcggacag ttttgatgag cagggaagtc tgaggccag aggacggggg 1500
gcccagcgtt aaggatgtc cgcactccc ttagccattg ccttggacca aacacagctg 1560
tttgacgctt ttctacgagc caggggggag gtteccctgg attaaggctc caataaagc 1620
actgtgttcc cagacacaaa aaaaacaaaa 1650

```

<210> 15

<211> 1647

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7472777CBI

<400> 15

```

atgtggtctg cggcggtctc cagcacagca gtgaccgggc tctctctggc gcagaccttc 60
ctccccctct ttctgggttc cggccagggg cctctgtccc cagcaggcgg cagggcgggc 120
gtgatgtctc tgggtgtgtc ctgctggcgc tgggctctgt cctctgtggg ccaactcttc 180
aacacagacc cggagctctt ctactaatg gagcccgctt ggcactgttg gaccaccttg 240
tcgcaggcca cggcgccaac gctgcacatg gctgtggcgc acctgtgtgc ctccgtcttc 300
ctgtcgacca tggacgtgtt tgaatgctat ctgcttggc gcccaacct gtccgacctc 360
ttccagtggg ccgtgagccg tgcactgtgc tggccaccgc cctgcagtgc ctttccccga 420
ggcgccatca gaaagcaggg cgtgtgcaag ccaactgtgc cggggcaylc ctttaccttg 480
gcccaggagg cctgcagctc ctacagccac gtggtgtcca aggaagtggc ctttctcaac 540
ctgcagggtc ttatccgctt gctcagcgac cccgcgtcca acctacgat cgtgcacctg 600
gtgcagcgac cggagcgctg gctgagctcc cgggagcaga cagccaaggg tctggcgctg 660
gcaacaggca tctgtctggg caccaacggc acctgggttg agggcgccc cggcctggcg 720
gtggtcgccg aggtctgccc tagccacgta cgcacgccg aggcgcgcc actcaagcgt 780
ccacctcttc tgcgcggccg ctacccgctg gtgcgcttc aggaacctgg cggggagccg 840
ctggcagaaa tctgtgctct ctacgcttc actgggttca gtctcagcc acagctcgag 900
gcttggtacc ataacatcac ccaaggatct ggaacctgtg cgcgcggcga agccttcagg 960
actctgtccc ggaatgctct caacgtctcc caggcctggc gccatgcgtt gccctttggc 1020
aagatccgcc gcgtcagga actgtgcgtt ggtgcgtgc agctgtctgg ctaccggcct 1080
gtgtactctg aggacagcca ggcgaacctc gcccttgatc tgggtgtgct acgagggcgt 1140
aacggcttca cttgggcata atccaccgcc tgcacccccc gaaatttgtg gaggccacag 1200
ttgtagcagg cgttaggccc gggaggagag tgcattgtgc agggggggtt ggggggacag 1260
ggaagcagg ttctatattt gacccaggag ttgtgagaa cctgctgtct gctcctttgc 1320
ttcgaccttc cgcctctgce cgggagaaa cccaggccag cctgctggac aagcagagac 1380
catgagagg agagttcagg ggtcccaaac cagcccatcc tagaccagcc agctccagct 1440
gatccgcaag cagccacttc ggtcaccttc tactggccaa agggagtcct agggctcacc 1500
cagattcaga ggtggggaaa ctgagtcac cacttgagaa gagttagtat aaagacat 1560
gagcagggcc agctgagccc agcactgccc ccaagtcgaa gctttaggag caataaaagt 1620
gcttatgttg ttctagtcaa aaaaaaaa 1647

```

<210> 16

<211> 2620

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1272843CBI

<400> 16

```

cgcggagacc tgggagagga gaaggagccg acctgcccag atggaggcga ccggcacctg 60
ggcgctgctg ctggcgctgg cgtctctcct gctgctyag ctggcgctgt ccgggaccag 120
ggcccgagcc caactgcccc cggggccccc gccgtacca ctgctgggaa acctcttcca 180
gctacggccc gggcgctctt attcaggctt catcggtctt agtaagaggt acggacoggt 240
gttccacctc tacctgggac cctggcgccc tgtggtgttc ctggttgggc aggaagcgtt 300

```

WO 01/79468

15

PCT/US01/11869

```

gggggaggcc ctgggaggtc agggctgagga gttcagcggc gggggaaccg tagcgatgct 360
ggaggggact tttgatggcc atgggggtttt ctttccaaac ggggaggggt ggaggcagct 420
gaggaagtgtt accatgcttg ctctcgccgga cctgggcatg ggggagcgag aaggcgagga 480
gttgatccag gggagggccc ggtgtctggt agagacatc caggggacag aaggagccc 540
attgatccc tccctgtgct tggcccaagg cactccaaac gtagtctgct cctctcctt 600
tggcctccgc ttctctctatg aggataaagg gttccaggcc gtggtccggg cagctgggtg 660
tacctgtgct ggagtcagct cccagggggg tcagacctac gagatgttct cctggttcc 720
ggggccctgt ccaggccccc accagcagct cctccaccac ctgagcactt tggctgctt 780
cacactccgg caggtgcagc agccaccagg gaacctggat gcttcgggcc ccgacgtga 840
ccttgtcgat gccctcctgc tgaagatggc acaggagga caaaaaccag gcacagatt 900
caccaacaag aacatgctga tgacagtca tlatltgctg tttgctggga cgtatgaggt 960
cagcaccacg gtccgctata cctcctctgt cctgatgaaa taccctcatg tccaaaaagt 1020
ggtacgtgag gagctgaatc gggagctggg ggcctggccag gcaccaagcc taggggacg 1080
taccgctcct ccttaccagc acgctgttct gcatgaggcg cagcgctgct tggcgctggt 1140
ggcctggga ataccocgca cctcctgctg gacccaccgc ttccgagggt acacctgccc 1200
ccagggcagc gaggctctcc cctcctctgg cctccatcctg catgacccca acatcttcaa 1260
gcaaccagaa gagttcaacc cagaccgttt cctggatgca gatggacggt tcaaggagca 1320
tgaagcttc ctgacctct ctttagggaa ggtgtctgct ctggagaggy gccctggcaa 1380
agcggagctc ttctcttctt tcccccact cctcaaggcc ttctcctgtg agagccctgt 1440
ccgcctggac accctgagcc tcaagccacc cgtcagtggc cttttcaaca ttcccaccg 1500
cttcagctg caagtcctgc ccactgaact tcactccacc acgcagacca gatgaaggaa 1560
ggcactctgg aagtgggtgg tgcacaggac ggtgcctcaa gctcaacag tggcgatgga 1620
caggggttaat gcttccagag tgtacactgc aggcagccac atttacaccg ctgcagtgt 1680
tttcggaggt ctgtcccaag gcccaacgct tcaactgaat catgctgcta agatgcacaa 1740
cgcacacccc atacacaact acaaggccca caaagcaact gctgggttag ctttccacag 1800
acataaatat agtccatctg caatcacaa cactatgcaa ggtaacccac caactccct 1860
ggatctcag cccacacgtg ggagctctgc tgtcaacctc caaagccaca gaaacggcca 1920
cacatgttca cagctcaaac gccctcttca ttcatgaaac ttctcagtt cctgttcc 1980
ggtcctctgc acagggaaca gcctgcctcc tccgggttca tgcacacag agactgtgc 2040
tgctatggc cccacacat gctcctctc ttggttacc cactctcca gctctgacc 2100
accgatgtcc acacaccccc aaccacttgt ccaacaagct acccagctac gacatgctc 2160
tggtccccc gagtatcttc ccactgagac acgcgcctcc caccagggca cagtccccag 2220
ccactctgc aactgcagcc ctacgtcacc ctttttaag caccctgati ctaccaaag 2280
caaacacatc tgggtctgag attatgcaca gagactttg acatacagg accctcagc 2340
cggaggaaca cctgcacac cccacacgt cttatgttaa ccactggaa agcgccct 2400
gctgcctct cacaacaca tacaacctca ctgatata cccctgttc ggcgtcag 2460
tccccactag acccagtgga aggggttaga gaccaagtat gggccagttt ccaattcacc 2520
ctgtcagga gtgagccgga tctgacgttc cttgtgactt aagggtccg cttgggaatt 2580
aaagtgtgt tctggccttt agcctaaaaa aaaaaaaa 2620

```

<210> 17

<211> 690

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7472790C01

<400> 17

```

atgaacatcc gcaacgctca gccagacgac ctgatgaata tgcaacactg caacctcctt 60
tgcttctcty agaactaaca gatgaataac tatattatct atggcctttc ctggcccccag 120
ctttcttata tegtgtggga tgaggacyyy aagattgttg gctatgttct ggccaaaatg 180
gaggaggaa cagatgagct cccgcatggc atataccact cactggcgtt gaaggttca 240
caacggggcc tgggtctggc ccagagctg atggaccag cctccaggc catgatagag 300
aactttaag ccaaataggt gtctctgcat gtcagagaa gtaaccggcc agccttgac 360
ctttattcta accacctcaa ctttcagatt agtgaagtg aaactaant ctatgcyat 420
ggggagagatg cttatgtcat gaagccggat cctctgcaga tggcagatga gctgagacga 480
caaatggacc tgaagaggyy cgggtatgtg gctctgggt ccaggagaa ccaggagacc 540
cagggcagca cactttctga ttcttaagag gctctgcag aaaaagacc ggtaccgaa 600
gaaagtggca gtgacagcaa agaacctaa gagtctgtgg agagaccaa cgtccaggac 660
agctcagaaa gctcggattc caactcctag

```

WO 01/79468

16

PCT/US01/11869

<210> 18
 <211> 1510
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 7473944CB1

<400> 18
 gcggcgcgga agcgactccg atccctccctc tgagccttgc tcagctctgc cccgcgcctc 60
 cgggctcccg gtcgcgcgag cggggtccct gctcctgcgc cccgggcgcg ctccccggac 120
 atcccgctcc cgcgcgcgag gacaaagcca tgaagccagc gctgctggaa gtgatgagga 180
 tgaacagaat ctgcgcgagc gtgctggcca ctgcttggg atcccttato ctggtcatct 240
 lctatttcca aatcatgagg aggaatccct ttggtgtgga catctgctgc cgggaagggt 300
 cccgaagccc cctgcaggaa ctctacaacc caatccagct ggagctctca aacactgctg 360
 tccctgcacca gatgcgcgag gaccagggtga cagacacgtg ccgagccacc agcgcacaaa 420
 gcgtaagcg gagggtgctg acccccaagc acctgaagca ctggtggtg gatgaggacc 480
 acgagctcat ctactgctac gtgcctcagg tgcctcgac caactggagc cggctcatga 540
 tggctccgac cgggcggggg aagtacagcg accccatgga gatcccgccc aacgaggcac 600
 acgtctccgc caactgaag accctgaacc agtatacgat cccagaatc aaccacccct 660
 tgaataagcta catgaagttc ctgtttgtcc gggagccctt cgagaggcta gtgtccgctt 720
 accgaacaaa gttcaccagc aagtacaaca tctccttcca caagcgtac ggcaccaaga 780
 tcatcaaacg ccagcggagc aacgcacccc aggaagccct gcgcaaggg gacgatgtca 840
 aatctcagga gtttgtggcc tatctcatcg acccacacac ccagcggggg gaggctttca 900
 acgaacactg gcaaacgctc tactcactct gccatccctg ccacatccac tatgacctcg 960
 tgggcacgta cgagacactg gaagaggatt ctaattactg cctgcagctg gcaggagtgg 1020
 gcagctacct gaagttcccc acctatgcaa agtctacgag aactactgat gaaatgacca 1080
 cgaatctctt cngaacatc agctcaggag accaaagcca gctgtacgaa gtctacaaac 1140
 tggatttttt aatgttcaat tactcagttg caagctactc gaaattgga taaggggggt 1200
 ggggagaggg agagaatcat gcttttttat ttaagatttt tatttgtaa aagaattata 1260
 tggatatttg gttattttgt aaattaatat ttctttgggg atgattgctg gagcagcata 1320
 gtgaataatc tttaaaatcc ttgtaggga aggacagctg tttttcagg gaaatagga 1380
 tgggtcgtcc ttgctgttag aagtcaatca tgcaacactg tctcaagggt ttctgtgttt 1440
 ctggtgaatt ccatgaattg tgcattccat aaattctaat taattatttt tatagttatt 1500
 taataaaaaa 1510

<210> 19
 <211> 1701
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2244136CB1

<400> 19
 gagctcggct cactatgtac gtgcgcagtg tgctggcaag gggagaagta gaaagcacc 60
 tgagatgacc ccggtctctc caccaggagc ggggcgggag ggggctccc tagcgggctt 120
 cgcgggggtg gcgtctctgg ggcctgggga ccccgccgag gcgctgacc cgcgcctct 180
 gcccccagcg ctgtgcttgc ccgtgagcgc ctgcctgctg ctgagctgac tgggtccccc 240
 cgcgctgctg ggcctgcgct actactacag ccgcaaggta atccgcccct acctggagt 300
 cgccttgcac acggacatgg cggacataga gctgtactac atgaagccgc ccggtctctg 360
 ctctcgggtg ggcctgctgg atggcaagct ggtgggcatt atgctgcac gggcccaaga 420
 gagggaacac ccgctggagc tgctgcggat gctgtggac tcaagtttcc gaggcaagg 480
 catgcacaag gcgctgggac ggaaggtgct ggagtttccc gtggtgcaca actactcgc 540
 gttgggtatg ggcacgacgg ccgtcaaggt gcccccac aaactctacg agtgcctgg 600
 ctctcagcac atggggccca gtgaccacta cgtctgccc ggcctgacc tctcgtggc 660
 tgagcgcctc ttcttcaggg tccgtaccca ccgtacrcg ctgaagctgc gcgaggagt 720
 accgcgcgag ctgcgcggcc cgcgccccc ggcgcctctg ccgcctttgc cgccttgc 780
 gcccccgggc gggcctgctt ttcagacgct cacttggcgt ttgtgttggg ttctcccttt 840
 tcaacatcct gcggttgtct ggcgtgttcc ggggggtgag gggctgttgt ttctgggga 900
 caetttgtgt ggggtgggtt gtttgtgccc atagccccc gcccttccc acctgcctg 960

WO 01/79468

17

PCT/US01/11869

```

gggggttggc aactgaagag tggcatcttg gaccacccgc gctgtccatg acgtgtccct 1020
gcccgcggcc actgggggaag gccacgcctt gctcaccacg cacagaacct ctgcagcaga 1080
tcccggggcc aggtctccggc cccgcctggg gccccagcgc cactgcctgt ggaggcccca 1140
gctggccacg gctgtgcttt gctccgggca tggcgaaggt gtggcccgcc tgagcatggc 1200
gcattgcacac agccctcgcc tcccgccttg cccagactgg acccggaac ccgggtggtg 1260
gaagcgcctt gtcccagcgc cccagctggc tgtgggaagg cctgcacctg cccccccttc 1320
ctggaggggc tggcttgccc cggcccgccc ggtctgttcc acacctgctt tgccttgacc 1380
ccctccattt cctgggtccc ggcctccccc ctgcttgggc tgtgtgactt ggtgtccata 1440
cccaggtggc tcccatggcg tccgtggcac agccagggtg dgggtccatg ggacccctct 1500
cccagtgccc cactggatcg tgcctggccc tcccagatgt ccccggggac cctcgtgccc 1560
tggctgacgg tccacctgtt gaattctbct agcccaggtt gctgcacaca gcgccacagc 1620
cacagcttct cccagcttga aaccacacaa ttttctata agttatttag acagatagc 1680
actctgcatg actttaattc ttgggacaaa acggtagttt gtaccttaag acacagtctc 1740
tggccacgtg tgaagggggt ggttggccgg gtggtggag cgttttgtgt ttggaaacct 1800
ggagggaaac cctgattgga tgtcatttcc tgcctggag caagcctccc agccttgccc 1860
tgcagggtgg gacgggtggg gggcaaggga gtcgcagccc tccgggagga ggggcagggc 1920
gctgccttgg gctgggtggg aagaggggtg gcccctcgg cttccgctgg ccatgctctt 1980
gctctctctt tccggagtc acagggcagg gctgcctcgg accggggggg ggggggttgg 2040
cctgaagagg gaggacaggg tggagggtgg cacaggctgc accctcagct tgaagtgga 2100
cctggttttg gtggcaggag aagaataaac acttgcccag acccctttgt gtgggggaat 2160
tggggagggg tctgtgcagg cggggtgggc cccggaactt ggtccaggc atcaaggcca 2220
cgtgcaggac catggaggga tctctctcac gaggctcttc agaagcgagc gaaggacag 2280
agaagccctg cgtccaggg ccttttgtcc tgttagcaat tgaggtgtgc agagcactgt 2340
acagacccca ccccctgta cattctctcc tggaggtgac cgttcccgcc ttggggatgg 2400
gagttttgta gactgtacag aatctggcac cctattttct tgcagctcag attttgttaa 2460
ctgtgaatat acagacagac gtaaatgttt tttagcaaat ggaacacaa acgttgtgct 2520
tttctctctt ttggttttgt ttgggttttg cctggggctg gtcccagttg gtgcggggca 2580
tctgtggggc agggggggca gggccggcca ggtggagtca ggttttggg gtgtccatgc 2640
ggggtgctgc cagcgtccct tggctctatg cctttgaacc cttgggtccc tggggtgag 2700
gtgtcttggg gatgctgttt gtagccaccc gacgcaggg caggccctgg agcccgagg 2760
ctgtccagtg tctcttgat atlgatctta gtagatcccc cctctggggg ttctgagagt 2820
cctgggcagt gtgcctctct ctctctcttg tggcatctgc gcccttgagt gacctcttcc 2880
ttggtgtcac tgcctgttgg gtgtgtgagc gcttggcctt ttttttgtct gccaggactt 2940
gttagagatg gcaggaaagg tgttcggggt ggttgtgtgc agaaccctgc cgcctccag 3000
gtggtcagag gaactggtgt gcttgcaggg ctggggcgga gctgcccaga acccttgcca 3060
ggccaatgtg tagcttggtg ggcctcttga agccacttga ggggttaggtg tgcctcccc 3120
cggggctgga ggcgggtgct ctgggtgggtg agcctgaact gggccacctc atccctccag 3180
cctgggtatc cactgtcttg cauctgtgca ctgacgtgaa ttttatacgt ttgtagaaac 3240
tctgtatbaa cttcttctac ctttgaagg cctcctggg cctgtgtgtc acgtggctgg 3300
tccgcttccc ctagtgcttg gccctctgccc cgtcccccat gggcaccac cactcaaat 3360
clacgttggt agggccgggg aaagccaggg cgggcccctag cactgtctag gggccctggg 3420
ctcctgggag tccgtgtgt ttgggaagct ctgtttttgg agtggggggc gtggaaggcg 3480
gaggggtctg acagcctctg gggccagagg tgggcgaagg aggaatctct cactggagca 3540
gcactgcagg cctgacgaga ggttcccggc agctagggcc ttaagacctt cgtgatctc 3600
cccatggccc aagcccccgt tctgttttgc tggggtgggt ttgcccgcgc gattgacctg 3660
tcccctgggt ttcttgacc agtctctcgg gggagggcct c 3701

```

<210> 20

<211> 960

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 7474327CB1

<400> 20

```

gcagccatgg ccagccacat tgtgtctaac aatggcacca agatgcccat cctggggcta 60
ggcactctga attccctccc aggcacagta actgaggctg tgaagggtgc cattaatggt 120
gggtactgcc acatcgactg tgcacacgtg taccagatag aagatgactg gggggtggcc 180
attcgggaga agctcaggga gcagggtggtg aagtgtgagg agctcttcat caccagcaag 240
ctgtggtgct cgtaccatga gaagggcctg gtgaaggag cctgcacaga gatgtcatt 300
gacctgaacc tggactacct ggcctctcac cttattgctt ggcccaacag cttcaagcct 360

```

WO 01/79468

18

PCT/US01/11869

```
gggaaggaat ttttccatt ggatgagcca ggtaatggta atgtggttcc cagtaacagt 420
aacattctgg acacatgggc gggcatggaa gagctggtgg atgaagggtt ggtgaaagct 480
attggcatct ccaacttcaa ccatctccag gttgagagga tcttaacaa acctgactta 540
aagtataagc cggtaggttaa tcagattgag tggcaccgtt acctcactca ggaagaagta 600
atccagtact gccagttcaa aggcacatg gtgactgctt acagctcctt cagctccccc 660
gacaggccct gggccaagcc tgaaggacct tccctctctg aagattccag gatcaaggcg 720
atcacggcca aacacataa aactacagcc caggttctga ttggttccc catgcagagg 780
aacttggtgg tgatcccca gtctgtgaca ccagaatcca ttgctgagaa ctttaaggtc 840
ttcaactttg aactgaacag ccaggatatg accccttat tcagttacea cagaaactgg 900
agggctcttg cttgtgtgag ctgtgctctc cacaaggatt accccttcca tgaagagttt 960
```

【国際公開パンフレット（コレクトバージョン）】

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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
25 October 2001 (25.10.2001)

PCT

(10) International Publication Number
WO 01/079468 A3(51) International Patent Classification: C12N 15/52,
9/00, 15/63, 5/10, A01K 67/027, C12P 21/00, C07K
16/40, C12Q 1/68, A61K 38/43, C12Q 1/00, G01N 33/573

(21) International Application Number: PCT/US01/11869

(22) International Filing Date: 12 April 2001 (12.04.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/197,590 13 April 2000 (13.04.2000) US
60/198,103 19 April 2000 (19.04.2000) US
60/200,185 28 April 2000 (28.04.2000) US
60/202,234 5 May 2000 (05.05.2000) US
60/203,509 11 May 2000 (11.05.2000) US(43) Related by continuation (CON) or continuation-in-part
(CIP) to earlier applications:US 60/198,403 (CIP)
Filed on 19 April 2000 (19.04.2000) US
60/197,590 (CIP)
Filed on 13 April 2000 (13.04.2000) US
60/200,185 (CIP)
Filed on 28 April 2000 (28.04.2000) US
60/202,234 (CIP)
Filed on 5 May 2000 (05.05.2000) US
60/203,509 (CIP)
Filed on 11 May 2000 (11.05.2000) 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): POLICKY, Jen-
nifer, L. (US/US); 1511 Jarvis Court, San Jose, CA
95118 (US); HAFALIA, April (US/US); 2227 Calle de
Prinavera, Santa Clara, CA 95051 (US); BURFORD,
Neil (GB/US); 105 Wildwood Circle, Durham, CT 06422
(US); KING, Huijun, Z. (US/US); 625 Orange Avenue,
Los Altos, CA 94022 (US); LAL, Preeti (IN/US); P.O.
Box 5142, Santa Clara, CA 95056 (US); TRIBOULEY,
Catherine, M. (FR/US); 1121 Tennessee Street, #5, San
Francisco, CA 94107 (US); YAO, Monique, G. (US/US);
111 Frederick Court, Mountain View, CA 94043 (US);
YUK, Henry (US/US); 826 Lois Avenue, Sunnyvale, CA94087 (US); TANG, Y., Tom (US/US); 4230 Ranwick
Court, San Jose, CA 95118 (US); PATTERSON, Chandra
(US/US); 490 Sherwood Way #1, Menlo Park, CA
94025 (US); DAS, Dehpriva (IN/US); Apartment 3, 1179
Bonita Avenue, Mountain View, CA 94040 (US); SAN-
JANWALA, Madhu, S. (US/US); 210 Sylvia Court, Los
Altos, CA 94024 (US); GANDHI, Ameeta, R. (US/US);
837 Roble Avenue, #1, Menlo Park, CA 94025 (US);
REDDY, Rupa (US/US); 1233 W. McKinley Avenue, #3,
Sunnyvale, CA 94086 (US); KILIAN, Farrah, A. (IN/US);
333 Escuela Avenue, #221, Mountain View, CA 94040
(US); RAUGHN, Mariah, R. (US/US); 14244 Santiago
Road, San Leandro, CA 94577 (US); RAMKUMAR,
Jaya (IN/US); 34359 Maybird Circle, Fremont, CA 94555
(US); GRIFFIN, Jennifer, A. (US/US); 181 Irene Court,
#11, Belmont, CA 94002 (US); AI-YOUNG, Janice
(US/US); 233 Golden Eagle Lane, Brisbane, CA 94005
(US).(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).(81) Designated States (national): AE, AG, AL, AM, AT, AU,
A7, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, GR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

with international search report
— before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments(88) Date of publication of the international search report:
18 July 2002For two-letter codes and other abbreviations, refer to the "Guide-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning 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, reagents, and antagonists. The invention also provides
methods for diagnosing, treating, or preventing disorders associated with aberrant expression of DME.

WO 01/079468 A3

【国際調査報告】

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER		International Application No. PCT/US 01/11869
IPC 7	C12N15/52 C12N9/00 C12N15/63 C12N5/10 A01K67/027 C12P21/00 C07K16/40 C12Q1/68 A61K38/43 C12Q1/09 G01N33/573	
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		
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)		
SEQUENCE SEARCH, EPO-Internal, WPI Data, PAJ, BIOSIS		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevance to claim No.
X	DATABASE EM HUM [Online] EMBL; 16 January 2000 (2000-01-16) WATERSTON R.H.: "Homo sapiens chromosome 4 clone RP11-468N14, complete sequence" retrieved from EBI, accession no. AC021146 Database accession no. AC021146 XP002186102 the whole document ---	3,12
A	WO 97 44466 A (ENDORECHERCHE INC ; LEVESQUE ERIC (CA); BEAULIEU MARTIN (CA); BELAN) 27 November 1997 (1997-11-27) Seq Id Nos 1,2 abstract; claims page 1, line 20 -page 2, line 4 --- -/--	1-45,55
☒ Further documents are listed in the continuation of box C. ☒ Patent family members are listed in annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of an earlier citation or other special reason (as specified) "O" document relating 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		Date of mailing of the international search report
8 January 2002		07.05.02
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL-6280 HV Rijswijk Tel. (+31-70) 340 2040, Tx. 31 651 490 nl Fax. (+31-70) 340 2018		Authorized officer CEDER O.

Form PCT/ISA/210 (second sheet) (July 1989)

INTERNATIONAL SEARCH REPORT

International Application No.
PCT/US 01/11869

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 0 905 232 A (KIRIN BREWERY) 31 March 1999 (1999-03-31) abstract page 1, line 1 -page 5, line 24 ---	
A	EP 0 463 395 A (HOFFMANN LA ROCHE) 2 January 1992 (1992-01-02) ---	
A	KIMURA S ET AL: "THE HUMAN DEBRISOQUINE 4-HYDROXYLASE (CYP2D) LOCUS: SEQUENCE AND IDENTIFICATION OF THE POLYMORPHIC CYP2D6 GENE, A RELATED GENE, AND A PSEUDOGENE" AMERICAN JOURNAL OF HUMAN GENETICS, UNIVERSITY OF CHICAGO PRESS, CHICAGO,, US, vol. 45, no. 6, 1 December 1989 (1989-12-01), pages 889-904. XP000570035 ISSN: 0002-9297 ---	
P,X	DATABASE EM_HUM [Online] EMBL; 29 September 2000 (2000-09-29) SUGANO ET AL: "Homo sapiens cDNA: FLJ21934 fis, clone HEP04364" retrieved from EBI, accession no. AK025587 Database accession no. AK025587 XP002186103 the whole document -& DATABASE SWALL [Online] 1 March 2001 (2001-03-01) KAWABATA ET AL: "cDNA: FLJ21934 fis, clone HEP04364" retrieved from EBI, accession no. Q9H6S4 Database accession no. Q9H6S4 XP002186764 the whole document -----	3,11,12
P,X		1

INTERNATIONAL SEARCH REPORT	International application No. PCT/US 01/11869
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, 23, 24 because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically: see FURTHER INFORMATION sheet PCT/ISA/210 3. ☐ Claims Nos.: - because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).	
Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)	
This International Searching Authority found multiple inventions in this International Application, as follows: see additional sheet 1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims. 2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee. 3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.: 4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: 1-44 all partially, 45, 55 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 01/11869

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

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

1. Claims: 1-44 all partially and 45 and 55 completely

An isolated polypeptide, an isolated polynucleotide and their uses where the polypeptide and polynucleotide are Seq Id Nos 1 and 11, respectively.

2. Claims: 1-44 all partially and 46 and 56 completely

Idem as for invention 1 above but for Seq Id Nos 2 and 12.

3. Claims: 1-44 all partially and 47 and 57 completely

Idem as for invention 1 above but for Seq Id Nos 3 and 13.

4. Claims: 1-44 all partially and 48 and 58 completely

Idem as for invention 1 above but for Seq Id Nos 4 and 14.

5. Claims: 1-44 all partially and 49 and 59 completely

Idem as for invention 1 above but for Seq Id Nos 5 and 15.

6. Claims: 1-44 all partially and 50 and 60 completely

Idem as for invention 1 above but for Seq Id Nos 6 and 16.

7. Claims: 1-44 all partially and 51 and 61 completely

Idem as for invention 1 above but for Seq Id Nos 7 and 17.

8. Claims: 1-44 all partially and 52 and 62 completely

Idem as for invention 1 above but for Seq Id Nos 8 and 18.

9. Claims: 1-44 all partially and 53 and 63 completely

Idem as for invention 1 above but for Seq Id Nos 9 and 19.

10. Claims: 1-44 all partially and 54 and 64 completely

Idem as for invention 1 above but for Seq Id Nos 10 and 20.

International Application No. PCT/US 01/11869

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 18,21 and 24 are directed to methods of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

Continuation of Box I.1

Rule 39.1(iv) PCT - Method for treatment of the human or animal body by therapy

Continuation of Box I.2

Claims Nos.: 20, 21, 23, 24

Claims 20,21,23,24 refers to agonists/antagonists of the polypeptide of claim 1 identified by the methods of claims 19 and 22 respectively, without giving a true technical characterization. Moreover, no such compounds are specifically defined in the description. It is only indicated that they "may include proteins" (such as antibodies), 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." (page 31 lines 22-25; page 33 lines 23-27). In consequence the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT). No search can be carried out for such claims whose wording is, in fact, a mere recitation of the results to be achieved.

Claim 8 has only been search in as far as it concerns non-human transgenic organisms.

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

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No.

PCT/US 01/11869

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9744466	A	27-11-1997	AU 2687897 A	09-12-1997
			WO 9744466 A1	27-11-1997
			EP 0904382 A1	31-03-1999
			JP 2000511046 T	29-08-2000
			US 6287834 B1	11-09-2001
EP 0905232	A	31-03-1999	US 0905232 A1	31-03-1999
			EP 2001024814 A1	27-09-2001
			CN 1214732 A	21-04-1999
			WO 9826053 A1	18-06-1998
			EP 0463395	A
AU 645091 B2	06-01-1994			
AU 7848791 A	02-01-1992			
CA 2045012 A1	23-12-1991			
JP 5211895 A	24-08-1993			
NZ 238610 A	25-11-1993			
US 5648482 A	15-07-1997			
US 5844108 A	01-12-1998			
DE 69125064 D1	19-06-1997			
DE 69125064 T2	28-08-1997			
ES 2103756 T3	01-10-1997			

フロントページの続き

(51)Int.Cl. ⁷	F I	テーマコード (参考)
A 6 1 P 1/00	A 6 1 P 1/16	4 B 0 6 5
A 6 1 P 1/16	A 6 1 P 5/00	4 C 0 8 4
A 6 1 P 5/00	A 6 1 P 27/02	4 C 0 8 5
A 6 1 P 27/02	A 6 1 P 29/00	4 H 0 4 5
A 6 1 P 29/00	A 6 1 P 35/00	
A 6 1 P 35/00	A 6 1 P 37/06	
A 6 1 P 37/06	C 0 7 K 16/40	
C 0 7 K 16/40	C 0 7 K 19/00	
C 0 7 K 19/00	C 1 2 N 1/15	
C 1 2 N 1/15	C 1 2 N 1/19	
C 1 2 N 1/19	C 1 2 N 1/21	
C 1 2 N 1/21	C 1 2 N 9/00	
C 1 2 N 5/10	C 1 2 P 21/08	
C 1 2 N 9/00	C 1 2 Q 1/00	
C 1 2 N 15/02	C 1 2 Q 1/02	
C 1 2 P 21/08	C 1 2 Q 1/68	A
C 1 2 Q 1/00	G 0 1 N 33/15	Z
C 1 2 Q 1/02	G 0 1 N 33/50	Z
C 1 2 Q 1/68	G 0 1 N 33/53	D
G 0 1 N 33/15	G 0 1 N 33/53	M
G 0 1 N 33/50	G 0 1 N 33/566	
G 0 1 N 33/53	C 1 2 N 15/00	C
G 0 1 N 33/566	C 1 2 N 5/00	A
	A 6 1 K 37/02	

(31)優先権主張番号 60/202,234

(32)優先日 平成12年5月5日(2000.5.5)

(33)優先権主張国 米国(US)

(31)優先権主張番号 60/203,509

(32)優先日 平成12年5月11日(2000.5.11)

(33)優先権主張国 米国(US)

(81)指定国 AP(GH,GM,KE,LS,MW,MZ,SD,SL,SZ,TZ,UG,ZW),EA(AM,AZ,BY,KG,KZ,MD,RU,TJ,TM),EP(AT,BE,CH,CY,DE,DK,ES,FI,FR,GB,GR,IE,IT,LU,MC,NL,PT,SE,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GW,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AT,AU,AZ,BA,BB,BG,BR,BY,BZ,CA,CH,CN,CR,CU,CZ,DE,DK,DM,DZ,EE,ES,FI,GB,GD,GE,GH,GM,HR,HU,ID,IL,IN,IS,JP,KE,KG,KP,KR,KZ,LC,LK,LR,LS,LT,LU,LV,MA,MD,MG,MK,MN,MW,MX,MZ,NO,NZ,PL,PT,RO,RU,SD,SE,S G,SI,SK,SL,TJ,TM,TR,TT,TZ,UA,UG,US,UZ,VN,YU,ZA,ZW

(72)発明者 ハファリア、エープリル

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 バーフォード、ニール

アメリカ合衆国コネチカット州0 6 4 2 2・ダラム・ワイルドウッドサークル 1 0 5

(72)発明者 リング、ヒュイジュン、ジュー

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 ラル、プリーティ

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 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 ヤオ、モニーク・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 ・サニーバイル・ルイスアベニュー 8 2 6
- (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 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 ダス、デボプリバ
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・ボニータアベニュー 1 1 7 9
・アパートメント 3
- (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 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 レディ、ルーパ
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・# 3 ・ウェストマッキンレーアベ
ニュー 1 2 3 3
- (72)発明者 カーン、ファラ・エイ
アメリカ合衆国カリフォルニア州 9 4 0 4 0 ・マウンテンビュー・# 2 2 1 ・エスケラアベニュー
3 3 3
- (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 9 4 5 7 7 ・サンレアンドロ・サンティアゴロード 1 4 2 4 4
- (72)発明者 ランクマール、ジャヤ
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 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 0 0 2 ・ベルモント・# 1 1 ・アイリーンコート 1 8 1
- (72)発明者 オウーヤング、ジャニス
アメリカ合衆国カリフォルニア州 9 4 0 0 5 ・ブリスベーン・ゴールデンイーグルレーン 2 3 3
- F ターム(参考) 2G045 AA40 DA12 DA13 DA14 DA36 FB02 FB03
4B024 AA01 AA11 BA07 BA53 BA61 CA03 CA04 CA07 CA09 CA12
CA20 DA02 DA06 DA12 EA02 EA04 FA02 GA03 GA11 HA03
HA04 HA11 HA13 HA14
4B050 CC01 CC03 CC05 DD11 FF14E LL01 LL03 LL05
4B063 QA01 QA05 QA19 QQ21 QQ41 QQ61 QQ89 QQ95 QR08 QR32
QR35 QR40 QR42 QR56 QR62 QR77 QR80 QR84 QS16 QS25
QS34 QS36 QX01 QX02
4B064 AG27 CA10 CA20 CC01 CC24 CE12 DA01 DA13
4B065 AA26X AA58X AA72X AA90X AA91X AA93Y AB01 AB05 AC14 BA02
BA08 BD14 CA25 CA27 CA44 CA46
4C084 AA02 AA03 AA06 AA07 AA17 BA01 BA08 BA22 BA23 CA18
MA52 MA56 MA57 MA59 MA60 MA63 MA66 NA14 ZA332 ZA662
ZA752 ZB082 ZB112 ZB262 ZC032
4C085 HH03 HH11 KA03 KA05 KA26 LL05 LL15 LL18
4H045 AA11 AA20 AA30 BA41 CA40 DA75 DA76 EA20 EA50 FA71
FA72 FA74 GA26

专利名称(译)	药物代谢酶		
公开(公告)号	JP2004512818A	公开(公告)日	2004-04-30
申请号	JP2001577452	申请日	2001-04-12
[标]申请(专利权)人(译)	洞察Genomics公司		
申请(专利权)人(译)	洞察基因组公司		
[标]发明人	ポリッキージェニファーエル ハファリアエープリル バーフォードニール リングヒュイジュンジー ラルプリーティ トリボレーキャサリーンエム ヤオモニークジー ユエヘンリー タングワイトム パターソンチャンドラ ダスデボプリバ サンジャンワラマデューエス ガンディーアミーナアール レディルーパ カーンファアラエイ ボーグンマライアアール ランクマールジャヤ グリフィンジェニファーエイ オウヤングジャニス		
发明人	ポリッキー、ジェニファー・エル ハファリア、エープリル バーフォード、ニール リング、ヒュイジュン、ジー ラル、プリーティ トリボレー、キャサリーン・エム ヤオ、モニーク・ジー ユエ、ヘンリー タング、ワイトム パターソン、チャンドラ ダス、デボプリバ サンジャンワラ、マデュー・エス ガンディー、アミーナ・アール レディ、ルーパ カーン、ファアラ・エイ ボーグン、マライア・アール ランクマール、ジャヤ グリフィン、ジェニファー・エイ オウ・ヤング、ジャニス		
IPC分类号	A01K67/027 A61K38/00 A61K45/00 A61K49/00 A61P1/00 A61P1/16 A61P5/00 A61P27/02 A61P29/00 A61P35/00 A61P37/06 C07K16/40 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12N5/10 C12N9/00 C12N9/02 C12N9/04 C12N9/10 C12N15/02 C12N15/09 C12P21/08 C12Q1/00 C12Q1/02 C12Q1/68 G01N33/15 G01N33/50 G01N33/53 G01N33/566		

CPC分类号	A01K2217/05 A61K38/00 A61P1/00 A61P1/16 A61P5/00 A61P27/02 A61P29/00 A61P35/00 A61P37/06 C12N9/0006 C12N9/0077 C12N9/1029 C12N9/1051 C12N9/13 C12Y101/01021 C12Y204/01017 C12Y204/01045
FI分类号	C12N15/00.ZNA.A A01K67/027 A61K45/00 A61K49/00.A A61P1/00 A61P1/16 A61P5/00 A61P27/02 A61P29/00 A61P35/00 A61P37/06 C07K16/40 C07K19/00 C12N1/15 C12N1/19 C12N1/21 C12N9/00 C12P21/08 C12Q1/00 C12Q1/02 C12Q1/68.A G01N33/15.Z G01N33/50.Z G01N33/53.D G01N33/53.M G01N33/566 C12N15/00.C C12N5/00.A A61K37/02
F-TERM分类号	2G045/AA40 2G045/DA12 2G045/DA13 2G045/DA14 2G045/DA36 2G045/FB02 2G045/FB03 4B024/AA01 4B024/AA11 4B024/BA07 4B024/BA53 4B024/BA61 4B024/CA03 4B024/CA04 4B024/CA07 4B024/CA09 4B024/CA12 4B024/CA20 4B024/DA02 4B024/DA06 4B024/DA12 4B024/EA02 4B024/EA04 4B024/FA02 4B024/GA03 4B024/GA11 4B024/HA03 4B024/HA04 4B024/HA11 4B024/HA13 4B024/HA14 4B050/CC01 4B050/CC03 4B050/CC05 4B050/DD11 4B050/FF14E 4B050/LL01 4B050/LL03 4B050/LL05 4B063/QA01 4B063/QA05 4B063/QA19 4B063/QQ21 4B063/QQ41 4B063/QQ61 4B063/QQ89 4B063/QQ95 4B063/QR08 4B063/QR32 4B063/QR35 4B063/QR40 4B063/QR42 4B063/QR56 4B063/QR62 4B063/QR77 4B063/QR80 4B063/QR84 4B063/QS16 4B063/QS25 4B063/QS34 4B063/QS36 4B063/QX01 4B063/QX02 4B064/AG27 4B064/CA10 4B064/CA20 4B064/CC01 4B064/CC24 4B064/CE12 4B064/DA01 4B064/DA13 4B065/AA26X 4B065/AA58X 4B065/AA72X 4B065/AA90X 4B065/AA91X 4B065/AA93Y 4B065/AB01 4B065/AB05 4B065/AC14 4B065/BA02 4B065/BA08 4B065/BD14 4B065/CA25 4B065/CA27 4B065/CA44 4B065/CA46 4C084/AA02 4C084/AA03 4C084/AA06 4C084/AA07 4C084/AA17 4C084/BA01 4C084/BA08 4C084/BA22 4C084/BA23 4C084/CA18 4C084/MA52 4C084/MA56 4C084/MA57 4C084/MA59 4C084/MA60 4C084/MA63 4C084/MA66 4C084/NA14 4C084/ZA332 4C084/ZA662 4C084/ZA752 4C084/ZB082 4C084/ZB112 4C084/ZB262 4C084/ZC032 4C085/HH03 4C085/HH11 4C085/KA03 4C085/KA05 4C085/KA26 4C085/LL05 4C085/LL15 4C085/LL18 4H045/AA11 4H045/AA20 4H045/AA30 4H045/BA41 4H045/CA40 4H045/DA75 4H045/DA76 4H045/EA20 4H045/EA50 4H045/FA71 4H045/FA72 4H045/FA74 4H045/GA26
优先权	60/197590 2000-04-13 US 60/198403 2000-04-19 US 60/200185 2000-04-28 US 60/202234 2000-05-05 US 60/203509 2000-05-11 US
外部链接	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。